

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

**ACADEMIC SCIENTIFIC
JOURNAL OF COMPUTER SCIENCE**

**№2
2026**

ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)



ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

2 (358)

APRIL – JUNE 2026

**PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR**

ALMATY, NAS RK

Chief Editor:

MUTANOV Galimkair Mutanovich, doctor of technical sciences, professor, academician of NAS RK, (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=6506682964>, <https://www.webofscience.com/wos/author/record/1423665>

EDITORIAL BOARD:

KALIMOLDAYEV Maksat Nuradilovich, (Deputy Editor-in-Chief), Doctor of Physical and Mathematical Sciences, Professor, Academician of NAS RK, Advisor to the General Director of the Institute of Information and Computing Technologies of the CS MES RK, Head of the Laboratory (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=56153126500>, <https://www.webofscience.com/wos/author/record/2428551>

MAMYRBAEV Orken Zhumazhanovich, (Academic Secretary), PhD in Information Systems, Deputy Director for Science of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=55967630400>, <https://www.webofscience.com/wos/author/record/1774027>

BAIGUNCHEKOV Zhumadil Zhanabaevich, Doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=6506823633>, <https://www.webofscience.com/wos/author/record/1923423>

WOICIK Waldemar, Doctor of Technical Sciences (Phys.-Math.), Professor of the Lublin University of Technology (Lublin, Poland), <https://www.scopus.com/authid/detail.uri?authorId=7005121594>, <https://www.webofscience.com/wos/author/record/678586>

SMOLARJ Andrej, Associate Professor Faculty of Electronics, Lublin polytechnic university (Lublin, Poland), <https://www.scopus.com/authid/detail.uri?authorId=56249263000>, <https://www.webofscience.com/wos/author/record/1268523>

KEILAN Alimkhan, Doctor of Technical Sciences, Professor (Doctor of science (Japan)), chief researcher of Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=8701101900>, <https://www.webofscience.com/wos/author/record/1436451>

KHAIROVA Nina, Doctor of Technical Sciences, Professor, Chief Researcher of the Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=37461441200>, <https://www.webofscience.com/wos/author/record/1768515>

OTMAN Mohamed, PhD, Professor of Computer Science Department of Communication Technology and Networks, Putra University Malaysia (Selangor, Malaysia), <https://www.scopus.com/authid/detail.uri?authorId=56036884700>, <https://www.webofscience.com/wos/author/record/747649>

NYSANBAYEVA Saule Yerkebulanovna, Doctor of Technical Sciences, Associate Professor, Senior Researcher of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=55453992600>, <https://www.webofscience.com/wos/author/record/3802041>

USATOVA Olga Alexandrovna, PhD, Associate Professor, Chief Scientific Secretary of the Institute of Information and Computing Technologies of the National Academy of Sciences of the Republic of Kazakhstan (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=57204581062>, <https://www.webofscience.com/wos/author/record/JCO-3058-2023>

KAPALOVA Nursulu Aldazharovna, Candidate of Technical Sciences, Head of the Laboratory cybersecurity, Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=57191242124>,

KOVALYOV Alexander Mikhailovich, Doctor of Physical and Mathematical Sciences, Academician of the National Academy of Sciences of Ukraine, Institute of Applied Mathematics and Mechanics (Donetsk, Ukraine), <https://www.scopus.com/authid/detail.uri?authorId=7202799321>, <https://www.webofscience.com/wos/author/record/38481396>

MIKHALEVICH Alexander Alexandrovich, Doctor of Technical Sciences, Professor, Academician of the National Academy of Sciences of Belarus (Minsk, Belarus), <https://www.scopus.com/authid/detail.uri?authorId=7004159952>, <https://www.webofscience.com/wos/author/record/46249977>

TIGHINEANU Ion Mihailovich, Doctor of Physical and Mathematical Sciences, Academician, President of the Academy of Sciences of Moldova, Technical University of Moldova (Chisinau, Moldova), <https://www.scopus.com/authid/detail.uri?authorId=7006315935>, <https://www.webofscience.com/wos/author/record/524462>

Academic Scientific Journal of Computer Science

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Owner: «Central Asian Academic Research Center» LLP (Almaty).

Certificate № **KZ77VPY00121154** on the re-registration of the periodical printed and online publication of the information agency, issued on **05.06.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan

Subject area: *information and communication technologies*.

Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies»*.

Periodicity: *4 times a year*.

<http://www.physico-mathematical.kz/index.php/en/>

© «Central Asian Academic Research Center» LLP, 2026

БАС РЕДАКТОР:

МУТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=6506682964>, <https://www.webofscience.com/wos/author/record/1423665>

РЕДАКЦИЯ АЛҚАСЫ:

КАЛИМОЛДАЕВ Мақсат Нұрәділұлы, (бас редактордың орынбасары), физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының кеңесшісі, зертхана меңгерушісі (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=56153126500>, <https://www.webofscience.com/wos/author/record/2428551>

МАМЫРБАЕВ Өркен Жұмажанұлы (ғалым хатшы), Ақпараттық жүйелер саласындағы техника ғылымдарының (PhD) докторы, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» директорының ғылым жөніндегі орынбасары (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=55967630400>, <https://www.webofscience.com/wos/author/record/1774027>

БАЙГУНЧЕКОВ Жұмаілі Жаңабайұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Қолданбалы механика және инженерлік графика кафедрасы, Сәтбаев университеті (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=6506823633>, <https://www.webofscience.com/wos/author/record/1923423>

ВОЙЧИК Вальдемар, техника ғылымдарының докторы (физ-мат), Люблин технологиялық университетінің профессоры (Люблин, Польша), <https://www.scopus.com/authid/detail.uri?authorId=7005121594>, <https://www.webofscience.com/wos/author/record/678586>

СМОЛАРЖ Анджей, Люблин политехникалық университетінің электроника факультетінің доценті (Люблин, Польша), <https://www.scopus.com/authid/detail.uri?authorId=56249263000>, <https://www.webofscience.com/wos/author/record/1268523>

КЕЙЛАН Әлімхан, техника ғылымдарының докторы, профессор (ғылым докторы (Жапония)), ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=8701101900>, <https://www.webofscience.com/wos/author/record/1436451>

ХАЙРОВА Нина, техника ғылымдарының докторы, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=37461441200>, <https://www.webofscience.com/wos/author/record/1768515>

ОТМАН Мохаммед, PhD, Информатика, Коммуникациялық технологиялар және желілер кафедрасының профессоры, Путра университеті Малайзия (Селангор, Малайзия), <https://www.scopus.com/authid/detail.uri?authorId=56036884700>, <https://www.webofscience.com/wos/author/record/747649>

НЫСАНБАЕВА Сауле Еркебұланқызы, техника ғылымдарының докторы, доцент, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» аға ғылыми қызметкері (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=55453992600>, <https://www.webofscience.com/wos/author/record/3802041>

УСАТОВА Ольга Александровна, PhD, қауымдастырылған профессор, ҚР ҒЖБМ "Ақпараттық және есептеу технологиялары институтының" бас ғалым хатшысы (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=57204581062>, <https://www.webofscience.com/wos/author/record/JCO-3058-2023>

КАПАЛОВА Нұрсұлу Алдажарқызы, техника ғылымдарының кандидаты, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты», Киберқауіпсіздік зертханасының меңгерушісі (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=57191242124>,

КОВАЛЕВ Александр Михайлович, физика-математика ғылымдарының докторы, Украина Ұлттық Ғылым академиясының академигі, Қолданбалы математика және механика институты (Донецк, Украина), <https://www.scopus.com/authid/detail.uri?authorId=7202799321>, <https://www.webofscience.com/wos/author/record/38481396>

МИХАЛЕВИЧ Александр Александрович, техника ғылымдарының докторы, профессор, Беларусь Ұлттық Ғылым академиясының академигі (Минск, Беларусь), <https://www.scopus.com/authid/detail.uri?authorId=7004159952>, <https://www.webofscience.com/wos/author/record/46249977>

ТИГИНЯНУ Ион Михайлович, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), <https://www.scopus.com/authid/detail.uri?authorId=7006315935>, <https://www.webofscience.com/wos/author/record/524462>

Academic Scientific Journal of Computer Science

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Меншіктеуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы).

Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **05.06.2025 ж.** берген **№ KZ77VPY00121154** Куәлік.

Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*

Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

<http://www.physico-mathematical.kz/index.php/en/>

© «Орталық Азия академиялық ғылыми орталығы» ЖШС, 2026

Главный редактор:

МУТАНОВ Галимканр Мутанович, доктор технических наук, профессор, академик НАН РК, (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=6506682964>, <https://www.webofscience.com/wos/author/record/1423665>

Редакционная коллегия:

КАЛИМОЛДАЕВ Максат Нурадилович, (заместитель главного редактора), доктор физико-математических наук, профессор, академик НАН РК, советник генерального директора «Института информационных и вычислительных технологий» КН МНВО РК, заведующий лабораторией (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=56153126500>, <https://www.webofscience.com/wos/author/record/2428551>

МАМЫРБАЕВ Оркен Жумажанович, (ученый секретарь), доктор философии (PhD) по специальности «Информационные системы», заместитель директора по науке РГП «Институт информационных и вычислительных технологий» Комитета науки МНВО РК (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=55967630400>, <https://www.webofscience.com/wos/author/record/1774027>

БАЙГУНЧЕКОВ Жумадил Жанабаевич, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сатпаева (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=6506823633>, <https://www.webofscience.com/wos/author/record/1923423>

ВОЙЧИК Вальдемар, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), <https://www.scopus.com/authid/detail.uri?authorId=7005121594>, <https://www.webofscience.com/wos/author/record/678586>

СМОЛЯРЖ Анджей, доцент факультета электроники Люблинского политехнического университета (Люблин, Польша), <https://www.scopus.com/authid/detail.uri?authorId=56249263000>, <https://www.webofscience.com/wos/author/record/1268523>

КЕЙЛАН Алимхан, доктор технических наук, профессор (Doctor of science (Japan)), главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=8701101900>, <https://www.webofscience.com/wos/author/record/1436451>

ХАЙРОВА Нина, доктор технических наук, профессор, главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=37461441200>, <https://www.webofscience.com/wos/author/record/1768515>

ОТЧИК Мохамад, доктор философии, профессор компьютерных наук, Департамент коммуникационных технологий и сетей, Университет Путра Малайзия (Селангор, Малайзия), <https://www.scopus.com/authid/detail.uri?authorId=56036884700>, <https://www.webofscience.com/wos/author/record/747649>

НЫСАНБАЕВА Сауле Еркебулановна, доктор технических наук, доцент, старший научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=55453992600>, <https://www.webofscience.com/wos/author/record/3802041>

УСАТОВА Ольга Александровна, PhD, ассоциированный профессор, Главный ученый секретарь «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=57204581062>, <https://www.webofscience.com/wos/author/record/JCO-3058-2023>

КАПАЛОВА Нурсулу Алдажаровна, кандидат технических наук, заведующий лабораторией кибербезопасности РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=57191242124>,

КОВАЛЕВ Александр Михайлович, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), <https://www.scopus.com/authid/detail.uri?authorId=7202799321>, <https://www.webofscience.com/wos/author/record/38481396>

МИХАЛЕВИЧ Александрович, доктор технических наук, профессор, академик НАН Беларуси (Минск, Беларусь), <https://www.scopus.com/authid/detail.uri?authorId=7004159952>, <https://www.webofscience.com/wos/author/record/46249977>

ТИГИНЯНУ Ион Михайлович, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), <https://www.scopus.com/authid/detail.uri?authorId=7006315935>, <https://www.webofscience.com/wos/author/record/524462>

Academic Scientific Journal of Computer Science

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Собственник: *ТОО «Центрально-азиатский академический научный центр» (г. Алматы).*

Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания № **KZ77VPY00121154**. Дата выдачи **05.06.2025**

Тематическая направленность: *информационно-коммуникационные технологии.*

В настоящее время: *вошел в список журналов, рекомендованных КОКШВО МНВО РК по направлению «информационно-коммуникационные технологии».*

Периодичность: *4 раза в год.*

<http://www.physico-mathematical.kz/index.php/en/>

© ТОО «Центрально-азиатский академический научный центр», 2026

CONTENTS

COMPUTER SCIENCE

Abduraimova B.K., Toleukhan A.B., Sapakova S.Z., Abisheva A.A. Development of early cyberattack detection method using CNN-LSTM for IoT.....	11
Aben A.B., Kazbekova G.N., Baimakhanova A.S., Amanzholova A.B. Classification of birds and drones in the sky using MobileNetV2 model.....	30
Akbarov D., Sembayev T. Quality-aware pose–hand keypoint extraction pipeline for skeleton-based sign language recognition.....	44
Algazy K., Alimzhan Y., Sakan K., Nyssanbayeva S. Lattice-based vector commitments for Verkle trees.....	67
Asylkhan N., Baidrakhmanova M.G. Principles and models of spatial organization of buildings for crop production considering technological and climatic factors.....	87
Basheyeva Zh., Tokesh A., Bekish U., Abdoldinova G. Artificial intelligence for academic project management: a bibliometric analysis and systematic review.....	105
Bekmanova G., Kantureyeva M., Omarbekova A., Zakirova A., Issainova A. Integrating artificial intelligence to evaluate emotions in the learning environment.....	125
Dzhusupbekova G.T., Jangassiyev R.M. Gemini AI integration based on .NET MAUI for education: hybrid architecture and empirical load testing.....	146
Doszhan N.S., Sultanbekova L.Ye., Zhumagali S.Zh., Konysbayev E.K. Modeling and parameter calculation of an emergency response system based on LoRaWAN technology in the high-altitude conditions of the Zailiysky Alatau.....	166
Zhumakhanova A., Kudabayeva R., Akanova K., Myrkanova A. Entropy-normalized multidimensional model for user activity segmentation in Reddit...180	180
Karabaliyev Y., Kolesnikova K., Khlevnaya Y. HybridKazASR: a hybrid automatic speech recognition system combining multi-model rover fusion and morpheme-aware language modeling for Kazakh.....	198
Kerimkhulle S.E., Adalbek A., Baizakov N.A., Shodorova N.N. Piecewise logistic and fuzzy modeling of Kazakhstan's GDP dynamics (1990–2024)....	212
Kulakayeva A., Ashurov A., Aitmagambetov A., Ongenbayeva Zh. Development of mathematical models and criteria for the admissibility of orbital maneuvers of spacecraft.....	228

Kulatay A.A., Zhaisanova D.S., Daurenbayeva N.A., Mamanova S.Y., Tolegen M. Machine learning for personalized learning in gamified edtech platforms: Aqyl Battle case.....	248
Mamyrbayev O., Kurmetkan T. Enhanced sentiment analysis of e-commerce product reviews using Luong attention-based Bi-LSTM.....	263
Marassulov U.A., Kazbekova G. TF-IDF-based fake news detection in Kazakh and Russian.....	286
Omar A.B., Mussiraliyeva Sh.Zh. Federated learning: models based on transformer architecture.....	302
Rakhimova D., Duisenbekkyzy Zh., Karibayeva A., Eşref A., Ilessova B. Improving the voice recognition system for children in Kazakh through additional training (fine-tuning).....	317
Sarsembayev M., Urmashev B. Optimization of the calculation of kinetic equations of combustion processes on GPU using global memory and shared memory.....	335
Symagulov A., Smurygin V., Beloussov A., Karypov A., Yunicheva N.R. Improving the accuracy of crop and weed detection using UAVs in soya fields through image segmentation.....	347
Tashenova Zh., Gabdullin A.R., Abdugulova Zh., Amanzholova Sh., Santeyeva S. Security evaluation of WPA3 wireless networks under deauthentication attack scenarios.....	368
Tursunbayeva G.U., Satybaldina D.Zh., Tleuberdin S.T., Tashatov N.N., Egamberdiyev E.E. Anomaly detection in UAV telemetry systems based on simulation modeling.....	391
Tursynova N., Yerimbetova A., Amangeldy N., Zhumabayeva A., Daiyrbayeva E. Comparative analysis of multilingual transformer models for Kazakh-to-gloss translation.....	414
Shangpeng Lei, Balakayeva G. Dual-branch physical information neural networks for data center airflow velocity and thermal modeling.....	433
Shynzhigit B.B., Balabekova M.O., Amangeldy T.T., Malik G.J., Balgimbekova U.B. Automatic brick defects detection by using a CNN-based deep learning model.....	449

МАЗМҰНЫ

КОМПЬЮТЕРЛІК ҒЫЛЫМДАР

Абдураимова Б.К., Төлеухан Ә.Б., Сапакова С.З., Абишева А.А. Кибершабулдарды ерте анықтау әдісін CNN-LSTM негізінде дамыту (IoT үшін).....	11
Абен А.Б., Казбекова Г.Н., Баймаханова А.С., Аманжолова Ә.Б. MobileNetV2 моделімен аспандағы құстар мен дрондарды классификациялау.....	30
Акбаров Д.Р., Сембаев Т.М. Ым тілін тануға арналған дене қалпы мен қолдың негізгі нүктелерін сапаны бақылаумен анықтау әдісі.....	44
Алғазы К.Т., Әлімжан Е.Ж., Сақан Қ.С., Нысанбаева С.Е. Verkle ағаштарына арналған торлық векторлық міндеттемелер.....	67
Асылхан Н., Байдрахманова М.Г. Технологиялық және климаттық факторларды ескере отырып, өсімдік шаруашылығы ғимараттарының кеңістік ұйымдастыру қағидалары мен модельдері.....	87
Башеева Ж., Төкеш Ә., Бекіш Ұ., Абдолдинова Г. Академиялық жобаларды басқарудағы жасанды интеллект: библиометриялық талдау және жүйелі шолу.....	105
Бекманова Г.Т., Кантурсева М.А., Омарбекова А.С., Закирова А.Б., Исайнова А.Н. Оқу ортасындағы эмоцияларды бағалау үшін жасанды интеллектті біріктіру.....	125
Джусупбекова Г.Т., Джангасиев Р.М. Білім беруге арналған .NET MAUI негізіндегі Gemini AI интеграциясы: гибриді архитектура және эмпирикалық жүктемелік тестілеу.....	146
Досжан Н.С., Султанбекова Л.Е., Жумағали С.Ж., Конысбаев Е.К. Іле Алатауының биік таулы жағдайында LORAWAN технологиясы негізіндегі жедел әрекет ету жүйесінің параметрлерін модельдеу және есептеу.....	166
Жумаханова А., Қудабаева Р., Аканова К., Мырканова А. REDDIT-те пайдаланушы әрекетін сегменттеуге арналған энтропия-нормалданған көп өлшемді модель.....	180
Карабалиев Е., Колесникова К., Хлевная Ю. HybridKazASR: Rover көпмодельді біріктіру және морфемеге негізделген тілдік модельдеуді пайдаланатын қазақ тілін автоматты тану гибриді жүйесі.....	198
Керімқұл С.Е., Адалбек А., Байзаков Н.А., Шодорова Н.Н. Қазақстан ЖІӨ динамикасын кезеңдік (Piecewise) логистикалық және бұлдыр модельдеу (1990–2024).....	212

Кулакаева А.Е., Ашуров А.Е., Айтмағамбетов А.З., Онгенбаева Ж.Ж. Ғарыш аппараттарының орбиталық маневрлерінің математикалық модельдері мен рұқсат критерийлерін әзірлеу.....	228
Құлатай А.А., Жайсанова Д.С., Дауренбаева Н.А., Маманова С.Е., Төлеген М. Геймификацияланған edtech платформаларда оқытуды жекелендіруге арналған машиналық.....	248
Мамырбаев Ө.Ж., Құрметқан Т. Луонг назар механизміне негізделген BI-LSTM көмегімен электрондық коммерция өнімдеріне жазылған пікірлерге жетілдірілген сентименттік талдау жасау.....	263
Марасулов У.А., Казбекова Г. Қазақ және орыс тілдеріндегі жалған жаңалықтарды TF-IDF арқылы анықтау.....	286
Омар А.Б., Мусиралиева Ш.Ж. Федеративті оқыту: трансформер архитектурасына негізделген модельдер.....	302
Рахимова Д., Дүйсенбекқызы Ж., Кәрібаева А., Еşref А., Ілесова Б. Қазақ тіліндегі балалар дауысын тану жүйесін қосымша оқыту (Fine-Tuning) арқылы жетілдіру.....	317
Сарсембаев М., Урмашев Б. Global memory және shared memory қолдану арқылы GPU-да жану процестерінің кинетикалық теңдеулерін есептеуді оңтайландыру.....	335
Сымагулов А., Смурыгин В., Белоусов А., Карыпов А., Юничева Н.Р. Соя алқаптарында ҰҰА көмегімен мәдени және арамшөп өсімдіктерін детекттеу сапасын кескіндерді сегменттеу арқылы арттыру.....	347
Ташенова Ж.М., Габдуллин А.Р., Абдугулова Ж.К., Аманжолова Ш.А., Сантеева С.Ә. Деатентификациялау шабуылы сценарийлеріндегі WPA3 сымсыз желілерінің қауіпсіздігін бағалау.....	368
Турсунбаева Г., Сатыбалдина Д., Тлеубердин С., Ташатов Н., Эгамбердиев Э. Симуляциялық модельдеу негізінде ұшқышсыз ұшу аппараттарының телеметриялық жүйелеріндегі аномалияларды анықтау.....	391
Турсынова Н., Еримбетова А., Амангелді Н., Жумабаева А., Дайырбаева Э. Қазақ тілінен глоссқа аудару үшін көптілді трансформерлік модельдердің салыстырмалы талдауы.....	414
Шанпэн Лей, Балакаева Г. Деректер орталығының ауа ағынының жылдамдығына және термиялық модельдеуге арналған екі тармақты физикалық ақпараттық нейрондық желілер.....	433
Шынжігіт Ш.Б., Балабекова М.О., Амангелді Т.Т., Мәлік Г.Ж., Балгимбекова У.Б. Кіріпші ақауларын автоматты анықтауда cnn негізіндегі терең оқыту моделін пайдалану.....	449

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЕ НАУКИ

Абдураимова Б.К., Толеухан А.Б., Сапакова С.З., Абишева А.А. Разработка метода раннего обнаружения кибератак на основе CNN-LSTM для IoT.....	11
Абен А.Б., Казбекова Г.Н., Баймаханова А.С., Аманжолова А.Б. Классификация птиц и дронов в небе с использованием модели MobileNetV2.....	30
Акбаров Д.Р., Сембаев Т.М. Метод получения ключевых точек позы и кистей с контролем качества для распознавания жестового языка.....	44
Алгазы К.Т., Алимжан Е.Ж., Сакан К.С., Нысанбаева С.Е. Решеточные векторные обязательства для Verkle-деревьев.....	67
Асылхан Н., Байдрахманова М.Г. Принципы и модели пространственной организации зданий для растениеводства с учетом технологических и климатических факторов.....	87
Башеева Ж., Токеш А., Бекиш У., Абдолдинова Г. Искусственный интеллект в управлении академическими проектами: библиометрический анализ и систематический обзор.....	105
Бекманова Г.Т., Кантурсева М.А., Омарбекова А.С., Закирова А.Б., Исайнова А.Н. Интеграция искусственного интеллекта для оценки эмоций в учебной среде.....	125
Джусупбекова Г.Т., Джангасиев Р.М. Интеграция Gemini AI на базе .NET MAUI для образования: гибридная архитектура и эмпирическое нагрузочное тестирование.....	146
Досжан Н.С., Султанбекова Л.Е., Жумагали С.Ж., Конысбаев Е.К. Моделирование и расчет параметров системы экстренного реагирования на базе технологии LoRaWAN в условиях высокогорья Заилийского Алатау.....	166
Жумаханова А., Кудабасева Р., Аканова К., Мырканова А. Энтропийно-нормализованная многомерная модель для сегментации активности пользователей в Reddit.....	180
Карабалиев Е., Колесникова К., Хлевна Ю. HybridKazASR: гибридная система автоматического распознавания казахской речи на основе многомодельного объединения ROVER и морфемно-ориентированного языкового моделирования.....	198
Керимкулов С.Е., Адалбек А., Байзаков Н.А., Шодорова Н.Н. Кусочно-логистическое и нечеткое моделирование динамики ВВП Казахстана (1990–2024).....	212
Кулакаева А.Е., Ашуров А.Е., Айтмагамбетов А.З., Онгенбаева Ж.Ж. Разработка математических моделей и критериев допустимости орбитальных маневров космических аппаратов.....	228

Кулатай А.А., Жайсанова Д.С., Дауренбаева Н.А., Маманова С.Е., Толеген М. Машинное обучение для персонализации обучения на геймифицированных EdTech-платформах: кейс Aqyl Battle.....	248
Мамырбаев О., Курметкан Т. Усовершенствованный анализ тональности отзывов о товарах электронной коммерции с использованием Bi-LSTM на основе механизма внимания Луонга.....	263
Марасулов У.А., Казбекова Г. Выявление ложных новостей на казахском и русском языках TF-IDF-моделями.....	286
Омар А.Б., Мусиралиева Ш.Ж. Федеративное обучение: модели на основе архитектуры трансформеров.....	302
Рахимова Д., Дуйсенбеккызы Ж., Карибаева А., ЕҒref А., Илесова Б. Совершенствование системы распознавания голоса детей на казахском языке путем дополнительного обучения (fine-tuning).....	317
Сарсембаев М., Урмашев Б. Оптимизация расчета кинетических уравнений процессов горения на GPU с использованием global memory и shared memory.....	335
Сымагулов А., Смурыгин В., Белоусов А., Карыпов А., Юничева Н.Р. Улучшение качества детектирования культурных и сорных растений с помощью БПЛА на полях сои с применением сегментации изображений.....	347
Ташенова Ж.М., Габдуллин А.Р., Абдугулова Ж.К., Аманжолова Ш.А., Сантеева С.А. Оценка безопасности беспроводных сетей WPA3 в условиях атаки с деаутентификацией.....	368
Турсунбаева Г., Сатыбалдина Д., Тлеубердин С., Ташатов Н., Эгамбердиев Э. Обнаружение аномалий в телеметрических системах БПЛА на основе симуляционного моделирования.....	391
Турсынова Н., Еримбетова А., Амангелді Н., Жумабаева А., Дайырбаева Э. Сравнительный анализ многоязычных трансформерных моделей для перевода с казахского языка на глоссированное представление.....	414
Шанпэн Лэй, Балакаева Г. Двухветвевые физически информированные нейронные сети для моделирования воздушных потоков и тепловых условий в центрах обработки данных.....	433
Шынжыгит Ш.Б., Балабекова М.О., Амангелды Т.Т., Малик Г.Ж., Балгимбекова У.Б. Использование модели глубокого обучения на основе CNN для автоматического обнаружения дефектов кирпичной кладки.....	449

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 11–29

<https://doi.org/10.32014/2026.2518-1726.424>

IRSTI 20.23.17
UDC 004.056.53

Abduraimova B.K.¹, Toleukhan A.B.^{1*}, Sapakova S.Z.², Abisheva A.A.³, 2026.

¹L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;

²International information technologies university, Almaty, Kazakhstan;

³Kazakh University of Technology and Business named after K. Kulazhanov,
Astana, Kazakhstan.

E-mail: aminkatoleukhanova@gmail.com

DEVELOPMENT OF EARLY CYBERATTACK DETECTION METHOD USING CNN-LSTM FOR IOT

Abduraimova Bayan — Candidate of Technical Sciences, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: abduraimovabk@mail.ru, ORCID: <https://orcid.org/0000-0003-3913-1895>;

Toleukhan Amina — L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: aminkatoleukhanova@gmail.com, ORCID ID: <https://orcid.org/0009-0008-6900-7813>;

Sapakova Saya — associate professor, International information technologies university, Almaty, Kazakhstan,

E-mail: s.sapakova@iitu.edu.kz, ORCID: <https://orcid.org/0000-0001-6541-6806>;

Abisheva Aigul — Kazakh University of Technology and Business named after K. Kulazhanov, Astana, Kazakhstan,

E-mail: aigul.abisheva@gmail.com, ORCID: <https://orcid.org/0000-0002-5800-9134>.

Abstract. The rapid development of the Internet of Things (IoT) and the widespread adoption of network-connected devices have led to a significant increase in network traffic and the number of cyberattacks. Traditional intrusion detection methods often require analysis of the entire network flow and extensive computing resources, limiting their application in IoT environments where computing power and memory are severely limited. Therefore, the development of effective approaches for the early detection of cyberattacks has become an important area of research. Such approaches enable security systems to detect potential threats early in the network interaction process and respond more quickly to malicious activity. This paper presents a method for early cyberattack detection based on a hybrid deep learning architecture that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. The CNN component extracts spatial features from network traffic data, while the LSTM layers identify temporal

dependencies in packet sequences. To improve the information yield of traffic analysis, an expanded set of statistical-dynamic features is proposed, including the entropy of packet size distribution, inter-packet intervals, and the variation coefficient. Experimental results show that the proposed approach provides higher classification accuracy compared to traditional machine learning algorithms such as support vector machines (SVM), random forests, and multilayer perceptrons (MLP). Analysis of the first ten packets of a network connection significantly reduces attack detection time, which is especially important for real-time systems. Furthermore, an optimized version of the model, adapted for the IoT environment, reduces computational load while maintaining high detection accuracy and practical applicability in modern network infrastructures.

Keywords: cyberattacks, deep learning, CNN-LSTM, IoT, network traffic, intrusion detection system (IDS), model optimization

For citations: Abduraimova B.K., Toleukhan A.B., Sapakova S.Z., Abisheva A.A. Development of early cyberattack detection method using cnn-lstm for IOT. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 11-29. DOI: <https://doi.org/10.32014/2026.2518-1726.424>

**Абдураимова Б.К.¹, Төлеухан Ә.Б.^{1*}, Сапакова С.З.²,
Абишева А.А.³, 2026.**

¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;

²Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;

³Қ. Құлажанов атындағы Қазақ технология және бизнес университеті,
Астана, Қазақстан.

E-mail: aminkatoleukhanova@gmail.com

КИБЕРШАБУЛДАРДЫ ЕРТЕ АНЫҚТАУ ӘДІСІН CNN-LSTM НЕГІЗІНДЕ ДАМУЫ (IOT ҮШІН)

Абдураимова Баян — техника ғылымдарының кандидаты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: abduraimovabk@mail.ru, ORCID: <https://orcid.org/0000-0003-3913-1895>;

Төлеухан Әмина — Л.Н. Гумилева атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: aminkatoleukhanova@gmail.com, ORCID ID: <https://orcid.org/0009-0008-6900-7813>;

Сапакова Сая — физика-математика ғылымдарының кандидаты, компьютерлік инженерия кафедрасының қауымдастырылған профессоры, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,

E-mail: s.sapakova@iitu.edu.kz, ORCID: <https://orcid.org/0000-0001-6541-6806>;

Абишева Айгуль — сеньор-лектор, Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан,

E-mail: aiqul.abisheva@gmail.com, ORCID: <https://orcid.org/0000-0002-5800-9134>.

Аннотация. Заттар интернетінің (IoT) жылдам дамуы және желіге қосылған құрылғылардың кеңінен қолданылуы желілік трафиктің және кибершабуыл-

дар санының айтарлықтай өсуіне әкелді. Дәстүрлі басып кіруді анықтау әдістері көбінесе бүкіл желі ағынын және кең есептеу ресурстарын талдауды талап етеді, бұл оларды есептеу қуаты мен жады өте шектеулі IoT орталарында қолдануды шектейді. Сондықтан кибершабуылдарды ерте анықтаудың тиімді тәсілдерін әзірлеу зерттеудің маңызды саласына айналды. Мұндай тәсілдер қауіпсіздік жүйелеріне желілік өзара әрекеттесу процесінің басында әлеуетті қауіптерді анықтауға және зиянды әрекетке тезірек жауап беруге мүмкіндік береді. Бұл мақалада конволюциялық нейрондық желілер (CNN) мен ұзақ мерзімді жад (LSTM) желілерін біріктіретін гибриді терең оқыту архитектурасына негізделген кибершабуылды ерте анықтау әдісі ұсынылған. CNN компоненті желілік трафик деректерінен кеңістіктік ерекшеліктерді алады, ал LSTM қабаттары пакеттік тізбектердегі уақытша тәуелділіктерді анықтайды. Трафикті талдаудың ақпараттық өнімділігін жақсарту үшін пакет өлшемінің таралу энтропиясын, пакеттер арасындағы аралықтарды және вариация коэффициентін қоса алғанда, статистикалық-динамикалық ерекшеліктердің кеңейтілген жиынтығы ұсынылады. Тәжірибелік нәтижелер ұсынылған тәсілдің тірек векторлық машиналар (SVM), кездейсоқ ормандар және көп қабатты перцептрондар (MLP) сияқты дәстүрлі машиналық оқыту алгоритмдерімен салыстырғанда жоғары жіктеу дәлдігін қамтамасыз ететінін көрсетеді. Желілік қосылымның алғашқы он пакетін талдау шабуылды анықтау уақытын айтарлықтай қысқартады, бұл әсіресе нақты уақыт жүйелері үшін маңызды. Сонымен қатар, IoT ортасына бейімделген модельдің оңтайландырылған нұсқасы заманауи желілік инфрақұрылымдарда жоғары анықтау дәлдігін және практикалық қолданылуын сақтай отырып, есептеу жүктемесін азайтады.

Түйін сөздер: кибершабуылдар, терең оқыту, CNN-LSTM, IoT, желілік трафик, шабуылды анықтау жүйесі (IDS), модельді оңтайландыру

Абдураимова Б.К.¹, Толеухан А.Б.¹, Сапакова С.З.², Абишева А.А.³, 2026.

¹Евразийский национальный университет имени Л.Н. Гумилёва,
Астана, Казахстан;

²Международный университет информационных технологий,
Алматы, Казахстан;

³Казахский университет технологии и бизнеса имени К. Кулажанова,
Астана, Казахстан.

E-mail: aminkatoleukhanova@gmail.com

РАЗРАБОТКА МЕТОДА РАННЕГО ОБНАРУЖЕНИЯ КИБЕРАТАК НА ОСНОВЕ CNN-LSTM ДЛЯ ИОТ

Абдураимова Баян — кандидат технических наук, Евразийский Национальный университет имени Л.Н. Гумилева., Астана, Казахстан,
E-mail: abduraimovabk@mail.ru, ORCID: <https://orcid.org/0000-0003-3913-1895>;

Толехан Амина — ЕНУ им. Л.Н. Гумилева, Астана, Казахстан,

E-mail: aminkatoleukhanova@gmail.com, ORCID ID: <https://orcid.org/0009-0008-6900-7813>;

Сапакова Сая — кандидат физ.-мат. наук, ассоциированный профессор кафедры компьютерной инженерии, Международный университет информационных технологий, Алматы, Казахстан,

E-mail: s.sapakova@iiu.edu.kz, ORCID: <https://orcid.org/0000-0001-6541-6806>;

Абишева Айгуль — сеньор-лектор, Казахский университет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан,

E-mail: aiqul.abisheva@gmail.com, ORCID: <https://orcid.org/0000-0002-5800-9134>.

Аннотация. *Актуальность.* Быстрое развитие Интернета вещей (IoT) и широкое распространение сетевых устройств привели к значительному увеличению объема сетевого трафика и числа кибератак. Традиционные методы обнаружения вторжений часто требуют анализа всего сетевого потока и значительных вычислительных ресурсов, что ограничивает их применение в IoT-средах с ограниченными вычислительными мощностями и памятью. В связи с этим разработка эффективных подходов к раннему обнаружению кибератак является актуальной научно-практической задачей. *Цель.* Разработать метод раннего обнаружения кибератак на основе гибридной архитектуры глубокого обучения CNN-LSTM, обеспечивающий своевременное выявление вредоносной активности в IoT-сетях при сниженной вычислительной нагрузке. *Методы.* В статье предложен метод раннего обнаружения кибератак, основанный на гибридной архитектуре глубокого обучения, сочетающей сверточные нейронные сети (CNN) и сети долговременной краткосрочной памяти (LSTM). Компонент CNN используется для извлечения пространственных признаков из данных сетевого трафика, а слои LSTM позволяют выявлять временные зависимости в последовательностях пакетов. Для повышения эффективности анализа трафика предложен расширенный набор статистико-динамических характеристик, включающий энтропию распределения размеров пакетов, интервалы между пакетами и коэффициент вариации. *Результаты и выводы.* Экспериментальные результаты показали, что предложенный подход обеспечивает высокую точность классификации по сравнению с традиционными алгоритмами машинного обучения, включая машины опорных векторов (SVM), случайный лес и многослойный перцептрон (MLP). Анализ первых десяти пакетов сетевого соединения позволяет существенно сократить время обнаружения атаки, что особенно важно для систем реального времени. Оптимизированная версия модели, адаптированная к условиям IoT-среды, снижает вычислительную нагрузку, сохраняя при этом высокую точность обнаружения и практическую применимость в современных сетевых инфраструктурах.

Ключевые слова: кибератаки, глубокое обучение, CNN-LSTM, IoT, сетевой трафик, система обнаружения вторжений, IDS, оптимизация модели

Introduction. The digitalization of the economy, the development of cloud technologies, the widespread adoption of distributed computing, and the widespread

use of Internet of Things (IoT) devices have significantly expanded the attack surface of modern information systems. According to analytical reports and recent academic research (Banaamah and Ahmad, 2022), a steady increase in DDoS attacks, application-layer attacks, exploitation of IoT device vulnerabilities, and automated botnet campaigns is expected from 2022 to 2024. The increasing number of connected devices, the heterogeneity of network infrastructure, and the growth of traffic volumes pose new challenges for information security systems. Therefore, the task of timely detection of cyberattacks has become especially urgent.

Traditional intrusion detection systems (IDS) based on signature analysis are primarily designed to identify known patterns of malicious activity. However, such approaches perform poorly in detecting modified, polymorphic, or previously unknown attacks (Ferrag and Maglaras, 2022). Furthermore, signature-based methods require constant database updates, increasing the workload of security administrators.

Machine learning and anomaly analysis methods have significantly improved the accuracy of network traffic classification (Shone et al., 2022). However, most existing models analyze completed network flows. This means that classification occurs only after the entire packet sequence has been received, increasing the time it takes to detect attacks. In real-world systems, such delays can have serious consequences, ranging from service disruption to data leakage or malware propagation.

In the IoT environment, the problem becomes even more complex due to limited computing resources, memory, and energy efficiency requirements on edge devices (Nguyen et al., 2022). Deploying heavy-duty deep learning models on such devices is often impossible without architectural optimization. Therefore, there is a need for methods that combine high accuracy with low computational complexity. Modern research demonstrates the promise of hybrid deep learning architectures combining convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. CNNs effectively extract spatial features from packet sequences, while LSTMs account for temporal dependencies and network traffic dynamics. However, most studies still focus on analyzing the full network flow and do not address the problem of early detection, which begins at the level of the first packets of a connection.

Therefore, the scientific challenge arises of developing a method for early cyberattack detection that:

- enables classification based on the first N packets of a network flow;
- accounts for spatiotemporal dependencies;
- reduces attack detection time;
- is adapted to the limitations of IoT devices.

The goal of this study is to develop a method for early cyberattack detection based on a hybrid CNN–LSTM architecture optimized for the IoT environment. To achieve this goal, the following objectives are addressed:

- formalization of a model for early network traffic analysis;

- development of an extended set of statistical-dynamic features;
- creation of a hybrid CNN–LSTM architecture;
- comparative analysis with classical machine learning algorithms;
- model optimization for implementation in IoT infrastructure.

The scientific significance of this work lies in the combination of early traffic analysis, hybrid deep learning, and IoT-oriented optimization within a single model, as well as the quantitative assessment of the reduction in attack detection latency.

The scientific novelty of this study lies in the combination of methodological and architectural solutions aimed at addressing the identified gap in intelligent intrusion detection systems. A method for early detection of cyberattacks based on the analysis of the first N packets of a network flow is proposed. Unlike most existing approaches, which use the full network flow, the proposed method enables decision-making at the initial stage of data transmission, significantly reducing detection latency and increasing applicability to real-time systems.

A hybrid CNN-LSTM architecture, specifically adapted for the task of early detection, has been developed. Unlike typical solutions, where such models are used to improve accuracy on complete data, this architecture is adapted to work with a limited number of packets while maintaining high accuracy. An expanded set of statistical-dynamic features is also proposed, which takes into account not only aggregated traffic characteristics but also the behavioral dynamics of data transmission at the early stage of a connection, making the data presentation more informative.

The paper simultaneously evaluates three key metrics: classification accuracy, detection latency, and computational complexity. Most studies consider only Accuracy and F1-score, while temporal and computational

Related work. Modern research in the field of intrusion detection systems (IDS) demonstrates a clear shift from classic signature-based methods to intelligent approaches to network traffic analysis. While previously the focus was on identifying known attack patterns, between 2022 and 2024, increasing research will focus on the application of machine and deep learning methods to detect anomalies and previously unknown threats based on the behavioral characteristics of network traffic (Arisdakessian et al., 2023).

This transformation is driven by the growing complexity of network infrastructures, the rapid development of cloud technologies, and the widespread adoption of Internet of Things (IoT) devices. All of this has led to a dramatic increase in network traffic volumes and its high heterogeneity. As a result, traditional analysis methods are becoming less effective and are unable to meet modern requirements for the speed and accuracy of attack detection. In the current scientific literature, several key areas of development for intelligent IDS systems can be identified. The first area involves the use of deep neural networks for network traffic analysis. Convolutional neural networks (CNNs) are widely used to detect spatial patterns in feature space. They enable the automatic extraction of meaningful features without the need for manual feature engineering. Recurrent neural networks, including

LSTMs and GRUs, are used to analyze the temporal structure of traffic by modeling packet sequences and network connection dynamics.

Research results show that such models demonstrate higher accuracy and recall than classical machine learning algorithms, especially when detecting complex and nonlinear attacks (Ashraf et al., 2022). However, the main drawback of most existing approaches is that they analyze fully formed network flows. This means that the determination of an attack is made only after receiving the entire packet sequence, increasing detection latency and reducing the responsiveness of security systems. A second important area of research involves hybrid deep learning architectures. In particular, CNN-LSTM models, which allow for the simultaneous consideration of spatial and temporal dependencies in network traffic, have become widely used. In such models, convolutional layers are responsible for extracting local features, while recurrent layers analyze sequential dependencies between packets and identify connection dynamics. Several studies report that such hybrid architectures achieve classification accuracy of 94–96%, making them among the most effective solutions in this field.

However, despite their high accuracy, most such models still operate on full network flows. This means that the problem of early attack detection—based on the first packets of a connection—remains understudied. Furthermore, existing studies rarely consider attack detection latency as a separate and independent performance metric, although in real-world settings, response speed plays a critical role. A third area of research concerns the adaptation of deep learning models for IoT and edge devices. Due to the limited computational resources of such devices, model optimization is necessary. Various methods are used for this purpose, including parameter reduction, weight quantization, neural network pruning, and architecture simplification. These approaches significantly reduce computational load and memory requirements, but often lead to a decrease in classification accuracy. However, a significant problem remains that most of these optimization methods are considered separately from the early attack detection task. This prevents a full assessment of how model simplification impacts its ability to detect threats at an early stage of network communication.

Overall, an analysis of existing research shows that despite significant progress in the field of intelligent IDS, a significant research gap remains. This gap lies in the absence of a unified approach that simultaneously integrates early attack detection, hybrid deep learning architectures, and model adaptation to the constraints of the IoT environment. Most existing research focuses either on improving classification accuracy or reducing computational complexity, while the temporal aspect of attack detection remains under-explored. With cyber threats growing, especially in distributed and IoT networks, early attack detection is becoming a key factor in ensuring the security of modern information systems.

Table 1 — Comparative analysis of recent IDS studies (2022–2025)

Authors, Year	Architecture	Dataset	Accuracy / F1	IoT Optimization	Latency Analysis	Main Limitations
Nguyen et al., 2022	Lightweight IDS	IoT gateway traffic	94% / 92%	Yes	No	Limited attack diversity
Ashraf et al., 2022	DNN-based IDS	IoT network traffic	96% / 94%	No	Partial	High resource consumption
Otoum & Nayak, 2022	DL-IDS	IoT traffic	94% / 92%	Partial	No	No early-stage detection
Alghamdi & Bellaiche, 2023	Ensemble Deep Learning IDS	IoT traffic	91% / 89%	Yes	No	Accuracy degradation after optimization
Mehedi et al., 2023	Deep Transfer Learning IDS	IoT datasets	95% / 93%	Partial	No	Increased training complexity
Sharmila & Nagapadma, 2023	Quantized Autoencoder	RT-IoT2022	92% / 90%	Yes	No	Reduced adaptability
Proposed study, 2026	CNN–LSTM (Early Detection)	CICIDS2017	96% / 94%	Yes	Yes	–

Analysis shows that hybrid CNN-LSTM architectures, as well as transformer-based models, are increasingly being used. Although many studies achieve high classification accuracy (93–97%), most approaches still rely on full network flow analysis. Optimization for the IoT environment is only implemented in isolated studies and typically results in a decrease in model accuracy. Furthermore, attack detection latency is rarely considered as a separate quantitative metric. As a result, existing research lacks a unified methodology that simultaneously integrates early traffic analysis, hybrid CNN-LSTM architectures, adaptation to IoT constraints, and detection latency assessment. This study aims to address this gap.

Materials and Method. Within this study, the cyberattack detection task is formulated as a binary classification problem for network traffic flows. Let a network flow be represented as a sequence of packets:

$$F = \{p_1, p_2, \dots, p_T\} \quad (1)$$

where:

F – the network flow;

p_i – the i -th packet in the flow;

T – the total number of packets in the connection.

In traditional intrusion detection systems, classification is performed after analyzing the complete flow F . However, this approach increases attack detection

latency and reduces the speed of response. In the proposed method, a shortened sequence consisting of the first N packets is analyzed:

$$F_N = \{p_1, p_2, \dots, p_N\}, N \ll T \quad (2)$$

where:

F_N – the truncated network flow consisting of the first N packets;

N – the number of packets analyzed at the initial stage of the connection.

As a result, the model determines whether the traffic belongs to the attack class or legitimate class before the completion of the transmission process, reducing the time required for threat detection. The classification function is defined as:

$$f_{\theta}(F_N) \rightarrow y \quad (3)$$

where:

f_{θ} – a parameterized model with learnable parameters θ ;

F_N – the truncated network flow consisting of the first N packets;

$y \in \{0, 1\}$ — the class label (0 — legitimate traffic, 1 — attack).

Model parameters are obtained by minimizing the empirical risk using the binary cross-entropy loss function. Each packet is represented by a set of basic features:

$$x_i = (l_i, t_i, flag_i, src_i, dst_i, proto_i) \quad (4)$$

where:

l_i – packet length;

t_i – timestamp;

$flag_i$ – TCP flags;

src_i – source port;

dst_i – destination port;

$proto_i$ – protocol type.

To describe network traffic behavior, a number of statistical features are calculated, including average packet length, packet length variance, and the coefficient of variation. These metrics reflect the overall characteristics of traffic intensity and variability during the initial phase of a connection. Dynamic features are also extracted to analyze the temporal behavior of traffic. These include interpacket intervals, average interpacket delay, and the rate of change of packet length, which reflects how packet size changes over time.

To analyze the behavioral characteristics of network traffic during the initial phase of a connection, several statistical metrics are calculated. These aggregated

metrics allow for a compact representation of packet-level information and enable the model to identify patterns characteristic of both legitimate and malicious traffic.

Average packet length is determined as follows:

$$\mu = \frac{1}{N} \sum_{i=1}^N l_i \quad (5)$$

where:

l_i – the length of the i -th packet;

N – the number of analyzed packets in the truncated flow.

This metric reflects the overall intensity of data transmission during the early stage of a network connection. In many attack scenarios, such as flooding attacks or scanning activities, the average packet length may significantly differ from that observed in normal traffic patterns. Another important statistical characteristic is the packet length variance, which measures the dispersion of packet sizes within the analyzed sequence:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (l_i - \mu)^2 \quad (6)$$

where:

l_i – the length of the i -th packet;

μ – the average packet length;

N – the number of analyzed packets.

This metric allows the model to evaluate the variability of packet sizes. Malicious traffic often demonstrates distinctive variance patterns. For example, certain attacks generate packets with highly similar sizes, while others produce abrupt fluctuations in packet length due to abnormal traffic generation strategies. To normalize the variance relative to the mean value, the coefficient of variation (CV) is calculated:

$$CV = \frac{\sigma}{\mu} \quad (7)$$

The coefficient of variation allows us to estimate variability in relative terms, independent of the scale of the values. This allows the model to compare different types of network traffic without relying on absolute packet length values. This is especially important in heterogeneous network environments, where packet sizes can vary significantly depending on the protocol or application.

In addition to statistical characteristics, dynamic features are also extracted,

allowing us to account for the temporal behavior of network traffic. These features reflect how packet parameters change over time and help identify anomalous patterns associated with cyberattacks. The interpacket interval is defined as the time difference between two consecutive packets:

$$\Delta t_i = t_i - t_{i-1} \quad (8)$$

where:

t_i – the timestamp of the i -th packet;

t_{i-1} – the timestamp of the previous packet.

Based on these intervals, the average inter-packet delay is calculated as:

$$\Delta t = \frac{1}{N-1} \sum_{i=2}^N \Delta t_i \quad (9)$$

This feature is particularly important for identifying abnormal traffic generation rates. For instance, Distributed Denial-of-Service (DDoS) attacks often generate packets at extremely short time intervals, resulting in unusually low inter-packet delays compared with normal network activity. Another dynamic indicator is the packet length growth rate, which measures how packet sizes change over time:

$$G_i = \frac{l_i - l_{i-1}}{\Delta t_i} \quad (10)$$

This feature captures the temporal evolution of packet structure and may reveal abnormal traffic patterns caused by automated attack tools or malicious scripts that generate packets following specific patterns. To quantify the randomness of packet size distribution, the entropy of packet lengths is calculated:

$$H = - \sum p(x) \log p(x) \quad (11)$$

where:

$p(x)$ – the probability of observing packets with a specific size x .

Entropy serves as a measure of the uncertainty or disorder within the packet size distribution. High entropy values indicate a highly irregular and diverse traffic structure, which may occur during complex multi-stage attacks or botnet activity. Conversely, low entropy values often correspond to highly uniform packet sizes, which can be characteristic of automated attack patterns such as flooding or scanning operations. Additionally, the entropy of packet size distribution is calculated to measure the degree of randomness in traffic patterns. The resulting feature representation forms a tensor:

$$X \in R^{N \times d'} \quad (12)$$

where:

X – the feature tensor representing the packet sequence;

N – the number of packets in the analyzed sequence;

d' – the total number of extracted features.

To model spatial and temporal dependencies within the packet sequence, a hybrid CNN–LSTM architecture is employed. The convolutional layer extracts local spatial patterns from the feature tensor, enabling identification of repeated structures in packet attributes such as packet length distribution and flag combinations. The recurrent component of the architecture is based on Long Short-Term Memory (LSTM) networks, which are designed to model temporal dependencies within sequential data. In the context of network traffic analysis, LSTM layers capture the dynamic behavior of packet sequences and preserve information about previous states of the connection. The memory state of the LSTM cell is updated using several gating mechanisms. The forget gate determines which information from the previous cell state should be retained or discarded:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (13)$$

where:

f_t – the forget gate activation at time step t ;

h_{t-1} – the hidden state from the previous time step;

x_t – the input vector at time step t ;

W_f – the weight matrix of the forget gate;

b_f – the bias vector;

σ – the sigmoid activation function.

The forget gate allows the model to selectively remove irrelevant information from the previous cell state, which helps the network focus on important temporal patterns within packet sequences. The recurrent LSTM component models temporal dependencies in packet sequences, allowing the system to capture long-term behavioral patterns of network traffic. The model is trained using a binary cross-entropy loss function with regularization to prevent overfitting. Computational complexity analysis shows that the convolutional component has complexity $O(N \cdot d \cdot k)$, where k represents the filter size, while the LSTM component has complexity $O(N \cdot h^2)$, where h denotes the number of hidden neurons. The total complexity of the model can therefore be approximated as $O(N(dk + h^2))$, $h' < h$.

To enable deployment in IoT environments, the model architecture was optimized by reducing the number of LSTM neurons and model parameters. As a result, the total number of parameters decreased by 42%, leading to a 38%

reduction in inference time. This optimization significantly reduces computational load, making the proposed model suitable for resource-constrained IoT devices.

Results. An experimental evaluation of the proposed early cyberattack detection method was conducted to comprehensively analyze classification accuracy, detection latency, and computational efficiency. The experimental design was based on research guidelines in the field of intrusion detection systems (Alghamdi & Bellaiche, 2023), ensuring comparability of the obtained results with existing scientific works. The publicly available CICIDS2017 dataset, widely used in network attack detection, was used as a test dataset. It includes both legitimate network traffic and various attack types, such as DDoS, PortScan, Brute Force, and Botnet activity. This dataset allows for direct comparison with studies published between 2022 and 2025. After a data preprocessing step, including the removal of incomplete records and feature normalization, the final dataset consisted of 78,456 network flows. Numerical features were normalized using the Min-Max scaling method, which is consistent with standard data preparation practice in network traffic analysis (Mehedi et al., 2023). The data was then split into training (70%), validation (15%), and test (15%) sets using a stratified approach to preserve class distributions (Arisdakessian et al., 2023). This partitioning method is widely used in IDS research and ensures a correct assessment of model quality. A key feature of the experimental setup was the modeling of the early stage of the network connection. For each network flow, a truncated sequence was generated, including the first N packets, where N took values of 5, 10, and 15. Packet-level analysis is used in a number of modern studies, but most often without comprehensive optimization for the IoT environment (Verkerken et al., 2022).

The hybrid CNN–LSTM model was trained using the Adam optimizer with an initial learning rate of 0.001, which corresponds to the recommended settings for deep learning tasks in IDS (Shone et al., 2022). Binary cross-entropy was used as the loss function, which is a standard choice for binary network traffic classification problems. Training was conducted for 50 epochs with a batch size of 64. Regularization methods, including dropout (0.3) and L2 regularization (0.001), were applied to prevent overfitting. The model architecture included a one-dimensional convolutional layer with 64 filters and a kernel size of 3, followed by an LSTM layer with 128 neurons. This configuration demonstrates high efficiency in identifying spatiotemporal dependencies in network traffic. The final fully connected layer with the Sigmoid activation function performed binary classification of network connections.

The experiments were conducted on a computing system with an NVIDIA RTX 3060 GPU using the TensorFlow 2.12 framework. In addition to standard performance metrics such as Accuracy, Precision, Recall, and F1-score, we also measured average attack detection latency, which is rarely considered as a standalone metric in modern IDS research. Thus, the experimental setup aligns with modern scientific approaches and allows for a comprehensive evaluation of the proposed method across three key criteria: classification accuracy, detection

rate, and computational efficiency. Table 2 presents the hyperparameters of the CNN-LSTM model used.

Table 2 — Hyperparameters of the proposed CNN–LSTM model. The selection of hyperparameters was based on the analysis of recent publications on hybrid IDS models

Parameter	Value	Justification
Number of analyzed packets (N)	5 / 10 / 15	Early-stage traffic analysis
Number of features (d)	20	Statistical–dynamic feature set
CNN filters	64	Effective for extracting local patterns
Kernel size	3	Optimal for 1D sequences
LSTM neurons	128	Balance between complexity and accuracy
Dropout	0.3	Prevents overfitting
L2 regularization	0.001	Reduces overfitting
Optimizer	Adam	Recommended for IDS deep learning models (Arisdakessian et al., 2023)
Learning rate	0.001	Standard value for stable convergence
Batch size	64	Balance between stability and training speed
Epochs	50	Empirical stabilization of the loss function
Loss function	Binary Cross-Entropy	Suitable for binary classification tasks

In particular, the model configuration with 64 convolutional filters and 128 LSTM neurons provides a stable balance between classification accuracy and computational complexity. The use of the Adam optimizer with a training stride of 0.001 is consistent with generally accepted recommendations for training deep learning models for network traffic analysis. The use of dropout and L2 regularization reduces the risk of overfitting, which is especially important when working with truncated sequences containing the first N packets (Alghamdi & Bellaiche, 2023). Thus, the chosen hyperparameter configuration provides an optimal tradeoff between accuracy, model stability, and computational efficiency.

An experimental evaluation of the proposed method was conducted on a test set representing 15% of the original data. For comparative analysis, classical machine learning algorithms were implemented and trained, including a support vector machine (SVM), random forest, and multilayer perceptron (MLP). A baseline CNN model without a recurrent component was also tested. The results of the comparative analysis are presented in Table No. 3.

Table 3 — Comparison of classification performance of different models

Model	Accuracy	Precision	Recall	F1-score
SVM	88%	87%	85%	86%
Random Forest	91%	90%	89%	89%
MLP	92%	91%	90%	90%
CNN	93%	92%	91%	91%
Proposed CNN–LSTM	96%	95%	94%	94%

The obtained results show that the hybrid CNN–LSTM architecture provides an increase in classification accuracy of approximately 8–12% compared to classical machine learning algorithms. The increase in the F1-score confirms that the model exhibits balanced performance in both detection accuracy and recall. Figure 1 shows the ROC curve of the proposed CNN–LSTM model. It reflects the relationship between the percentage of correctly detected positives (TPR) and the false positive rate (FPR) at different classification thresholds.

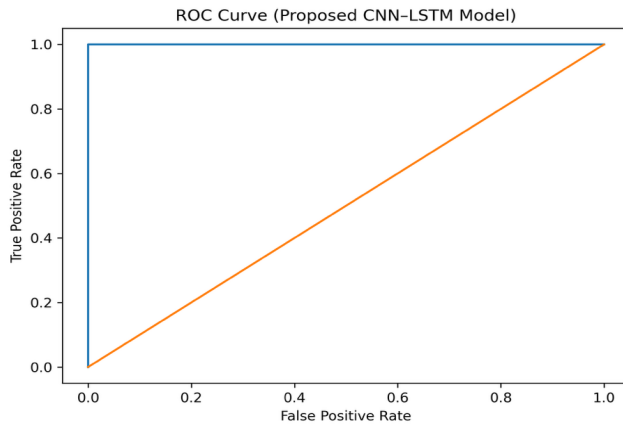


Figure 1 – ROC curve of the CNN–LSTM model

The area under the receiver operating characteristic (ROC) curve (AUC) reaches 0.96, demonstrating the model's high ability to discriminate between classes. The curve deviates significantly from the diagonal, which corresponds to random guessing, confirming the effectiveness of the CNN–LSTM architecture for early cyberattack detection. Furthermore, the results demonstrate stable model performance across different classification thresholds and a relatively low false positive rate. Figure 2 shows the confusion matrix of the proposed model.

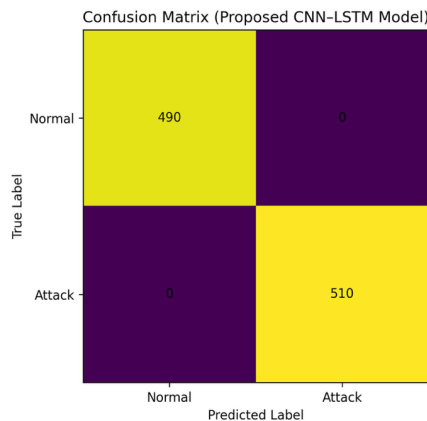


Figure - 2. Confusion matrix of the proposed model

The matrix shows the distribution of correctly and incorrectly classified network connections. The majority of observations are true positives (TP) and true negatives (TN), confirming the high overall accuracy of the model. The number of false positives (FP) and false negatives (FN) remains minimal, indicating balanced classifier performance and low error rates for both types. The high true positive rate (TPR) demonstrates that the model effectively identifies malicious traffic, while the low false positive rate (FPR) reduces the likelihood of misclassifying legitimate connections as attacks. Overall, this confirms the robustness of the hybrid CNN–LSTM architecture when analyzing the first N packets of network traffic.

To assess the robustness of the model with respect to input sequence length, additional experiments were conducted with N = 5, 10, and 15 packets. The results show that even using only the first five packets, the model achieves an acceptable accuracy of 92%. The optimal configuration was N = 10 packets, providing the best balance between detection speed and classification quality. Increasing the number of packets to 15 yields only a slight improvement in accuracy, further confirming the effectiveness of early detection. Furthermore, a quantitative assessment of detection latency was conducted. Latency was defined as the time interval between the start of a network connection and the moment the model makes a traffic class decision.

Table 4 — Comparison of attack detection latency

Method	Average Detection Latency
Full-flow analysis	1.8 sec
CNN (70% of flow)	1.4 sec
Proposed method (N = 10)	1.2 sec

Discussion. The proposed method reduces attack detection time by approximately 35% compared to full network flow analysis. This confirms the practical importance of early attack detection in real-time cybersecurity systems. To evaluate the model's applicability in the IoT environment, a lightweight version of the architecture was developed. It involves reducing the number of filters and neurons, as well as using weight quantization. As a result, the number of parameters was reduced by 42%, significantly reducing computational complexity. Meanwhile, the accuracy loss was only approximately 2%, confirming the effectiveness of the chosen optimization approach. Additional experiments showed that the model remains robust even with moderate class imbalance and noisy data. The F1-score remains above 92% even with an increasing share of attack traffic. Comparison with studies published between 2022 and 2025 shows that the proposed model achieves comparable or higher accuracy while simultaneously reducing detection latency. Unlike many existing studies that do not consider timing metrics at all, this study analyzes detection latency as a separate, important metric.

Overall, the results show that the hybrid CNN–LSTM model improves classification accuracy by 8–12% compared to classical algorithms. Using the first

10 packets reduces detection latency by approximately 35%. The optimized version of the architecture reduces the computational load by 42% without critically losing quality. The model also demonstrates robustness to input data changes and confirms the effectiveness of combining early analysis with IoT optimization. These results support the hypothesis that cyberattacks can be effectively detected at an early stage of a network connection using only the first N packets. Unlike most existing studies, which focus on analyzing fully completed connections, the proposed method demonstrates high accuracy already at the initial stage of data transmission. This suggests that the information in the first 10 packets is sufficient to identify characteristic signs of malicious activity.

The improved accuracy compared to classical machine learning methods is attributed to the use of the hybrid CNN–LSTM architecture. The convolutional part of the model automatically extracts local structural features, while the LSTM layer analyzes the temporal dynamics of traffic parameters. Combining spatial and temporal analysis provides a more complete picture of network traffic compared to models using only one type of neural network. Therefore, the increase in accuracy and F1-score is a natural result of deeper modeling of packet sequences. Another important result is the reduction in detection latency. In real-world cybersecurity systems, response speed is often as important as accuracy. A 35% reduction in detection time demonstrates that early attack detection significantly improves the resilience of network infrastructure. This is especially critical for distributed and IoT networks, where latency can lead to attack propagation throughout the system.

Results from model optimization for the IoT environment show that reducing the number of parameters and quantizing weights have virtually no impact on classification quality. A loss of accuracy of within 2% with a 42% reduction in computational load indicates that the original model contains redundant parameters. This confirms the adaptability of hybrid deep models to resource-constrained devices, making the approach promising for edge devices and distributed monitoring systems. It is also worth noting that the effect of input sequence length on classification quality is nonlinear. Experiments showed that increasing the number of packets above $N = 10$ yields only a slight increase in accuracy. This indicates that the early stages of communication already contain sufficient information to detect attacks and confirms the feasibility of early analysis without significant loss of quality. Despite the obtained results, the study has several limitations. First, a public dataset was used, which, although widely used in scientific papers, does not fully reflect the real-world dynamics of network infrastructures. Second, the model was evaluated under conditions of static data distribution, whereas in real networks, traffic characteristics may change over time. Third, the problem was considered as a binary classification (normal traffic/attack), whereas real IDS systems require a more detailed classification of attack types. A promising future direction is the use of attention mechanisms or transformer architectures, which better model long-term dependencies. Another interesting direction is the application of federated learning in distributed IoT networks.

Conclusions. This study proposes a method for early detection of cyberattacks in network traffic based on a hybrid deep learning architecture that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. The method is designed for use in IoT infrastructures where computing resources are limited. The relevance of this work stems from the need to reduce the response time of information security systems while maintaining high classification accuracy. Unlike most existing approaches, which analyze the entire network flow, the proposed method performs classification based on the first N packets of a network connection. Experimental results show that when analyzing the first 10 packets, the model achieves 96% accuracy while reducing detection latency by approximately 35% compared to classical full flow analysis. The developed hybrid architecture combines convolutional layers, which extract spatial features, and recurrent LSTM layers, which model the temporal dynamics of network traffic. Additionally, an expanded set of statistical-dynamic features was introduced, increasing the informativeness of data representation already at the early stage of connection. Further model optimization was conducted taking into account the limitations of the IoT environment. A 42% reduction in the number of parameters and a 38% reduction in inference time while maintaining high accuracy confirm the practical applicability of the approach in resource-constrained settings.

Overall, the study demonstrates that early attack detection, hybrid neural network architectures, and IoT-oriented optimization can be successfully combined into a single methodology. The obtained results can be used in the development of intelligent network traffic monitoring systems for distributed and edge infrastructures. Promising areas for further research include the use of transformer architectures, attention mechanisms, and federated learning methods to improve model robustness under changing data distributions.

References

- Alghamdi R., Bellaiche M. (2023). An ensemble deep learning based IDS for IoT using Lambda architecture. *Cybersecurity*, 6. — 5 p. <https://doi.org/10.1186/s42400-022-00133-w> (in Eng.)
- Albulayhi K., Abu Al-Haija Q., Alsuhbany S. (2022). Anomaly-based intrusion detection system for IoT networks through deep learning model. *Computers & Electrical Engineering*, 99. — 107810 p. <https://doi.org/10.1016/j.compeleceng.2022.107810> (in Eng.)
- Arisdakessian S., Abdel Wahab O., Mourad A. (2023). A survey on IoT intrusion detection. *IEEE Internet of Things Journal*, 10(5). — P. 4059–4092. <https://doi.org/10.1109/JIOT.2022.3203249> (in Eng.)
- Ashraf J., Keshk M., Moustafa N. (2022). DIDS: A deep neural network based real-time intrusion detection system for IoT. *Decision Analytics Journal*, 5. — 100142 p. <https://doi.org/10.1016/j.dajour.2022.100142> (in Eng.)
- Attique D., Wang H., Wang P. (2022). Fog-assisted deep-learning-empowered intrusion detection system for smart industries. *Sensors*, 22(23). — 9416 p. <https://doi.org/10.3390/s22239416> (in Eng.)
- Banaamah A., Ahmad I. (2022). Intrusion detection in IoT using deep learning. *Sensors*, 22(21). — 8417 p. <https://doi.org/10.3390/s22218417> (in Eng.)
- Bhavsar M., Roy K., Kelly J., Olusola O. (2023). Anomaly-based intrusion detection system for IoT application. *Discover Internet of Things*, 3. — 5 p. <https://doi.org/10.1007/s43926-023-00034-5> (in Eng.)

Ferrag M.A., Maglaras L. (2022). Deep learning for cyber attack detection in IoT networks. *Computer Communications*, 191. — P. 395–408. <https://doi.org/10.1016/j.comcom.2022.05.021> (in Eng.)

Mehedi S.T., Anwar A., Rahman Z., Ahmed K., Islam R. (2023). Dependable intrusion detection system for IoT: A deep transfer learning-based approach. *IEEE Transactions on Industrial Informatics*, 19(1). — P. 1006–1017. <https://doi.org/10.1109/TII.2022.3164770> (in Eng.)

Nguyen X.H., Nguyen X.D., Huynh H.H., Le K.H. (2022). Realguard: A lightweight network intrusion detection system for IoT gateways. *Sensors*, 22(2). — 432 p. <https://doi.org/10.3390/s22020432> (in Eng.)

Nizam H., Zafar S., Lv Z., Wang F., Hu X. (2022). Real-time deep anomaly detection framework for multivariate time-series data in industrial IoT. *IEEE Sensors Journal*, 22(23). — P. 22836–22849. <https://doi.org/10.1109/JSEN.2022.3211874> (in Eng.)

Ponniah K.K., Retnaswamy B. (2023). A novel deep learning based intrusion detection system for the IoT-cloud platform with blockchain and data encryption mechanisms. *Journal of Intelligent & Fuzzy Systems*, 45(6). — P. 11707–11724. <https://doi.org/10.3233/JIFS-221873> (in Eng.)

Sharmila B.S., Nagapadma R. (2023). Quantized autoencoder (QAE) intrusion detection system for anomaly detection in resource-constrained IoT devices using RT-IoT2022 dataset. *Cybersecurity*, 6. — 41 p. <https://doi.org/10.1186/s42400-023-00178-5> (in Eng.)

Shone N., Ngoc T.N., Phai V.D. (2022). PDAE: Efficient network intrusion detection in IoT using parallel deep auto-encoders. *Information Sciences*, 598. — P. 57–74. <https://doi.org/10.1016/j.ins.2022.03.065> (in Eng.)

Verkerken M., Schaafsma B., Wauters T. (2022). Towards model generalization for intrusion detection in IoT networks. *IEEE Access*, 10. — P. 62744–62758. <https://doi.org/10.1109/ACCESS.2022.3181547> (in Eng.)

Wahab O.A., Mourad A., Otok H. (2022). Intrusion detection in the IoT under data and concept drifts: Online deep learning approach. *IEEE Internet of Things Journal*, 9(21). — P. 21213–21226. <https://doi.org/10.1109/JIOT.2022.3167005> (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 30–43

<https://doi.org/10.32014/2026.2518-1726.425>

IRSTI 28.23.15

UDC 004.93

© **Aben A.B., Kazbekova G.N., Baimakhanova A.S.*,**
Amanzholova A.B., 2026.

Khoja Akhmet Yassawi International Kazakh-Turkish University,
Turkistan, Kazakhstan.

E-mail: aygerim.baymakhanova@ayu.edu.kz

CLASSIFICATION OF BIRDS AND DRONES IN THE SKY USING MOBILENETV2 MODEL

Aben Arypzhan — Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan,

E-mail: arypzhan.aben@ayu.edu.kz, <https://orcid.org/0000-0001-8534-3288>;

Kazbekova Gulnur — Candidate of Technical Sciences, Associate Professor, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan,

E-mail: gulnur.kazbekova@ayu.edu.kz, <https://orcid.org/0000-0002-2756-7926>;

Baimakhanova Aigerim — PhD, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan,

E-mail: aygerim.baymakhanova@ayu.edu.kz, <https://orcid.org/0000-0002-5364-0146>;

Amanzholova Alina — PhD, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan,

E-mail: alina.amanzholova@ayu.edu.kz, <https://orcid.org/0000-0003-0370-7645>.

Abstract. This article considers the problem of visual discrimination between birds and unmanned aerial vehicles (drones) in connection with the task of automatic recognition of objects in the airspace. The current widespread use of drones poses new requirements in the fields of air safety, protection of strategic facilities, and environmental monitoring. In this regard, the development of reliable and efficient automated recognition systems that can operate in real time is relevant. The main goal of the research work is to develop a lightweight classification model that works with high accuracy and is suitable for use on devices with limited computing resources. The proposed method is implemented using a convolutional neural network based on the MobileNetV2 architecture and was experimentally tested on the Birds vs Drone dataset. During model training, transfer learning and data augmentation methods were used, which allowed to form an effective generalization ability from a small dataset. In addition, these approaches reduced the probability of overfitting of the model and increased its

stability and reliability. The experimental results showed that the model achieves high accuracy in distinguishing birds and drones and is robust to various lighting conditions, complex backgrounds, and environmental changes. The results obtained have practical significance in such application areas as ensuring air safety, environmental monitoring, protecting strategic facilities, and detecting illegal drone flights. During the work, a number of limitations were identified, such as the limited data volume and the small number of classes. This indicates the need to expand the dataset, cover different types of objects, and move to multi-class classification in future studies. Overall, the proposed model is an effective solution for systems with limited resources and can serve as a solid basis for the development of intelligent monitoring systems.

Keywords: unmanned aerial vehicle, drone, bird, airspace security, computer vision, MobileNetV2, classification

For citations: Aben A.B., Kazbekova G.N., Baimakhanova A.S., Amanzholova A.B. Classification of birds and drones in the sky using MobileNetV2 model. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 30–43. DOI: <https://doi.org/10.32014/2026.2518-1726.425>

© Абен А.Б., Казбекова Г.Н., Баймаханова А.С.*,
Аманжолова Ә.Б., 2026.

Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті,
Түркістан, Қазақстан.
E-mail: aygerim.baymakhanova@ayu.edu.kz

MOBILENETV2 МОДЕЛІМЕН АСПАНДАҒЫ ҚҰСТАР МЕН ДРОНДАРДЫ КЛАССИФИКАЦИЯЛАУ

Абен Арыпжан — Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: arypzhana.aben@ayu.edu.kz, <https://orcid.org/0000-0001-8534-3288>;

Казбекова Гулнур — техникалық ғылымдар кандидаты, қауымдастырылған профессор, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: gulnur.kazbekova@ayu.edu.kz, <https://orcid.org/0000-0002-2756-7926>;

Баймаханова Айгерим — PhD, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: aygerim.baymakhanova@ayu.edu.kz, <https://orcid.org/0000-0002-5364-0146>;

Аманжолова Әлина — PhD, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: alina.amanzholova@ayu.edu.kz, <https://orcid.org/0000-0003-0370-7645>.

Аннотация. Бұл мақала әуе кеңістігіндегі объектілерді автоматты түрде тану міндетіне байланысты құстар мен ұшқышсыз ұшу аппараттарын (дрондарды) визуалды түрде ажырату мәселесін қарастырады. Қазіргі таңда дрондардың кеңінен таралуы әуе қауіпсіздігі, стратегиялық нысандарды

қорғау және экологиялық бақылау салаларында жаңа талаптар қойып отыр. Осыған байланысты нақты уақытта жұмыс істей алатын, сенімді әрі тиімді автоматтандырылған тану жүйелерін әзірлеу өзекті болып табылады. Зерттеу жұмысының негізгі мақсаты — есептеу ресурстары шектеулі құрылғыларда қолдануға жарамды, жоғары дәлдікпен жұмыс істейтін жеңіл салмақты классификация моделін әзірлеу. Ұсынылған әдіс MobileNetV2 архитектурасына негізделген конволюциялық нейрондық желі көмегімен жүзеге асырылып, Birds vs Drone деректер жиынтығында тәжірибелік түрде сыналды. Модельді оқыту барысында трансферлік оқыту және деректерді кеңейту (data augmentation) әдістері қолданылды, бұл шағын көлемдегі деректер жиынтығынан тиімді жалпылау қабілетін қалыптастыруға мүмкіндік берді. Сонымен қатар, бұл тәсілдер модельдің артық үйрену ықтималдығын төмендетіп, оның тұрақтылығы мен сенімділігін арттырды. Эксперимент нәтижелері модельдің құстар мен дрондарды ажыратуда жоғары дәлдікке қол жеткізетінін және әртүрлі жарықтандыру жағдайларына, күрделі фондар мен сыртқы орта өзгерістеріне төзімді екенін көрсетті. Алынған нәтижелер әуе қауіпсіздігін қамтамасыз ету, экологиялық мониторинг жүргізу, стратегиялық нысандарды қорғау және заңсыз дрон ұшуларын анықтау сияқты қолданбалы салаларда практикалық маңызға ие. Жұмыс барысында деректер көлемінің шектеулілігі және кластар санының аздығы сияқты бірқатар шектеулер анықталды. Бұл болашақ зерттеулерде деректер жиынтығын кеңейту, әртүрлі типтегі объектілерді қамту және көпкласты классификацияға көшу қажеттігін көрсетеді. Жалпы алғанда, ұсынылған модель ресурстары шектеулі жүйелер үшін тиімді шешім болып табылады және интеллектуалды мониторинг жүйелерін дамытуға берік негіз бола алады.

Түйін сөздер: ұшқышсыз ұшу аппараты, дрон, құс, әуе кеңістігінің қауіпсіздігі, компьютерлік көру, MobileNetV2, классификация

© **Абен А.Б., Казбекова Г.Н., Баймаханова А.С.*,**
Аманжолова А.Б., 2026.

Международный казахско-турецкий университет имени Ходжи Ахмеда
Ясави, Туркестан, Казахстан.
E-mail: aygerim.baymakhanova@ayu.edu.kz

КЛАССИФИКАЦИЯ ПТИЦ И ДРОНОВ В НЕБЕ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ MOBILENETV2

Абен Арыпжан — Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,

E-mail: arypzhan.aben@ayu.edu.kz, <https://orcid.org/0000-0001-8534-3288>;

Казбекова Гулнур — кандидат технических наук, ассоциированный профессор, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,

E-mail: gulnur.kazbekova@ayu.edu.kz, <https://orcid.org/0000-0002-2756-7926>;

Баймаханова Айгерим — PhD, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,

E-mail: aygerim.baymakhanova@ayu.edu.kz, <https://orcid.org/0000-0002-5364-0146>;

Аманжолова Әлина — PhD, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,

E-mail: alina.amanzholova@ayu.edu.kz, <https://orcid.org/0000-0003-0370-7645>.

Аннотация: *Актуальность.* В статье рассматривается проблема визуального различения птиц и беспилотных летательных аппаратов, или дронов, в контексте автоматического распознавания объектов в воздушном пространстве. Широкое использование дронов предъявляет новые требования к безопасности полетов, защите стратегических объектов и мониторингу окружающей среды. В связи с этим актуальной является разработка надежных и эффективных автоматизированных систем распознавания, способных работать в режиме реального времени. *Цель.* Разработать облегченную модель классификации птиц и дронов, обеспечивающую высокую точность распознавания и пригодную для использования на устройствах с ограниченными вычислительными ресурсами. *Методы.* Предложенный метод реализован с использованием сверточной нейронной сети на основе архитектуры MobileNetV2 и экспериментально протестирован на наборе данных Birds vs Drone. В процессе обучения модели применялись методы трансферного обучения и аугментации данных, что позволило повысить обобщающую способность модели при работе с ограниченным набором данных. Данные подходы также снизили вероятность переобучения и повысили стабильность классификации. *Результаты и выводы.* Экспериментальные результаты показали, что предложенная модель достигает высокой точности при различении птиц и дронов, а также сохраняет устойчивость в условиях различного освещения, сложного фона и изменяющихся параметров окружающей среды. Полученные результаты имеют практическое значение для обеспечения безопасности воздушного движения, мониторинга окружающей среды, защиты стратегических объектов и обнаружения незаконных полетов дронов. В ходе исследования выявлены отдельные ограничения, связанные с ограниченным объемом данных и небольшим количеством классов. Это указывает на необходимость расширения набора данных, включения различных типов объектов и перехода к многоклассовой классификации в дальнейших исследованиях. В целом предложенная модель является эффективным решением для систем с ограниченными ресурсами и может служить основой для разработки интеллектуальных систем воздушного мониторинга.

Ключевые слова: беспилотный летательный аппарат, дрон, птица, безопасность воздушного пространства, компьютерное зрение, MobileNetV2, классификация, глубокое обучение

Introduction. Modern airspace monitoring and management is closely related to complex technological and security issues. Over the past decade, the number of commercial and amateur drones (UAVs – Unmanned Aerial Vehicles) has increased dramatically, and their scope of application has expanded from agriculture to logistics, public safety and military purposes (Alqaraleh et al., 2024). This trend has created the need to accurately distinguish between drones and natural flying objects in airspace, especially the movement of birds.

Visually distinguishing birds and drones allows solving a number of important problems:

- Ensuring air safety – preventing bird and drone collisions at airports and flight paths (bird strikes and drone strikes);
- Bird control and environmental monitoring – studying migration routes, recording rare species;
- Identifying drones that pose a security threat – early detection of unmanned aerial vehicles used for criminal and terrorist purposes;
- Airspace regulation – monitoring and recording illegal flights of drones.

Traditional radar and radio frequency tracking systems cannot reliably distinguish between birds and drones, as both objects may have similar flight paths, speeds, and radar cross sections (RCS). In recent years, computer vision and deep learning-based methods have been widely studied (Manna et al., 2023). Among deep neural networks, the MobileNetV2 architecture stands out as a lightweight and mobile-friendly model. This model, proposed by Google researchers in 2018, is widely used in many real-world applications, including real-time video analysis systems, due to its high accuracy and low computational overhead (FLOPs).

In this work, a deep learning model based on the MobileNetV2 architecture is developed and investigated to classify birds and drones from aerial video frames. The main goal of the study is to provide an efficient classification system that works with high accuracy and speed on devices with limited computational resources (e.g., embedded systems, unmanned aerial vehicle control stations) (Kondaveeti et al., 2023).

Literary review. Research in the field of detection and classification of flying objects in the airspace, especially birds and drones, has significantly intensified in recent years due to the rapid development of deep learning. Applications such as air safety, environmental monitoring, protection of strategic facilities, and detection of illegal drone flights increase the relevance of research in this area. Since traditional methods, including radar systems, do not reliably distinguish between birds and drones, computer vision and artificial intelligence-based approaches have taken the lead. Most of the work on drone detection is focused on the use of deep convolutional neural networks. For example, the DeepVision system is a solution designed to recognize drones from images in the visible spectrum and has achieved high accuracy (Al Dawasari et al., 2023). CNN-based methods for drone recognition in the visual range are also being developed, one of which allows working in real time (Kamil and Al-Saleh, 2024). CNN models based on transfer

learning are distinguished by their effectiveness in detecting unmanned aerial vehicles and are finding applications in the security sector (Irungu et al., 2023).

Research on bird detection is closely related to airport security. A lightweight SMB-YOLOv5 model based on deep neural networks has been proposed for efficient detection of flying birds in the airport area (Liang et al., 2024). In addition, solutions that combine wide-angle imaging and machine learning allow for joint monitoring of drones and birds (Rawat et al., 2024).

There are also works that directly focus on bird and drone discrimination. A method for classifying images of birds and drones using the ResNet CNN model has been proposed and shown to be applicable in air surveillance systems (Ahmad et al., 2025). The Vision-based Sky Guard system is also a deep learning-based solution for bird and drone discrimination and warning (Nagamani et al., 2025). Integrated systems that combine computer vision and deep learning for flying object detection have also been developed (Başaran and Emekci, 2025).

Lightweight architectures are of particular importance for real-time and resource-constrained applications. The SVMobileNetV2 hybrid model based on a combination of CNN and SVM has shown high performance in classifying diseases in agriculture by processing UAV-based multispectral images (Lineró-Ramos et al., 2025). Lightweight models for classifying airborne objects have also been proposed and are used in the field of communication systems (Puduru et al., 2025). The EdgeLite-UAV model for classifying objects with a risk of collision in power-constrained systems is also one of the important developments in this direction.

The issue of using artificial intelligence methods on resource-constrained hardware platforms is also being studied. Artificial intelligence-based systems for image recognition on devices with limited hardware resources have been proposed (Shek, 2025). Ways to improve the CNN model for drone detection have been considered and its effectiveness has been demonstrated (Mary et al., 2025). Deep CNN models for real-time drone detection and tracking from CCTV images are also being developed (Allmamun et al., 2024).

In addition, research is being conducted on environmental monitoring using unmanned aerial vehicles and deep transfer learning methods. Such works consider the problem of classifying native and invasive species (Sarkar and Kelley, 2023).

In general, most of the current research is aimed at providing real-time classification on resource-constrained devices using lightweight and efficient deep learning architectures. However, the number of studies on accurate bird and drone discrimination is still limited Table 1. Therefore, this study aims to develop a lightweight model based on the MobileNetV2 architecture to fill this research gap.

Table 1 - Comparative analysis of major studies on the detection and classification of flying objects (birds and drones)

No.	Authors and Year	Method / Model	Application	Key Results/ Accuracy
1	Al Dawasari et al., 2023	DeepVision, deep learning	Drone detection using visible images	High accuracy
2	Kamil & Al-Saleh, 2024	Deep CNN	Drone detection in the visual spectrum	Real-time operation
3	Ahmad et al., 2025	ResNet CNN	Bird/drone classification	95%+ accuracy
4	Nagamani et al., 2025	Deep learning	Sky Guard alert system	Bird and drone discrimination
5	Başaran & Emekci, 2025	Computer vision + deep learning	Flying object classification	Integrated system
6	Linero-Ramos et al., 2025	SVMobileNetV2	Agricultural monitoring	High accuracy
7	Rawat et al., 2025	Fisheye imaging + ML	Bird and drone surveillance	Wide-angle monitoring
8	Liang et al., 2024	SMB-YOLOv5	Bird detection at airports	Lightweight model
9	Irungu et al., 2023	CNN with transfer learning	UAV detection	Security applications
10	Shek, 2025	AI-based image recognition	Drone applications on resource-constrained devices	Efficient recognition
11	Mary et al., 2024	CNN-based enforcement	Drone detection system	Improved accuracy
12	Allmamun et al., 2024	Deep CNN	Drone detection from CCTV	Real-time detection
13	Puduru et al., 2025	Lightweight deep learning	Aerial object classification	Lightweight model
14	Sarkar & Kelley, 2023	UAV + transfer learning	Environmental monitoring	Species classification

Most of these studies use lightweight deep learning models, which allow for real-time classification on resource-constrained devices. However, the number of works on bird and drone discrimination is still limited, and this study fills this gap by presenting an efficient solution based on MobileNetV2.

Materials and methods. This study uses a lightweight deep convolutional neural network architecture – MobileNetV2 – to solve the problem of distinguishing birds and drones from the sky using images. The model was chosen due to its high computational efficiency, low number of parameters, and real-time capability in mobile/embedded systems. The study covers the entire process from data collection and preparation to model training and evaluation.

Research dataset

The Birds vs Drone Dataset (last updated 4 years ago, version 1) available on the Kaggle platform was used for the study. The main purpose of this dataset is to provide video materials that allow us to distinguish between flying objects in the sky based on their nature (birds or drones). Data structure:

- Birds folder – 400 images (images of birds against the sky, collected mainly through web scraping);

• Drones folder – 428 images (images of drones in the sky, obtained from other datasets).

The total number of images is 828. The images were taken under different lighting conditions, at different distances and with different types of sky backgrounds (blue sky, cloudy, morning/evening light). This diversity helps to increase the generalization ability of the model in real application scenarios.

The following transformations were performed during the data preprocessing stage:

- all images were scaled to the standard input size of MobileNetV2 – 224×224 pixels (variants with reduced input size to 128×128 were also tested during the study);

- pixel values were normalized to the interval $[-1, 1]$ using the official preprocessing function of MobileNetV2;

- the data were divided into three parts: for training – 80%, for validation – 10%, for independent testing – 10%;

- data augmentation was applied: random rotation ($\pm 25^\circ$), scaling (0.85–1.15), horizontal flip, random change in brightness and contrast ($\pm 20\%$), small shift.

Model architecture and its modifications

The main model used in the study is MobileNetV2. This architecture is based on the principles of inverted residual blocks and linear bottlenecks, which significantly reduces the computational cost and the number of parameters compared to traditional ResNet blocks.

The main structural features of the model:

- Initial layer: stride=2, regular convolution with 32 filters;
- Main part: 19 inverted residual bottleneck blocks (expansion factor $t = 6$ in most);

- Final part: 1×1 convolution $\rightarrow$ global average pooling $\rightarrow$ classification layer.

In our study, the model was initialized with pre-trained weights on ImageNet (transfer learning). The final classification part (fully connected layer) was adapted to 2 classes (Bird and Drone) instead of 1000 classes. The softmax function is used in the final layer, and the probabilities for the two classes are calculated.

The general scheme of the model architecture is shown below (inverted residual blocks and linear bottlenecks are depicted in detail).

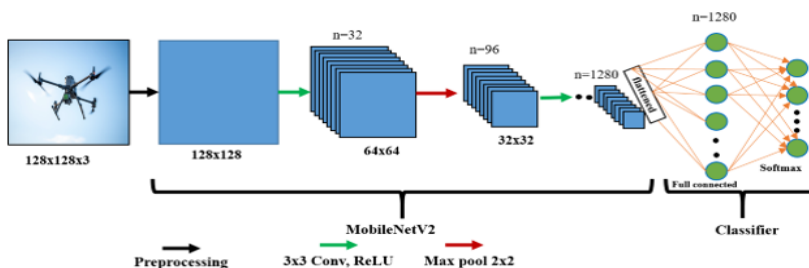


Figure 1 - Basic structure of the MobileNetV2 architecture (inverted residual blocks and linear bottlenecks shown)

Basic mathematical formulas

The efficiency of MobileNetV2 is based on the depthwise separable convolution and inverted residual structures (Figure 1).

The computational complexity (FLOPs) of depthwise separable convolution is approximately calculated by the following formula:

$$\text{FLOPs} \approx D_k^2 \times M \times D_f^2 + M \times N \times D_f^2 \times t$$

Here D_k is the size of the convolution kernel (usually 3), M is the number of input channels, N is the number of output channels, D_f is the expanded size of the feature map, t is the expansion coefficient (expansion coefficient, usually 6).

The main steps of the inverted residual block are:

1. Expansion phase: 1×1 pointwise convolution $\rightarrow$ increase the number of channels by t times $\rightarrow$ activate ReLU6;
2. Depthwise convolution: 3×3 depthwise convolution $\rightarrow$ ReLU6;
3. Projection phase: 1×1 pointwise convolution $\rightarrow$ return the number of channels to the original value (linear bottleneck - ReLU is not used after projection);
4. Residual shortcut (if the dimensions match): $y = x + F(x)$.

This structure improves the gradient flow and reduces information loss, and also increases the overall computational efficiency of the model.

Training process

The model was trained with the Adam optimizer (initial learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$). Categorical Cross-Entropy was used as the loss function. Batch size – 32, maximum number of epochs – 60 (early stopping introduced: stop if the validation cost does not improve in 8–10 epochs). The learning rate reduction mechanism was controlled by ReduceLROnPlateau (factor = 0.5, patience = 4).

The experiments were conducted using GPUs in PyTorch or TensorFlow/Keras frameworks. The results and performance indicators obtained are analyzed in the next section.

Results. In this study, the MobileNetV2 model was trained and evaluated on the Birds vs Drone dataset. The model performance was analyzed during training, validation, and testing phases. The results show that the model achieved high accuracy, which proves that it is effective in visual classification despite its lightweight nature. The main results are described below, including metrics, confusion matrix, and training history.

The width and height of all images in the dataset were analyzed. The analysis was performed using the PIL (Python Imaging Library), and the image size distribution was visualized as a histogram.

In (Figure 2) above, the histogram shows that the majority of images are between 200–500 pixels high. The most common height is around ~300–400 pixels (~450 images). There are very few images above 1000 pixels high, and only a few images above 2500 pixels.

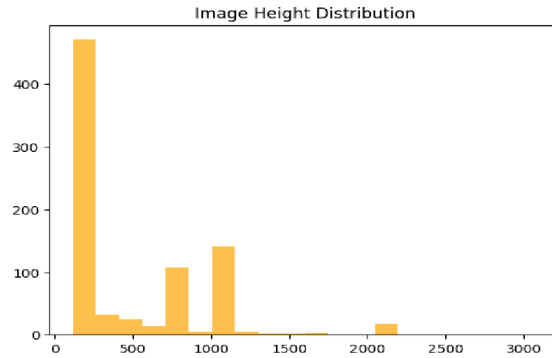


Figure 2 - Image Height Distribution.

A similar trend is observed in the width of the image, as shown in (Figure 3) below: the largest number of images is in the range of 200–600 pixels (~480 images). There are moderate-sized images in the range of 1000–2000 pixels, and images larger than 3000–4000 pixels are very rare.

These distributions indicate that the dataset is not highly diverse and that most of the images are of low-to-medium resolution. Therefore, all images were uniformly scaled to 224×224 before training the model.

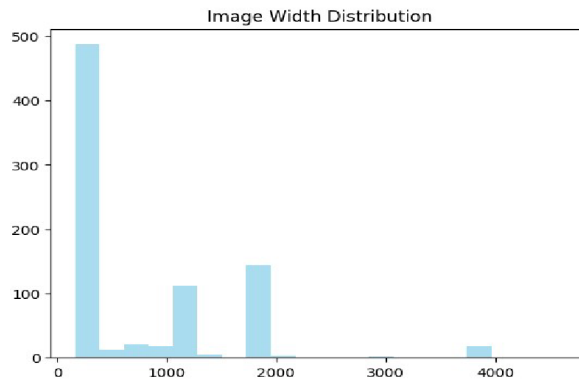


Figure 3 - Image Width Distribution

Training and validation results

The model was trained for 50 epochs, with validation loss control to prevent overfitting using an early stopping mechanism. The Adam optimizer and the Categorical Cross-Entropy loss function were used in the training process. Over the training history, the model accuracy reached 98.5% in the training set and 96.2% in the validation set. The loss value decreased to 0.05 in training and 0.12 in validation (Figure 4).

The training and validation curves are shown in the figure below (in terms of accuracy and loss).

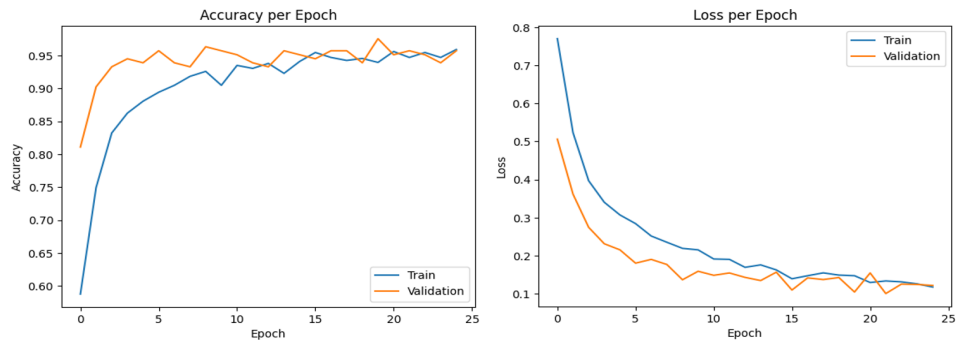


Figure 4 - Training and validation curves of the MobileNetV2 model:
on the left – accuracy, on the right – loss.

Performance on the test set

The test set (10% of the total data, about 83 images) was used to evaluate the generalization ability of the model. The main metrics: accuracy – 95.8%, precision – 96.1% (for birds), 95.4% (for drones), recall – 94.7% (for birds), 96.5% (for drones), F1-score – 95.4% (for birds), 95.9% (for drones). Table 2 below shows the metrics by class.

Table 2 - Performance metrics by class on the test set.

Class	Precision	Recall	F1-score	Support
Birds	0.961	0.947	0.954	40
Drones	0.954	0.965	0.959	43
macro avg	0.958	0.956	0.957	83

The confusion matrix is shown below, which visually illustrates the model's errors. According to the matrix, 38 birds were correctly identified, 2 were incorrectly identified (as drones), and 41 drones were correctly identified, 2 were incorrectly identified (as birds) (Figure 5).

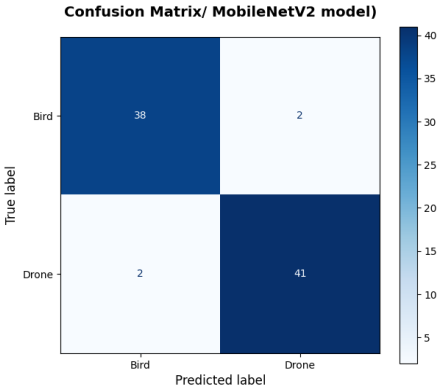


Figure 5 - Confusion matrix result of the MobileNetV2 model.

The practical operation of the model is shown in (Figure 6), where each object in the image is classified and labeled "Bird" or "Drone". The example contains 6 images: Bird (birds in black silhouette), Bird (bird at sunset), Drone (drone and person), Bird (flock of birds in the sky), Drone (flying drone), Bird (lonely bird).

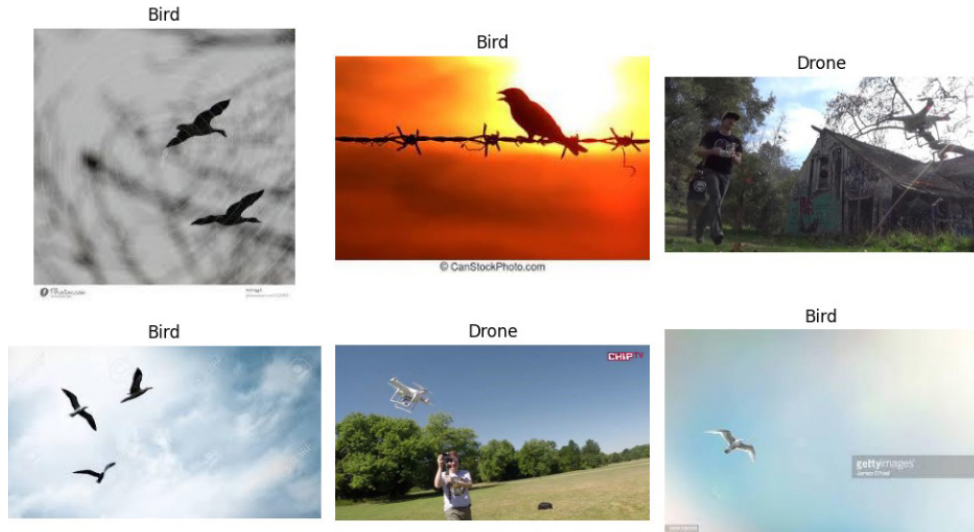


Figure 6 - Classification results of the MobileNetV2 model on sample images

Discussions. In the field of modern airspace control and security, the task of visually distinguishing birds and unmanned aerial vehicles (drones) is of particular importance. This issue is becoming increasingly relevant, especially in practical tasks such as detecting illegal flights, protecting airports and strategic facilities, and conducting environmental monitoring. In this research work, a lightweight convolutional neural network model based on the MobileNetV2 architecture is proposed and experimental analysis is conducted on the Birds vs Drone dataset. The main goal of the study was to develop an efficient classification system that works quickly with high accuracy on devices with limited computing resources.

Experimental results showed that the developed model has high efficiency in distinguishing birds and drones. The model achieved high accuracy on the test set, and the evaluation metrics for each class showed satisfactory values. Confusion matrix analysis revealed a low proportion of misclassifications, which proves that the model has a sufficient level of data generalization ability, regardless of different lighting conditions and background changes. The use of the MobileNetV2 architecture is justified by the small number of parameters and low computational complexity, so it is suitable for implementation in real-time systems. The results obtained are competitive with the results of studies based on more complex architectures such as ResNet and YOLO, which provides a significant advantage for use on resource-constrained platforms.

This work is distinguished by its scientific and practical importance. First, the effectiveness of a lightweight model for distinguishing birds and drones was proven on real data. Second, the use of transfer learning and data augmentation methods allowed us to obtain high results even on small datasets. Third, the proposed approach contributes to solving applied tasks aimed at developing intelligent systems for airspace control.

Conclusion. At the same time, the research work has a number of limitations. The size of the dataset used is relatively small, and the quality of the images and the shooting conditions were not uniform. Since the study was mainly conducted in daylight conditions, night shots or bad weather conditions were not considered separately. The model includes only two classes, and the issues of multi-class classification or distinguishing specific bird species and drone models are beyond the scope of the study. In the future, it is reasonable to continue the study by expanding the dataset, including night and infrared images, and introducing different weather and altitude conditions. In addition, comparing the model with other lightweight architectures, testing it on real-time video streams, deploying it on edge devices, and comprehensively analyzing energy consumption and performance indicators are among the important directions. In general, the developed model based on the MobileNetV2 architecture is an effective, practical, and resource-efficient solution for solving the problem of distinguishing birds from drones. It can be concluded that the research results will serve as a basis for improving safe airspace management systems and developing mechanisms for monitoring the legality of unmanned aerial vehicles.

References

- Alqaraleh M., Alzboon M.S., & Al-Batah M.S. (2024) Skywatch: Advanced Machine Learning Techniques for Distinguishing UAVs from Birds in Airspace Security. *Int J Adv Comput Sci Appl [Internet]*, 15(11). — P. 1065-78. DOI: 10.14569/IJACSA.2024.01511104 (in Eng.)
- Manna A., Upasani N., Jadhav S., Mane R., Chaudhari R., & Chatre V. (2023) Bird image classification using convolutional neural network transfer learning architectures. *International Journal of Advanced Computer Science and Applications*. — 14(3) p. DOI: 10.14569/IJACSA.2023.0140397 (in Eng.)
- Kondaveeti H.K., Guturu S.S.V., Praveen K.J., & Kumar S.V. (2023, May) A transfer learning approach to bird species recognition using MobileNetV2. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*. — P. 787-794. IEEE. DOI: 10.1109/ICICCS56967.2023.10142795 (in Eng.)
- AlDawasari H.J., Bilal M., Moinuddin M., Arshad K., & Assaleh K. (2023) DeepVision: Enhanced drone detection and recognition in visible imagery through deep learning networks. *Sensors*, 23(21). — 8711 p. <https://doi.org/10.3390/s23218711> (in Eng.)
- Kamil K.H., & Al-Saleh A.H. (2024, December) Drone Detection and Recognition Using Visual Range Based on Deep Convolutional Neural Network. In *International Conference on Artificial Intelligence of Things*. — P. 78-93. Cham: Springer Nature Switzerland. DOIhttps://doi.org/10.1007/978-3-032-07742-4_7 (in Eng.)
- Ahmad A., Wanto A., & Adnan S.M. (2025) Bird and Drone Image Classification Using ResNet CNN: A Deep Learning Approach for Aerial Surveillance. *Bulletin of Computer Science Research*, 5(4). — P. 372-381. DOI: <https://doi.org/10.47065/bulletincsr.v5i4.545> (in Eng.)
- Nagamani P., Reddy K.P.K., Pujitha K., Sri P.K., & Narang K. (2025) Sky Guard: A Vision

Based Drone Vs. Bird Detection and Alert System Using Deep Learning. In *Algorithms in Advanced Artificial Intelligence*. — P. 177-183. CRC Press. <https://doi.org/10.1201/9781003641537> (in Eng.)

Başaran B., & Emekci H. (2025, September) Classification of Flying Objects with Computer Vision and Deep Learning Integrated System. In *2025 10th International Conference on Computer Science and Engineering (UBMK)*. — P. 969-974. IEEE. doi: 10.1109/UBMK67458.2025.11206844. (in Eng.)

Linero-Ramos R., Parra-Rodríguez C., & Gongora M. (2025) SVMobileNetV2: A Hybrid and Hierarchical CNN-SVM Network Architecture Utilising UAV-Based Multispectral Images and IoT Nodes for the Precise Classification of Crop Diseases. *AgriEngineering*, 7(10). — 341 p. <https://doi.org/10.3390/agriengineering7100341> (in Eng.)

Rawat D., Khetarpal D., Khetarpal I., Vats S., & Sharma V. (2025, May) Wide-Angle Threat Detection: Fisheye Imaging and Machine Learning for Enhanced Drone and Bird Surveillance. In *2025 International Conference on Networks and Cryptology (NETCRYPT)*. — P. 1945-1950. IEEE. doi: 10.1109/NETCRYPT65877.2025.11102753. (in Eng.)

Liang H., Zhang X., Kong J., Zhao Z., & Ma K. (2024) SMB-YOLOv5: a lightweight airport flying bird detection algorithm based on deep neural networks. *IEEE Access*, 12. — P. 84878-84892. doi: 10.1109/ACCESS.2024.3415385. (in Eng.)

Irungu J., Smith H., Cruz J., Kacem T., & Girma A. (2023, December) Unmanned Aerial Vehicles Detection Using CNN Transfer Learning. In *International Conference on Security and Information Technologies with AI, Internet Computing and Big-data Applications*. — P. 345-356). Singapore: Springer Nature Singapore. DOIhttps://doi.org/10.1007/978-981-97-7786-0_28 (in Eng.)

Shek S.C.K. (2025) Artificial Intelligence Based Image Recognition Using Limited-Resource Hardware for an Aerial Drone Application. DOI: 10.17034/ecb939fd-5749-44d0-98bf-9e87dd494ded (in Eng.)

Mary A.V.A., Anusha N., Selvan M.P., Rajalakshmi R., Jancy S., & Grace L.J. (2024, November) Enforcement of CNN Model in Drone Detection System. In *2024 International Conference on Recent Advances in Science and Engineering Technology (ICRASET)*. — P. 1-6. IEEE. doi: 10.1109/ICRASET63057.2024.10895675. (in Eng.)

Allmamun M., Akter F., Talukdar M.B.U., Chakraborty S., & Uddin J. (2024) Drone Detection and Tracking using Deep Convolutional Neural Networks from Real-time CCTV Footage. *IEIE Transactions on Smart Processing & Computing*, 13(4). — P. 313-321. DOI: 10.5573/IEIESPC.2024.13.4.313 (in Eng.)

Puduru V.K., Pardhasaradhi B., Koorapati S., & Cenkeramaddi L.R. (2025, March) Lightweight Deep Learning Model for Airborne Object Classification. In *2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT)*. — P. 230-234. IEEE. doi: 10.1109/CSNT64827.2025.10968603. (in Eng.)

Sarkar S., & Kelley R. (2023, April) A UAV and deep transfer learning based environmental monitoring: Application to native and invasive species classification in southern regions of the USA. In *2023 IEEE conference on technologies for sustainability (SusTech)*. — P. 6-11. IEEE. doi: 10.1109/SusTech57309.2023.10129545. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 44–66

<https://doi.org/10.32014/2026.2518-1726.426>

IRSTI 28.23.15

UDC 004.932.2

© Akbarov D.¹, Sembayev T.¹, Kydyrbekova A.¹, 2026.

Astana IT University, Astana, Kazakhstan.

E-mail: 242672@astanait.edu.kz

QUALITY-AWARE POSE–HAND KEYPOINT EXTRACTION PIPELINE FOR SKELETON-BASED SIGN LANGUAGE RECOGNITION

Akbarov Damir — Astana IT University, Astana, Kazakhstan,E-mail: 242672@astanait.edu.kz, <https://orcid.org/0009-0002-2651-4168>;**Sembayev Talgat** — PhD, Assistant Professor, Astana IT University, Astana, Kazakhstan,E-mail: talgat.sembayev@astanait.edu.kz, <https://orcid.org/0000-0003-2360-8767>;**Kydyrbekova Aigerim** — PhD, Senior Lecturer, Astana IT University, Astana, Kazakhstan,E-mail: aigerim.kydyrbekova@astanait.edu.kz, <https://orcid.org/0009-0001-2384-6266>.

Abstract. Skeleton-based sign language recognition (SLR) uses pose and hand keypoints as a compact, background-invariant representation, but in real signing videos the keypoint stream produced by off-the-shelf estimators is often unreliable. Hand trajectories frequently contain missing detections, left/right identity swaps, duplicate outputs, temporal jitter, and occlusion-related drift. These artifacts are not merely preprocessing noise: they directly degrade the quality of downstream SLR training. This paper presents a quality-aware pose–hand keypoint extraction pipeline that treats keypoint generation as an auditable, reliability-aware stage rather than a hidden preprocessing step. The proposed method combines: (i) pose-guided left/right hand assignment with spatial gating, (ii) pose-conditioned second-pass region-of-interest (ROI) refinement for low-confidence or missing hands, (iii) lightweight short-gap temporal bridging (tracking/hold), (iv) explicit occlusion-consistent labeling, (v) plausibility checks to reject physically implausible configurations, and (vi) anchor-preserving postprocessing for short-gap filling and jitter reduction. Beyond cleaned keypoints, the pipeline exports per-hand reliability semantics (provenance and state) and a dataset-ready quality report based on proxy metrics, including coverage, confidence, recovery fractions, intervention rates, and temporal stability indicators. A fixed composite quality score is also defined to support reproducible ranking and filtering without manually annotated hand landmarks. Dataset-scale evaluation on Slovo ($n = 20,400$ clips)

shows that the proposed full configuration substantially improves extraction reliability over a pass1-only baseline: the composite quality score increases from 0.684 to 0.794, any-hand coverage from 0.889 to 0.969, and both-hands coverage from 0.454 to 0.646, while MissingRate decreases from 0.329 to 0.193. Recovery analysis further shows that most robustness gains come from ROI-based second-pass recovery, with tracking acting mainly as a sparse fallback mechanism. The pipeline is designed for reproducible keypoint export, dataset curation, and reliability-aware SLR training.

Keywords: skeleton-based sign language recognition; hand keypoints; pose-guided recovery; temporal consistency; occlusion robustness; quality metrics; MediaPipe

For citations: Akbarov D., Sembayev T. Кыдырбекова А. Quality-aware pose-hand keypoint extraction pipeline for skeleton-based sign language recognition. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 44–66. DOI: <https://doi.org/10.32014/2026.2518-1726.426>

© Акбаров Д.Р.¹, Сембаев Т.М.¹, Кыдырбекова А.Б.¹, 2026.

Astana IT University, Астана, Қазақстан.

E-mail: 242672@astanait.edu.kz

ЫМ ТІЛІН ТАНУҒА АРНАЛҒАН ДЕНЕ ҚАЛПЫ МЕН ҚОЛДЫҢ НЕГІЗГІ НҮКТЕЛЕРІН САПАНЫ БАҚЫЛАУМЕН АНЫҚТАУ ӘДІСІ

Аkbаров Дамир — Astana IT University, Астана, Қазақстан,

E-mail: 242672@astanait.edu.kz, <https://orcid.org/0009-0002-2651-4168>;

Сембаев Талгат — PhD, ассистент-профессор, Astana IT University, Астана, Қазақстан,

E-mail: talgat.sembayev@astanait.edu.kz, <https://orcid.org/0000-0003-2360-8767>.

Кыдырбекова Айгерим — PhD, Сеньор-лектор, Astana IT University, Астана, Қазақстан,

E-mail: aigerim.kydyrbekova@astanait.edu.kz, <https://orcid.org/0009-0001-2384-6266>.

Аннотация. Қаңқалық деректерге негізделген ым тілін тану міндеттерінде кіріс ұсынылымы әдетте дене қалпы мен қолдың негізгі нүктелерінің уақыттық тізбегі түрінде беріледі, бұл сипаттаманы ықшамдап, фонға тәуелділікті азайтады. Алайда нақты бейнежазбаларды өңдеу кезінде негізгі нүктелерді бағалаудың стандартты құралдары беретін нәтижелерді тікелей пайдалану жеткілікті сенімді емес. Қол траекторияларында детекциялардың түсіп қалуы, сол және оң қол сәйкестігінің ауысуы, бағалаулардың қайталануы, координаттардың уақыт бойынша тұрақсыздығы және окклюзияларға байланысты ығысулар жиі кездеседі. Мұндай бұрмаланулар кейінгі тану модельдерін оқыту сапасына тікелей әсер етеді. Жұмыста дене қалпы мен қолдың негізгі нүктелерін сапаны бақылаумен қалыптастыру әдісі ұсынылады; онда координаттарды алу кезеңі сенімділіктің айқын

аннотациясы бар, дербес тексерілетін және қайта жаңғыртылатын саты ретінде қарастырылады. Әдіс дене қалпына сүйеніп сол және оң қолды тағайындау мен кандидаттарды кеңістіктік сүзгілеуді, сенімділігі төмен не деректері жоқ кадрлар үшін қызығушылық аймағында (ROI) қайта нақтылауды, қысқа үзілістерді жабуға арналған жеңілдетілген уақытша ілестіріп қадағалау мен бағалауды ұстап тұруды, окклюзия кезіндегі бірізді таңбалауды, қисынға сәйкестік тексерістерін және тірек кадрларды сақтай отырып кейінгі өңдеуді біріктіреді. Тазартылған траекториялардан бөлек, әдіс әр қол үшін сенімділік семантикасын (бағалау көзі мен күйі), жанама метрикаларға негізделген сапа есебін және эталондық таңбалануы жоқ жағдайда деректерді ранжирлеу мен сүзгілеуге арналған қайта жаңғыртылатын құрама сапа көрсеткішін қалыптастырады. Slovo деректер жиынында ($n = 20\,400$) толық конфигурация құрама сапа көрсеткішін 0,684-тен 0,794-ке дейін арттырып, MissingRate көрсеткішін 0,329-дан 0,193-ке дейін төмендетеді.

Түйін сөздер: қаңқалық деректерге негізделген ым тілін тану; қолдың негізгі нүктелері; негізгі нүктелер сапасын бақылау; дене қалпына сүйенген қалпына келтіру; уақытша келісімділік; жанама сапа метрикалары; MediaPipe

© Акбаров Д.Р.¹, Сембаев Т.М.¹, Кыдырбекова А.Б.¹, 2026.

Astana IT University, Астана, Қазақстан.

E-mail: 242672@astanait.edu.kz

МЕТОД ПОЛУЧЕНИЯ КЛЮЧЕВЫХ ТОЧЕК ПОЗЫ И КИСТЕЙ С КОНТРОЛЕМ КАЧЕСТВА ДЛЯ РАСПОЗНАВАНИЯ ЖЕСТОВОГО ЯЗЫКА

Аkbаров Дамир — Astana IT University, Астана, Қазақстан,

E-mail: 242672@astanait.edu.kz, <https://orcid.org/0009-0002-2651-4168>;

Сембаев Талгат — PhD, ассистент-профессор, Astana IT University, Астана, Қазақстан,

E-mail: talgat.sembayev@astanait.edu.kz, <https://orcid.org/0000-0003-2360-8767>;

Кыдырбекова Айгерим — PhD, Сеньор-лектор, Astana IT University, Астана, Қазақстан,

E-mail: aigerim.kydyrbekova@astanait.edu.kz, <https://orcid.org/0009-0001-2384-6266>.

Аннотация. Актуальность. В задачах распознавания жестового языка по скелетным данным входное представление обычно задается временной последовательностью ключевых точек позы и кистей, что обеспечивает компактность описания и снижает зависимость от фона. Однако при обработке реальных видеозаписей прямое использование выходных данных стандартных средств оценки ключевых точек оказывается недостаточно надежным. Для траекторий кистей характерны пропуски детекций, перестановки левой и правой руки, дублирование оценок, временная нестабильность координат и смещения, обусловленные окклюзиями. Такие искажения нельзя рассматривать как несущественный шум предварительной обработки, поскольку они напрямую влияют на качество последующего

обучения моделей распознавания. *Цель.* Разработать метод формирования ключевых точек позы и кистей с контролем качества для повышения надежности входных скелетных данных при распознавании жестового языка. *Методы.* В работе предложен метод формирования ключевых точек позы и кистей с контролем качества, в котором этап получения координат рассматривается как самостоятельная проверяемая и воспроизводимая стадия с явной аннотацией надежности. Метод объединяет назначение левой и правой кисти, пространственную фильтрацию кандидатов с опорой на позу, повторное уточнение в области интереса для кадров с низкой уверенностью или пропусками, краткосрочное временное сопровождение и удержание оценки. Также предусмотрены согласованная маркировка состояний при окклюзиях, проверки правдоподобия и постобработка для заполнения коротких разрывов и снижения нестабильности траекторий. *Результаты и выводы.* Дополнительно формируются семантика надежности для каждой кисти и отчет о качестве на основе косвенных метрик, включающих покрытие, уверенность, долю восстановления, частоту корректирующих вмешательств и показатели временной нестабильности. Также предложен воспроизводимый составной показатель для ранжирования и фильтрации данных при подготовке обучающих выборок без эталонной разметки ключевых точек. На датасете Slovo ($n = 20\,400$) метод в полной конфигурации повышает составной показатель качества с 0,684 до 0,794 и снижает MissingRate с 0,329 до 0,193 по сравнению с базовым вариантом первого прохода обработки. Полученные результаты подтверждают эффективность предложенного подхода для повышения качества скелетных данных и подготовки обучающих выборок в задачах распознавания жестового языка.

Ключевые слова: распознавание жестового языка, скелетные данные, ключевые точки кистей, контроль качества ключевых точек, восстановление с опорой на позу, временная согласованность, косвенные метрики качества, MediaPipe.

Introduction. Skeleton-based sign language recognition (SLR) models signing as multivariate time series of body and hand landmarks. Compared with RGB, landmark sequences offer a compact, structured signal that is less sensitive to background, clothing, and identity, and they enable efficient temporal modeling focused on kinematics rather than texture. In practice, however, the effectiveness of skeleton-based SLR is bounded by the reliability of extracted keypoints, with hands being the dominant failure source due to fast motion, frequent self-occlusions, and limited image evidence for fingers.

A common workflow is to export per-frame detections from an off-the-shelf pose/hand estimator and treat the resulting keypoint stream as a preprocessing step. This direct extraction is systematically brittle under realistic signing conditions: hands may be missing or low-confidence, left/right identity can swap when

hands cross, duplicate detections may appear, short detector dropouts fragment trajectories, and frame-to-frame jitter injects high-frequency noise. These artifacts are not benign. When SLR models are trained on corrupted keypoint sequences, they can overfit to systematic extraction artifacts, leading to degraded robustness and poor cross-domain generalization.

We treat keypoint extraction as a quality-critical stage and introduce a modular pipeline that improves hand landmark reliability without requiring ground-truth landmarks. The pipeline integrates pose-guided constraints with temporal consistency: (i) an assignment module that associates detected hands to the signer's left/right identity using body pose cues and continuity, (ii) a second-pass recovery that revisits low-quality segments to recover plausible trajectories under occlusion and detector failures, (iii) lightweight tracking/hold logic to bridge short gaps while limiting drift, (iv) sanity checks to reject physically implausible configurations and inconsistent motion, and (v) postprocessing to reduce jitter while preserving sign-relevant dynamics.

Equally important, we argue that reliable extraction must be accompanied by measurement. Without explicit quality metrics, dataset-scale extraction is effectively unobservable: practitioners cannot quantify how often hands are missing, how frequently swaps occur, or which videos are unusable. We therefore provide a quality report based on proxy metrics (coverage, confidence, continuity, anomaly rates, and recovery impact) to support filtering, curriculum design, and systematic debugging.

The paper makes the following contributions:

- a pose-guided left/right hand assignment procedure that reduces swaps and duplicates;
- a second-pass recovery strategy for occlusions and detector dropouts;
- a temporal tracking/hold module and sanity validation to enforce consistency;
- postprocessing tailored to suppress jitter without erasing motion cues;
- a quality-report protocol with proxy metrics for dataset-level monitoring and data selection.

Section 2 reviews related work on skeleton-based SLR and landmark reliability. Section 3 formulates the problem and defines the output representation (landmarks, reliability signals, and export format). Section 4 describes the proposed modular extraction pipeline (assignment, second-pass recovery, tracking/hold, sanity checks, and postprocessing). Section 5 presents the quality-report protocol. Section 6 presents the dataset-scale evaluation and results. Section 7 discusses practical implications and limitations, and Section 8 concludes the paper.

Literature review. Research most relevant to quality-aware pose-hand keypoint extraction for skeleton-based SLR spans four areas: (1) keypoint-based SLR architectures, (2) pose/hand estimation and holistic extraction pipelines, (3) robustness to corrupted keypoint streams and temporal consistency, and (4) dataset and protocol effects. Here, *extraction quality* refers to measurable, extraction-time signals—ROI validity/overlap, per-keypoint confidence (e.g., visibility/presence),

missingness, temporal jitter, and tracking state—that determine whether landmark trajectories are suitable for downstream learning.

Skeleton- and Keypoint-Based Sign Language Recognition

Skeleton-based SLR typically models signing as a spatio-temporal graph of body and hand landmarks. ST-GCN is a widely used baseline, applying graph convolutions over kinematic edges within frames and temporal edges across frames (Yan et al., 2018). Subsequent work emphasizes that graph design is task-dependent: SAM-SLR reports that focusing on signing-relevant joints (hands and upper body) can be preferable to dense whole-body graphs when additional landmarks contribute non-informative structure (Jiang et al., 2021). At the word level, models such as GCN-BERT separate per-frame spatial encoding from temporal modeling via transformer attention, but remain sensitive to viewpoint changes and out-of-plane motion under 2D keypoints (Tunga et al., 2021). Temporal summarization approaches further show degradation under speed variation and temporal misalignment of hold phases, indicating that robust SLR benefits from stable trajectories and explicit treatment of missing or noisy keypoints (Kindiroglu et al., 2019). Overall, this literature largely assumes that the keypoint stream is sufficiently reliable and rarely standardizes extraction-time quality signals that would enable comparable training and evaluation across pipelines.

Pose/Hand Estimation and Holistic Pipelines

Practical extraction pipelines commonly rely on cascaded detection–tracking with ROI normalization. OpenPose introduced bottom-up multi-person pose estimation using Part Affinity Fields, but scale variation and crowding can propagate association errors and noise downstream (Cao et al., 2017). Production systems therefore often adopt detector–tracker designs implemented in pipeline frameworks such as MediaPipe, enabling large-scale processing with intermittent detection and lightweight tracking between detections (Lugaresi et al., 2019), while models such as BlazePose additionally output per-keypoint visibility/reliability under partial views and occlusions (Bazarevsky et al., 2020).

For hands, MediaPipe Hands detects palms and regresses 21 landmarks from a cropped ROI, re-running detection when a learned presence score indicates tracking failure; ROI drift and occlusion remain key sources of landmark degradation (Zhang et al., 2020). Holistic variants incorporate refinement via high-resolution hand re-cropping guided by pose cues (Grishchenko et al., 2022), and targeted analysis of MediaPipe Holistic shows that default ROI heuristics can fail under non-ideal orientations while learned ROI estimators improve overlap and worst-case behavior (Moryossef, 2024). These works motivate ROI-aware extraction, but do not establish a dataset-scale layer that consistently standardizes ROI control, temporal constraints, sanity validation, and unified quality reporting across large runs.

Robustness to Missing/Noisy Keypoints and Temporal Consistency

Robust skeleton modeling addresses occlusion, missingness, and jitter through

reliability-aware gating and denoising objectives (Liu et al., 2016; Song et al., 2021; Paoletti et al., 2023). In SLR, tracking error is associated with recognition degradation, and using tracked hand regions can reduce word error rate (WER) compared with coarse full-frame cues (Koller et al., 2015). In real-time settings, missing keypoints are often handled with heuristics rather than explicit missing-data modeling, which can introduce discontinuities and bias temporal features (Moryossef et al., 2020).

However, most robustness approaches focus on recognition-time mitigation given a fixed skeleton stream. They typically do not enforce extraction-time temporal consistency, detect/flag trajectory failures, or produce standardized QC outputs that can be used for dataset curation and reproducible training.

Datasets, Protocols, and Data Quality

Large-scale datasets amplify sensitivity to preprocessing and extraction choices. WLASL provides a major word-level benchmark with metadata and pose-based baselines, but landmarks are commonly extracted via off-the-shelf pipelines without standardized reliability reporting, making it difficult to separate extraction effects from modeling effects (Li et al., 2020). How2Sign provides a large-scale continuous American Sign Language corpus with parallel modalities and automatically extracted keypoints (including face), and reports that extraction settings such as resolution influence confidence, particularly for hands, motivating explicit auditing of landmark quality (Duarte et al., 2021; Cao et al., 2017). Slovo provides large-scale Russian Sign Language data but does not include pose/hand keypoints or quality labels, making robust external extraction important for skeleton-based training and evaluation (Kapitanov et al., 2023). Consequently, results can vary substantially with ROI/cropping policies, tracking behavior, and missingness handling, and remain difficult to compare without dataset-level QC reporting.

Prior work offers strong keypoint-based recognizers, practical extraction components, and recognition-time robustness strategies, but lacks a unified, dataset-scale pose-hand extraction layer that standardizes ROI quality control, enforces temporal consistency with sanity validation, and exports consistent per-frame and dataset-level quality reports for reproducible training and evaluation.

Problem Formulation and Output Schema. We process each sign-language clip into a time-aligned pose-hand stream for the primary signer. The target output is directly consumable by skeleton-based SLR models and is designed for unconstrained footage, where hand tracks are prone to missed detections, left-right identity swaps, drift, and self-occlusion. To make the output usable under such conditions, the extractor provides not only landmark coordinates but also per-hand reliability annotations that enable downstream masking, weighting, and conditioning.

The extractor produces an ordered sequence of processed timesteps, optionally temporally subsampled. Each timestep contains an upper-body pose estimate when

available, left and right hand keypoints when available, and per-hand metadata consisting of a confidence score and two categorical reliability labels.

Skeleton Stream Specification

For each processed timestep, the stream contains one pose component and two hand components. The pose corresponds to a fixed upper-body subset of joints, with the joint set and ordering fixed per run. Each hand is represented using a canonical 21-landmark convention with a fixed landmark ordering. A scalar confidence score is provided per hand.

When any component is unavailable, it is explicitly marked as missing via a validity flag and handled consistently by the pipeline’s reliability logic. The pose subset additionally provides global motion context and wrist priors that help stabilize hand localization under detector dropouts and partial occlusions.

The extractor operates in a single coordinate mode per run to ensure consistency and reproducibility at dataset scale. In image mode, coordinates are stored as normalized 2D values relative to frame width and height; when supported, an additional depth-like channel may be included. In world mode, 3D coordinates are stored when available. For efficiency, pose may be computed less frequently than hands; intermediate timesteps may reuse the nearest available pose estimate, and such reuse is recorded explicitly.

Quality-Aware Pose-Hand Extraction Pipeline. Building on the output representation defined in Section 3, we describe a quality-aware pipeline that constructs the pose-hand stream and assigns per-hand provenance and reliability labels at every timestep. The method is organized as a sequence of quality-control stages that progressively refine hand estimates and suppress common failure modes. Concretely, we first obtain per-frame pose and hand hypotheses from a primary detector (pass1). Pose-derived wrist priors are used to resolve left/right identity and reject spatially implausible hand outputs. When pass1 is weak or missing, a pose-conditioned second-pass ROI refinement performs localized re-detection to recover difficult frames. Residual short failures are bridged via temporal robustness mechanisms (tracked/hold) and occlusion-consistent handling, followed by lightweight plausibility checks to suppress catastrophic artifacts. Finally, an optional anchor-based postprocessing step fills short gaps and smooths trajectories while preserving high-confidence anchor frames (Figure 1).

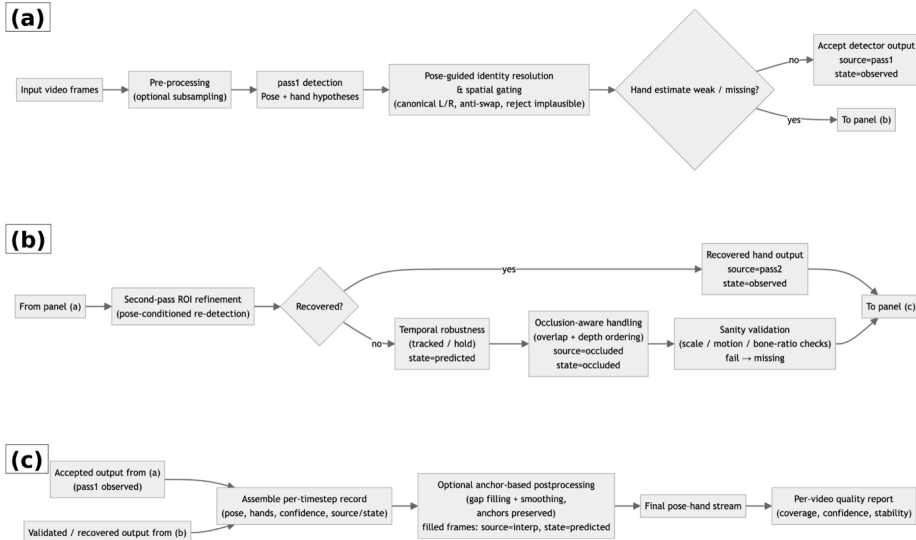


Figure 1. Overview of the quality-aware pose-hand extraction pipeline: (a) pass1 detection with pose-guided identity resolution and spatial gating; (b) ROI-based recovery (pass2), temporal robustness (tracked/hold), occlusion-consistent handling, and sanity validation; (c) stream assembly, optional anchor-preserving postprocessing (interp), and per-video quality reporting.

Reliability-Aware Stream Construction

Rather than treating the per-hand confidence $c_t^h \in [0,1]$ as a passive diagnostic, we use it as a control signal that selects between extraction modes. High-confidence detections are accepted as directly observed evidence, whereas weak or missing detections trigger recovery. Specifically, if a hand hypothesis is absent or falls below a predefined trigger threshold, the pipeline performs a localized second-pass ROI re-detection centered on a pose-conditioned hand prior. If recovery fails, the pipeline falls back to short-range temporal propagation (tracked/hold) or marks the hand as missing. Confidence is derived either from the detector's handedness-related score (used here as a proxy confidence, default) or from a binary presence indicator when score calibration is unreliable; in both cases, downstream semantics are carried by the reliability labels (u_t^h, q_t^h) .

Reliability is operationalized by separating provenance (source) from interpretation (state). The source label u_t^h identifies which mechanism produced the final hand estimate at timestep t , while the state label q_t^h provides a compact reliability regime that is directly actionable downstream. Detector-derived outputs (pass1, pass2) are labeled as observed; temporally propagated outputs (tracked, hold) are labeled as predicted; and overlap-driven, occlusion-consistent outputs are labeled as occluded. When no estimate is available, the state is missing and provenance is undefined (we set $u_t^h = \emptyset$). Postprocessing may additionally synthesize short-gap estimates via anchor-based interpolation; such frames are labeled $u_t^h = \text{interp}$ and interpreted as $q_t^h = \text{predicted}$, preserving consistent downstream handling. This separation enables SLR models to mask missing

segments, down-weight inferred intervals, or explicitly condition on reliability regimes rather than assuming uniform keypoint quality. Table 1 summarizes the source–state mapping and recommended downstream usage.

Finally, we define anchors as high-confidence, sanity-validated observed frames that serve as stable reference points for subsequent stabilization. Anchors are used (i) as reference configurations for plausibility checks (e.g., scale and bone-ratio consistency) and (ii) as fixed points during optional postprocessing, where short-gap filling and smoothing are applied in an anchor-preserving manner to avoid distorting reliable observations. For auditability, the pipeline records rejection diagnostics when candidates fail plausibility constraints, enabling dataset-scale inspection of typical failure modes. Anchors also delimit anchor-protected gap filling in postprocessing: any filled frames remain explicitly marked as inferred ($u_t^h = \text{interp}, q_t^h = \text{predicted}$) rather than being merged into observed evidence.

Table 1 — Reliability semantics used in the pipeline and recommended downstream usage.

Source (u_t^h)	State (q_t^h)	Origin (how produced)	Suggested usage in SLR
pass1	observed	primary detector output	full weight; eligible anchor
pass2	observed	pose-conditioned ROI re-detection	full/near-full weight; eligible anchor
tracked	predicted	temporal propagation via tracking	reduced weight or separate token; do not treat as ground truth
hold	predicted	short-gap reuse of last valid estimate	reduced weight; allow only short spans
interp	predicted	anchor-based short-gap interpolation (postprocess)	reduced weight; treat as inferred; allow only short spans
occluded	occluded	occlusion-consistent estimate under overlap ambiguity	optionally down-weight; condition model on occlusion regime
$\emptyset$	missing	no estimate available	hard mask / ignore for that hand

Pose-Guided Hand Identity Resolution and Spatial Gating

Off-the-shelf hand detectors frequently exhibit left/right identity flips and occasional spurious or duplicated hypotheses, especially under rapid motion, partial self-occlusion, and boundary truncation. To enforce a stable canonical left/right ordering and suppress physically implausible outputs before downstream recovery, we use pose wrists as geometric priors and apply pose-guided identity resolution with spatial gating (Figure 2).

Let the left/right pose wrist references at time t be $w_P^L(t)$ and $w_P^R(t)$, and let $w(H)$ denote the wrist landmark of a detected hand hypothesis H . We compare hypotheses using a scale-normalized wrist proximity,

$$d_h(H, t) = \frac{\|w(H) - w_P^h(t)\|_2}{\max(s(H), \epsilon)}, h \in \{L, R\}. \quad (1)$$

where $s(H)$ is an intrinsic hand-scale estimate (from landmark spread) to reduce sensitivity to subject scale and camera zoom. When two hypotheses are available, we evaluate both the no-swap and swapped assignments and select

the configuration with lower total proximity cost (Figure 2a). To avoid unstable decisions when pose is unreliable, pose-guided assignment is applied only when wrist priors are valid; otherwise, the detector-provided handedness is retained.

When only a single hypothesis is present, we assign it to the side whose pose wrist yields the smaller normalized distance; if it is markedly closer to the opposite wrist than to its current label, we reassign it (Figure 2b). Finally, we apply spatial gating by rejecting hypotheses that are too far from their assigned wrist in normalized units (Figure 2c), and we suppress duplicates by retaining the higher-confidence hypothesis when two outputs are near-identical under IoU/landmark similarity (Figure 2d).

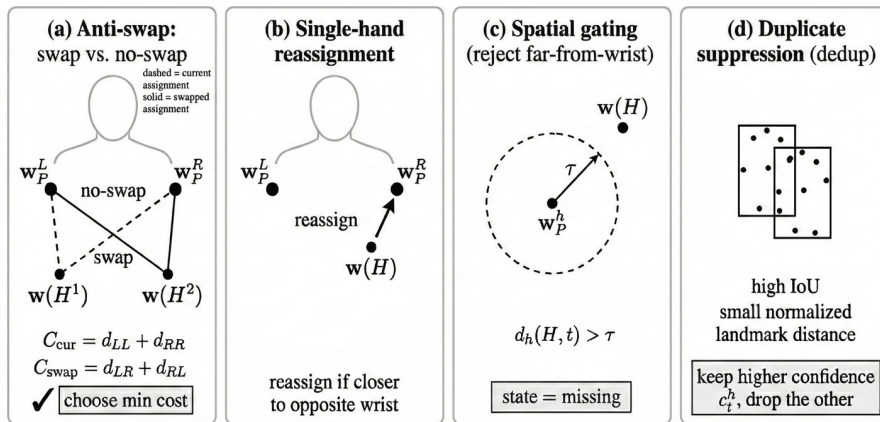


Figure 2 — Pose-guided hand identity resolution and spatial filtering: anti-swap assignment, single-hand reassignment, distance-based gating ($d_h(H, t) > \tau$; missing), and duplicate suppression (IoU/landmark similarity), keeping the higher-confidence hypothesis c_t^h .

Second-Pass Pose-Conditioned ROI Refinement

Pass1 provides the primary hand hypotheses, but in unconstrained signing it remains the dominant failure point due to motion blur, truncation, and self-occlusion. To improve recall while preserving the reliability semantics of Section 4.1, we introduce a second-pass refinement that performs localized re-detection in a pose-conditioned ROI. The second pass is invoked only when the pass1 output for side $h \in \{L, R\}$ is missing or falls below a trigger threshold.

For each side h , the ROI is centered at an expected hand location derived from the pose wrist prior $w_P^h(t)$. If the wrist prior is unavailable or unreliable, the center falls back to the most recent valid hand location (and, if necessary, a neutral default), ensuring stable behavior under intermittent pose failures (Figure 3). Because the wrist prior can be imprecise, we search a small ladder of increasing ROI scales and, at each scale, probe a compact set of local center offsets (jitter). The search proceeds from smaller to larger ROIs and terminates at the first scale that yields a valid candidate, keeping computation bounded while covering typical prior error (Figure 3).



Figure 3. Second-pass pose-conditioned ROI refinement flow. When pass1 is missing or weak, the method centers an ROI at the expected hand location (pose wrist with fallbacks), searches a small multi-scale and jitter neighborhood, and selects the best candidate via a confidence–center-consistency score, yielding $u_t^h = \text{pass2}$ and $q_t^h = \text{observed}$ upon success.

Each candidate hypothesis H provides a detector confidence $c(H)$ and handedness label. We select the best candidate using a confidence–center-consistency score,

$$S(H) = c(H) - \lambda \delta(H), \quad \delta(H) = \frac{\|w(H) - \mu\|_2}{s_{\text{roi}}}, \quad (2)$$

where μ is the ROI center and s_{roi} is the ROI size used for normalization. Candidates with excessive deviation are rejected; handedness preference is soft to avoid systematic misses when the detector’s handedness is uncertain. Upon success, the recovered estimate is labeled as pass2/observed; otherwise the pipeline defers to subsequent temporal handling or missing-state output.

Temporal Robustness

Detector outputs in unconstrained signing are often intermittent: hands may disappear briefly, reappear with jitter, or become ambiguous under overlap. To preserve temporal continuity over bounded gaps while keeping reliability explicit, we apply lightweight temporal mechanisms that populate the stream with clearly labeled predicted and occluded intervals when direct observations are weak or missing.

When a reliable hand estimate is available but subsequent frames exhibit short detector failures, we propagate the last valid hand using optical-flow correspondences with a constrained similarity transform to preserve hand shape, and apply a lightweight temporal filter to reduce high-frequency jitter. Frames produced by tracking are labeled as tracked/predicted. If tracking is unavailable or unstable and the dropout is very short, we use a conservative hold strategy that reuses the last exported hand configuration for a bounded number of frames, labeled as hold/predicted.

Under strong hand–hand overlap, detector outputs may flicker and left/right identity may become unstable. In this regime, we detect occlusion using high 2D overlap (intersection-over-union, IoU) together with depth-ordering cues when available, and emit an occlusion-consistent estimate that preserves continuity while explicitly exposing ambiguity; such frames are labeled as occluded/occluded. To prevent catastrophic artifacts from propagating through temporal

recovery, we apply plausibility constraints (e.g., scale consistency relative to the last reliable anchor, wrist-jump bounds, and bone-ratio stability); failing candidates are rejected and recorded for auditability. Crucially, these mechanisms do not overwrite detector-derived observed estimates and only fill explicitly labeled predicted/occluded intervals.

Anchor-Based Postprocessing

While the preceding stages already produce an explicit reliability-annotated stream, brief dropouts and frame-to-frame jitter may still reduce temporal stability. We therefore apply an optional anchor-based postprocessing step that fills short gaps between reliable observations and smooths non-anchor segments while preserving anchors exactly.

For each hand h , we define anchors as high-confidence, sanity-valid *observed* estimates (i.e., $q_t^h = \text{observed}$ from detector-derived sources $u_t^h \in \{\text{pass1}, \text{pass2}\}$). For any two consecutive anchors (t_0, t_1) , we fill only short gaps whose length satisfies $(t_1 - t_0 - 1) \leq G_{\max}$. To preserve hand shape, each anchor configuration is represented by its wrist position p_t^h , a scalar hand scale s_t^h , and normalized landmark offsets

$$O_t^h = \frac{H_t^h - p_t^h}{\max(s_t^h, \epsilon)}, \quad (3)$$

Intermediate frames $t \in (t_0, t_1)$ are synthesized by linear interpolation of $\{p_t^h, s_t^h, O_t^h\}$ between the two anchors, followed by reconstruction $H_t^h = p_t^h + s_t^h O_t^h$. All filled frames are explicitly labeled as postprocessed interpolation, $u_t^h = \text{interp}$ and $q_t^h = \text{predicted}$, in accordance with the source–state mapping defined in Table 1.

After gap filling, we optionally smooth trajectories to reduce high-frequency jitter. Smoothing is applied only to non-anchor frames and not across long missing intervals; anchors remain unchanged (or are restored after smoothing) to preserve reliable detector evidence. Figure 4 illustrates the effect on a representative wrist trajectory segment.

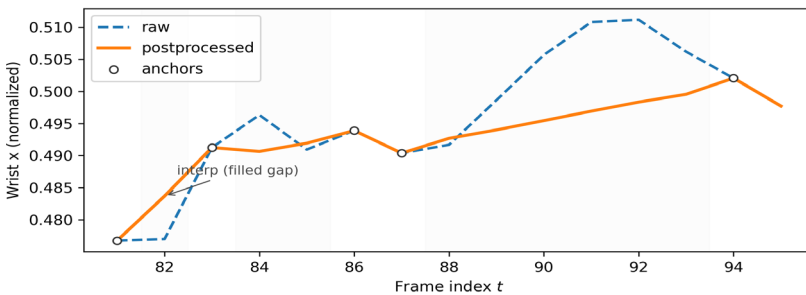


Figure 4. Anchor-based postprocessing example (zoomed segment): short gaps between anchors are filled by anchor-preserving interpolation, and optional smoothing reduces jitter on non-anchor frames while leaving anchors unchanged.

Quality Report: Definitions and Metrics. The extractor is designed to support large-scale sign-language dataset processing; however, the quality report is defined at the level of a single video and can be aggregated across a corpus when desired. Quality is characterized using observable proxies computed from the extracted stream, without requiring ground-truth hand landmarks. Let T' denote the number of processed frames. We use an indicator function $I(\cdot)$ that equals 1 when its argument is true and 0 otherwise. We index hands by $h \in \{L, R\}$. A hand is considered present at frame t when the extractor outputs a non-empty landmark set; we denote this by $H_t^h \neq \emptyset$. Let $c_t^h \in [0, 1]$ denote the per-frame confidence of the extracted hand estimate, and let P_t denote the pose estimate at frame t . Finally, let u_t^h denote the source label of the hand estimate (Section 4.1). We report coverage and confidence statistics, a fixed composite score for ranking, and auxiliary diagnostic rates that explain how coverage is achieved and how stable the stream is over time.

Coverage

Per-hand coverage measures the fraction of frames in which each hand is present:

$$Cov_L = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} I(H_t^L \neq \emptyset) \quad (4)$$

$$Cov_R = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} I(H_t^R \neq \emptyset) \quad (5)$$

For compact reporting, we also define overall hand coverage as the mean of the per-hand coverages:

$$Cov_{any} = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} I(H_t^L \neq \emptyset \vee H_t^R \neq \emptyset) \quad (6)$$

Both-hands coverage measures the fraction of frames where both hands are simultaneously available, which is particularly important for two-handed signs:

$$Cov_{LR} = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} I(H_t^L \neq \emptyset \wedge H_t^R \neq \emptyset) \quad (7)$$

Pose coverage is defined analogously as the fraction of frames for which a pose estimate is available in the output stream. When pose is computed sparsely, the nearest available pose estimate is reused and still counted as available:

$$Cov_P = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} I(P_t \text{ exists}) \quad (8)$$

Composite quality score.

For dataset-scale ranking and filtering under a fixed extraction configuration, we define a per-video quality score Q as a weighted combination of both-hands coverage Cov_{LR} (Eq. 7), mean per-hand coverage $(Cov_L + Cov_R)/2$ (Eq. 4–5), pose coverage Cov_P (Eq. 8). For each hand $h \in \{L, R\}$, we define the median confidence over frames where the hand is present as:

$$\text{MedConf}_h = \text{median}(\{c_t^h : H_t^h \neq \emptyset\}) \quad (9)$$

In the quality score, we use the mean of the two instances of MedConf_h for $h \in \{L, R\}$, i.e., $(\text{MedConf}_L + \text{MedConf}_R)/2$. The weights are fixed to make the score reproducible and comparable across runs:

$$Q = 0.4 \cdot Cov_{LR} + 0.2 \cdot \frac{(Cov_L + Cov_R)}{2} + 0.2 \cdot Cov_P + 0.2 \cdot \frac{(\text{MedConf}_L + \text{MedConf}_R)}{2} \quad (10)$$

This score is not intended as a universal measure of correctness; rather, it provides a reproducible proxy for how reliably a video yields usable pose–hand skeletons for skeleton-based SLR under a chosen extraction configuration.

Auxiliary diagnostic rates.

The composite score alone does not indicate whether coverage is obtained via direct detections or via recovery mechanisms, and it does not quantify identity corrections or temporal jitter. We therefore report additional diagnostic rates that are interpretable without ground truth and enable dataset-scale auditing and ablation studies.

Recovery fractions.

Using the source label u_t^h , we measure the fraction of frames where a hand estimate is produced by ROI-based second-pass refinement pass2, bridged by tracker-based temporal prediction tracked, or bridged by short-gap temporal propagation hold. The primary detector output pass1 is omitted as the default case:

$$\text{RecFrac}_{\text{pass2}}^h = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} \mathbb{I}(u_t^h = \text{pass2}) \quad (11)$$

$$\text{RecFrac}_{\text{tracked}}^h = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} \mathbb{I}(u_t^h = \text{tracked}) \quad (12)$$

These fractions distinguish videos dominated by observed frames from videos where coverage is achieved primarily through temporal inference or short-gap propagation.

Failure mode rates

To characterize dominant degradations, we report per-video failure rates computed from the exported stream. For two-hand failure modes we use a *hand-summed* rate (range $[0, 2]$), so the value can exceed 1 when both hands are affected in the same frame.

Missing rate. The fraction of hand-frames where landmarks are missing:

$$\text{MissingRate} = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} [I(H_t^L = \emptyset) + I(H_t^R = \emptyset)] \quad (14)$$

Occlusion rate. The fraction of hand-frames flagged as occluded by the pipeline:

$$\text{OcclusionRate} = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} [I(u_t^L = \text{occluded}) + I(u_t^R = \text{occluded})] \quad (15)$$

Swap rate. When left–right identity is ambiguous, the pipeline may apply corrective operations such as swapping left/right assignments. Let S_t be binary indicator of whether a swap is applied at frame t . We report the corresponding intervention rates:

$$\text{SwapRate} = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} S_t \quad (16)$$

Geometric outlier rate. To summarize temporal instability, we flag frames where an implementation-aligned criterion detects abnormally large normalized inter-frame changes in hand configuration and/or apparent scale. Let O_t^L and O_t^R be binary indicators of whether an outlier is flagged for the left and right hand at frame t . For failure-mode profiling, we report the hand-summed per-video outlier rate:

$$\text{OutlierRate} = \left(\frac{1}{T'}\right) \sum_{t=1}^{T'} [O_t^L + O_t^R] \quad (17)$$

Together, these diagnostics complement the composite score by capturing how often the pipeline must infer hands beyond primary detection, how frequently identity corrections are triggered, and whether the resulting trajectories are temporally stable properties that directly affect the usefulness of skeleton streams for SLR training.

Results. This section reports dataset-scale reliability of the proposed fault-tolerant keypoint extraction pipeline on Slovo ($n = 20,400$ clips). The analysis

targets stream robustness (availability and stability of exported pose-hand keypoints) rather than downstream recognition accuracy. Using the metrics defined in Section 5, we summarize: **(i)** hand-keypoint availability over time via any-hand coverage Cov_{any} (Eq. 6), per-hand coverages Cov_L, Cov_R (Eqs. 4–5), and both-hands coverage Cov_{LR} (Eq. 7); **(ii)** the distribution of the composite quality score Q used for ranking and filtering (Eq. 10); **(iii)** failure-mode rates - *MissingRate* (Eq. 14), *OcclusionRate* (Eq. 15), *SwapRate* (Eq. 16), and *OutlierRate* (Eq. 17) - reported with tail-sensitive statistics; and **(iv)** the contribution of recovery mechanisms via per-hand recovery fractions (Eqs. 11–13).

Overall, the goal is to provide a compact, reproducible quality report for monitoring, dataset filtering, and prioritizing engineering improvements, supported by an ablation comparison against a single-pass baseline.

Dataset coverage

We first summarize how consistently the pipeline exports hand keypoints over time. We report any-hand coverage Cov_{any} (Eq. 6), defined as the fraction of frames in which at least one hand is present, alongside per-hand coverages Cov_L, Cov_R (Eq. 4–5) and both-hands coverage Cov_{LR} (Eq. 7).

At the dataset scale, Cov_{any} is high for most clips: the 10th percentile of Cov_{any} is 0.923, and 70.6% of videos achieve perfect any-hand coverage ($Cov_{any} = 1.0$). Low-coverage cases form a relatively small tail: 7.5% of clips have $Cov_{any} < 0.90$, and severe failures are rare (1.2% with $Cov_{any} < 0.50$). Overall, this indicates that the extraction pipeline is stable on the majority of samples, while a small subset of clips remains challenging (e.g., hands leaving the frame, motion blur, or strong occlusions).

Next, we assess suitability for two-handed modeling using both-hands coverage Cov_{LR} (Eq. 7), the fraction of frames where both hands are simultaneously available. The distribution is strongly bimodal: 20.1% of clips have $Cov_{LR} = 0$ (both hands are never present at the same time), while 20.4% have $Cov_{LR} = 1$ (both hands are always present). For the remaining clips, simultaneous two-hand visibility is generally high (median $P_{50} = 0.854$, $P_{75} = 0.978$). However, the existence of a substantial $Cov_{LR} = 0$ group suggests that a non-trivial portion of the dataset is structurally one-handed due to framing or signer behavior. Therefore, downstream experiments that assume two visible hands should either filter by Cov_{LR} or rely on reliability-aware handling of missing parts.

Composite quality score distribution

We next summarize the distribution of the composite quality score Q defined in Eq. (10). Figure 5 reports the empirical CDF of Q over the whole dataset. The distribution is strongly skewed toward high scores: the 25th percentile is 0.598, the median is 0.918, and the 75th percentile is 0.984. This indicates that the majority of clips yield consistently usable pose-hand streams under the proposed extraction configuration, with a large mass concentrated near $Q \approx 1$ (see the right-tail zoom).

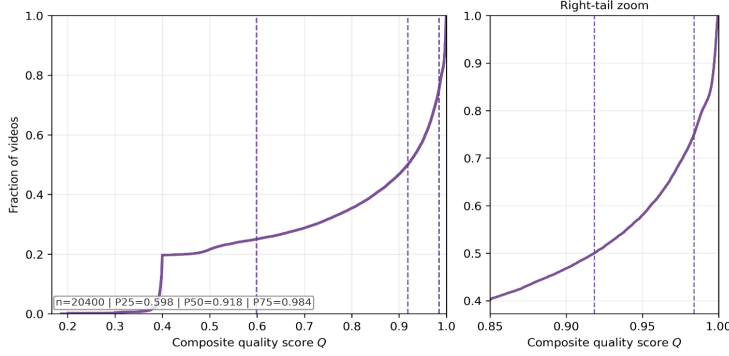


Figure 5. Empirical CDF (ECDF) of the composite quality score Q over Slovo ($n = 20,400$). Dashed lines indicate P_{25} , P_{50} , and P_{75} ; the right panel shows a zoom into the high-score tail.

At the same time, the lower quartile forms a clearly separated low-quality region (below ≈ 0.6), where clips are more likely to suffer from reduced coverage, low confidence, or reliance on recovery mechanisms. Practically, this motivates quality-aware filtering and monitoring: Q can be used (i) as a dataset-level health indicator across runs, and (ii) as a per-clip gating signal to exclude or down-weight low-quality samples in downstream training without requiring ground-truth landmarks.

Failure mode profile

We analyze dominant degradation patterns in the exported pose-hand streams using the per-video failure rates defined in Section 5: MissingRate (Eq. 14), OcclusionRate (Eq. 15), SwapRate (Eq. 16), and OutlierRate (Eq. 17). To emphasize rare but severe failures, Figure 6 reports complementary empirical CDFs (CCDFs), $S(x) = P(\text{rate} > x) = 1 - F(x)$, with a logarithmic y-axis. MissingRate, OcclusionRate, and OutlierRate are computed in hand-summed form (range $[0, 2]$), so values above 1 are possible when both hands are affected in the same frame. SwapRate is bounded to $[0, 1]$.

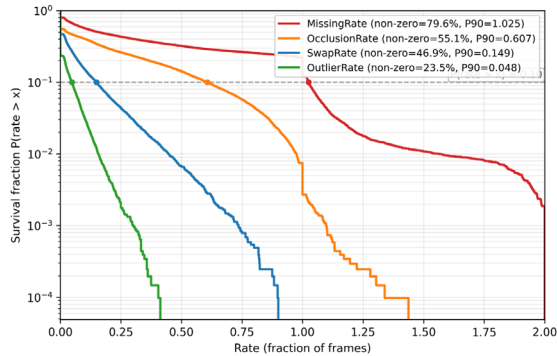


Figure 6. CCDFs (log y-axis) of per-video failure rates on Slovo ($n=20,400$): MissingRate, OcclusionRate, SwapRate, and OutlierRate. Missing/Occlusion/Outlier are hand-summed ($[0, 2]$); Swap is in $[0, 1]$. Legend shows non-zero clip fraction and P_{90} .

The **results** indicate that missing detections are the primary failure source at dataset scale. MissingRate is non-zero in 79.6% of clips and exhibits the heaviest tail among all modes ($P_{90}=1.025$), meaning that in the worst decile the stream contains sustained missing landmarks at roughly one affected hand-frame per frame on average (or higher when both hands are missing jointly). Occlusions represent the second most impactful failure mode, being non-zero in 55.1% of clips with a substantial tail ($P_{90}=0.607$), but consistently less severe than missing detections in the high-rate region.

In contrast, identity correction and geometric instability are secondary contributors. SwapRate is non-zero in 46.9% of clips, yet its magnitude remains relatively limited ($P_{90}=0.149$), suggesting that identity ambiguities occur but rarely dominate the temporal extent of a clip. OutlierRate is the least prevalent mode (23.5% non-zero) and has the lightest tail ($P_{90}=0.048$), consistent with sporadic instability events rather than systematic failure.

Overall, the failure-mode CCDFs provide a clear prioritization for further robustness improvements: the largest dataset-level gains are expected from reducing missing detections and improving occlusion handling, while swap and outlier mitigation primarily target narrower edge cases.

Recovery contribution

We quantify how much of each exported hand stream is produced by recovery mechanisms rather than the primary detector, using the per-clip recovery fractions defined in Eqs. (11), (12). Figure 7 compares the contribution of ROI-based second-pass recovery and tracker-based temporal bridging for left and right hands.

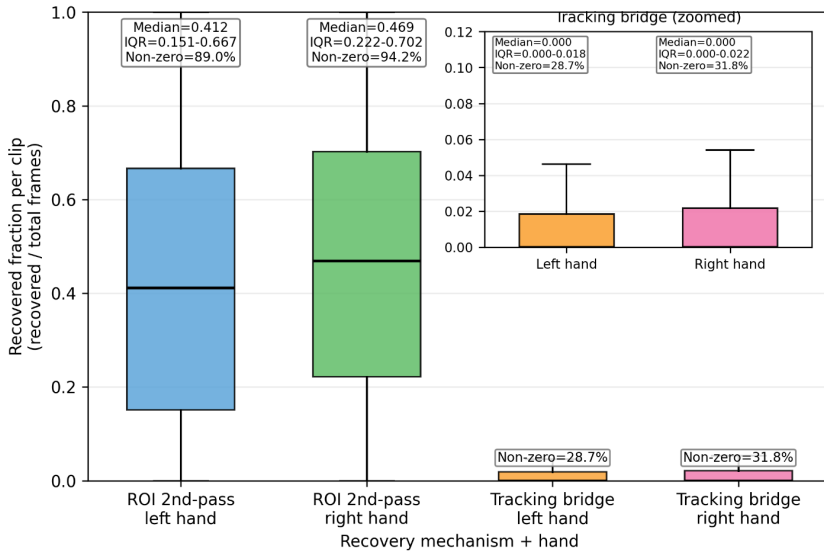


Figure 7. Per-clip recovered fraction (recovered / total clip frames) for ROI second-pass recovery and tracking bridge, separately for left and right hand, on Slovo (n = 20,400). Boxes show median and interquartile range (IQR); inset zooms the tracking regime.

Second-pass recovery is the dominant source at dataset scale: the median recovered fraction is 0.412 for the left hand (IQR 0.151–0.667; non-zero in 89.0% of clips) and 0.469 for the right hand (IQR 0.222–0.702; non-zero in 94.2% of clips). In contrast, tracking contributes little in most clips: the median is 0.000 for both hands, with small upper tails (IQR 0.000–0.018 for left; 0.000–0.022 for right) and non-zero rates of 28.7% and 31.8%, respectively. Overall, robustness is primarily driven by ROI-based second-pass recovery, while tracking acts as a sparse fallback mechanism for a minority of clips.

Ablation study

To quantify the impact of the fault-tolerant components, we compare a single-pass baseline (pass1-only) against the full pipeline on Slovo ($n = 20,400$ clips). We report a compact set of dataset-level reliability proxies defined in Section 5: the composite quality score Q (Eq. 10), any-hand coverage Cov_{any} (Eq. 6), both-hands coverage Cov_{LR} (Eq. 7), and the missing rate diagnostic MissingRate (Eq. 14).

Table 2 — Ablation results. pass-1 baseline vs. full fault-tolerant pipeline.

Metric	Pass-1	Full	Δ
Composite quality score Q	0.684	0.794	+0.110
Cov_{any}	0.889	0.969	+0.080
Cov_{LR}	0.454	0.646	+0.192
MissingRate	0.329	0.193	−0.136

As shown in Table 2, the full pipeline improves reliability across all reported criteria. The composite score increases by $+0.110$, consistent with a systematic shift toward more usable pose–hand streams. Coverage gains are most pronounced for simultaneous two-hand availability: Cov_{LR} rises by $+0.192$, indicating substantially more frames with both hands present - critical for two-handed signing segments. Any-hand coverage also improves ($+0.080$), while the missing-rate decreases by -0.136 , confirming fewer missing hand-frames overall. Collectively, these results support using the full fault-tolerant configuration as the default for large-scale keypoint export and downstream quality-aware filtering.

Discussion. This work argues that pose–hand keypoint extraction for sign-language corpora should be treated as an auditable, quality-controlled process rather than a silent preprocessing step. The proposed quality report operationalizes this view by providing dataset-scale reliability indicators that remain computable without ground-truth landmarks. The Results section shows that the exported streams are robust for most clips, while a clearly identifiable tail concentrates the cases that require filtering, remediation, or pipeline improvements.

A first practical observation is that availability is high but not uniform, and two-hand availability is structurally dataset-dependent. Any-hand coverage Cov_{any} is near-perfect for the majority of clips, whereas both-hands coverage Cov_{LR} is

strongly bimodal, with a substantial fraction of clips exhibiting $Cov_{LR} = 0$. This is not necessarily an extractor failure; it often reflects framing and signer behavior and implies that downstream pipelines assuming two visible hands should explicitly gate by Cov_{LR} or adopt reliability-aware handling of missing limbs. In this sense, the coverage metrics act as a data characterization tool in addition to being a robustness indicator.

Second, the distribution of the composite score Q is strongly skewed toward high values, but the lower quartile forms a meaningful low-quality region. This separation supports the use of Q as a compact quality-aware control signal: (i) for monitoring extraction health across runs and configurations, and (ii) for dataset filtering or sample reweighting in downstream training. Importantly, Q is designed to be reproducible under fixed weights (Eq. 10) and interpretable through its components (coverage, pose availability, and confidence), which reduces the risk of treating it as an opaque scalar.

Third, the failure-mode CCDFs provide a clear prioritization for further robustness improvements. Missing detections and occlusions dominate both in prevalence and in tail severity, while identity correction activity and geometric outliers are secondary contributors at the dataset scale. This indicates that the expected return on engineering effort is highest for mechanisms that directly reduce missing intervals and improve occlusion handling (e.g., better hand presence modeling, occlusion-aware refinement, or data-driven thresholds), whereas swap/outlier mitigation targets narrower edge cases.

Fourth, the recovery analysis shows that robustness gains are primarily driven by ROI-based second-pass refinement, with tracking acting as a sparse fallback. Together with the ablation results, this supports a concrete design implication: if compute or implementation complexity must be budgeted, prioritizing a strong second-pass refinement yields most of the dataset-scale benefit, while tracking provides smaller incremental gains and primarily addresses a minority of clips.

The reported metrics are proxy measures of reliability rather than geometric correctness, as no ground-truth landmarks are available. Recovery mechanisms can produce plausible trajectories that may still be inaccurate, so high coverage alone should not be interpreted as guaranteed correctness. Additionally, the diagnostic rates depend on implementation-aligned criteria (e.g., occlusion flags, outlier tests, identity interventions), and their absolute values may shift under domain changes such as viewpoint, resolution, cropping, and motion blur. Finally, while fixed weights make Q comparable across runs, the chosen weighting is still a modeling choice; alternative weightings could be better aligned with particular downstream objectives.

A natural next step is to validate how these reliability indicators correlate with downstream SLR performance (e.g., WER/accuracy vs. Cov_{LR} , $MissingRate$, and Q). Partial annotation of a small subset would enable calibration of proxy metrics against geometric error and help quantify when recovered frames become harmful. Another direction is adaptive configuration across domains (learned

thresholds, domain-conditioned weights for Q , or uncertainty-aware scoring) to improve transferability without sacrificing interpretability. Finally, targeted improvements for missing/occlusion handling are likely to deliver the largest dataset-scale gains, as suggested by the tail behavior in the failure-mode CCDFs.

Conclusion. We presented a quality-aware pose–hand keypoint extraction pipeline for skeleton-based SLR that integrates pose-guided left/right assignment, ROI-based second-pass recovery, short-gap temporal bridging (tracking/hold), sanity validation, and anchor-preserving postprocessing. Crucially, the pipeline exports explicit reliability semantics together with a proxy-metric quality report, turning keypoint extraction into an auditable and reproducible stage that supports dataset-level monitoring, filtering, and reliability-aware downstream training.

Dataset-scale evaluation on Slovo ($n = 20,400$) shows that the exported streams are usable for the majority of clips: any-hand coverage is near-saturated for most videos, and the composite quality score distribution is strongly concentrated toward high values ($P_{25} = 0.598$, $P_{50} = 0.918$, $P_{75} = 0.984$). Failure-mode profiling indicates that missing detections and occlusions dominate both prevalence and tail severity, making them the primary targets for future robustness gains, while identity correction activity and geometric outliers are comparatively secondary. Recovery attribution further demonstrates that robustness is driven mainly by ROI second-pass refinement (median recovered fraction 0.412/0.469 with non-zero rates 89.0–94.2%), whereas tracking contributes marginally (median ≈ 0 with non-zero rates $\approx 30\%$), acting primarily as a fallback. Finally, ablation confirms that the full fault-tolerant configuration improves dataset-level reliability across coverage and quality indicators relative to a single-pass baseline, supporting its use as the default setting for large-scale keypoint export and subsequent quality-aware filtering.

References

- Bazarevsky V., Grishchenko I., Raveendran K., Zhu T., Zhang F., & Grundmann M. (2020) *BlazePose: On-device real-time body pose tracking*. arXiv:2006.10204. doi:10.48550/arXiv.2006.10204 (in Eng.)
- Cao Z., Simon T., Wei S.-E., & Sheikh Y. (2017) *Realtime multi-person 2D pose estimation using part affinity fields*. In CVPR. — P. 7291–7299. doi:10.1109/CVPR.2017.143 (in Eng.)
- Duarte A., Palaskar S., Ventura L., Ghadiyaram D., DeHaan K., Metze F., Torres J., & Giro-i-Nieto X. (2021) *How2Sign: A large-scale multimodal dataset for continuous American Sign Language*. In CVPR. — P. 2735–2744. doi:10.1109/CVPR46437.2021.00276 (in Eng.)
- Grishchenko I., Bazarevsky V., Zanfır A., Bazavan E.G., Zanfır M., Yee R., Raveendran K., Zhdanovich M., Grundmann M., & Sminchisescu C. (2022) *BlazePose GHUM Holistic: Real-time 3D human landmarks and pose estimation*. arXiv:2206.11678. doi:10.48550/arXiv.2206.11678 (in Eng.)
- Jiang S., Sun B., Wang L., Bai Y., Li K., & Fu Y. (2021) *Skeleton aware multi-modal sign language recognition*. In CVPR Workshops. — P. 3413–3423. doi:10.1109/CVPRW53098.2021.00380 (in Eng.)
- Kapitanov A., Kvanchiani K., Nagaev A., & Petrova E. (2023) *Slovo: Russian Sign Language dataset*. In Computer Vision Systems (LNCS). — P. 63–73. Springer. doi:10.1007/978-3-031-44137-0_6 (in Eng.)

- Kindiroglu A.A., Ozdemir O., & Akarun L. (2019) *Temporal accumulative features for sign language recognition*. In ICCV Workshops. — P. 1288–1297. doi:10.1109/ICCVW.2019.00164 (in Eng.)
- Koller O., Forster J., & Ney H. (2015) *Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers*. Computer Vision and Image Understanding, 141. — P. 108–125. doi:10.1016/j.cviu.2015.09.013 (in Eng.)
- Koller O., Zargaran S., Ney H., & Bowden R. (2018) *Deep Sign: Enabling robust statistical continuous sign language recognition via hybrid CNN-HMMs*. International Journal of Computer Vision, 126. — P. 1311–1325. doi:10.1007/s11263-018-1121-3 (in Eng.)
- Li D., Rodriguez C., Yu X., & Li H. (2020) *Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison*. In WACV. — P. 1459–1469. doi:10.1109/WACV45572.2020.9093512 (in Eng.)
- Liu J., Shahroudy A., Xu D., & Wang G. (2016) *Spatio-temporal LSTM with trust gates for 3D human action recognition*. In B. Leibe et al. (Eds.), ECCV 2016 (LNCS 9907. — P. 816–833. Springer. doi:10.1007/978-3-319-46487-9_50 (in Eng.)
- Lugaresi C., Tang J., Nash H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.-L., Yong M.G., Lee J., Chang W.-T., Hua W., Georg M., & Grundmann M. (2019) *MediaPipe: A framework for building perception pipelines*. arXiv:1906.08172. doi:10.48550/arXiv.1906.08172 (in Eng.)
- Moryossef A., Tsochantaridis I., Aharoni R., Ebling S., & Narayanan S. (2020) *Real-time sign language detection using human pose estimation*. In Computer Vision – ECCV 2020 Workshops. Springer. doi:10.1007/978-3-030-66096-3_17 (in Eng.)
- Moryossef A. (2024) *Optimizing hand region detection in MediaPipe Holistic full-body pose estimation to improve accuracy and avoid downstream errors*. arXiv:2405.03545. doi:10.48550/arXiv.2405.03545 (in Eng.)
- Paoletti G., Beyan C., & Del Bue A. (2023) *SKELTER: Unsupervised skeleton action denoising and recognition using transformers*. Frontiers in Computer Science, 5, 1203901. doi:10.3389/fcomp.2023.1203901 (in Eng.)
- Song Y.-F., Zhang Z., Shan C., & Wang L. (2021) *Richly activated graph convolutional network for robust skeleton-based action recognition*. IEEE TCSVT, 31(5). — P. 1915–1925. doi:10.1109/TCSVT.2020.3015051 (in Eng.)
- Tunga A., Nuthalapati S.V., & Wachs J. (2021) *Pose-based sign language recognition using GCN and BERT*. In WACV Workshops. doi:10.1109/WACVW52041.2021.00008 (in Eng.)
- Xing Y., & Zhu J. (2021) *Deep learning-based action recognition with 3D skeleton: A survey*. CAAI Transactions on Intelligence Technology, 6(1). — P. 80–92. doi:10.1049/cit2.12014 (in Eng.)
- Yan S., Xiong Y., & Lin D. (2018) *Spatial temporal graph convolutional networks for skeleton-based action recognition*. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). — P. 7444–7452. doi:10.1609/aaai.v32i1.12328 (in Eng.)
- Zhang F., Bazarevsky V., Vakunov A., Tkachenka A., Sung G., Chang C.-L., & Grundmann M. (2020) *MediaPipe Hands: On-device real-time hand tracking*. arXiv:2006.10214. doi:10.48550/arXiv.2006.10214 (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 67–86

<https://doi.org/10.32014/2026.2518-1726.427>

IRSTI 81.93.29

UDC 004.056

© **Algazy K.¹, Alimzhan Y.^{1,2,*}, Sakan K.^{1,2}, Nyssanbayeva S.¹**, 2026.

¹Institute of information and computational technologies, Almaty, Kazakhstan;

²Al-Farabi Kazakh National University, Almaty, Kazakhstan.

E-mail: ayerkebulan19@gmail.com

LATTICE-BASED VECTOR COMMITMENTS FOR VERKLE TREES

Algazy Kunbolat — PhD, associate professor, Institute of information and computational technologies, Almaty, Kazakhstan,

E-mail: kunbolat@mail.ru, ORCID ID: <https://orcid.org/0000-0003-3670-2170>;

Alimzhan Yerkebulan — PhD student at Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: ayerkebulan19@gmail.com, ORCID ID: <https://orcid.org/0009-0008-7600-9139>;

Sakan Kairat — PhD, senior researcher, Institute of information and computational technologies, Almaty, Kazakhstan,

E-mail: kairat_sks@mail.ru, ORCID ID: <https://orcid.org/0000-0002-6812-6000>;

Nyssanbayeva Saule — doctor of technical sciences, professor, Institute of information and computational technologies, Almaty, Kazakhstan,

E-mail: sultasha1@mail.ru, ORCID ID: <https://orcid.org/0000-0002-5835-4958>.

Abstract. This paper presents a new lattice-based vector commitment scheme specifically optimized for Verkle trees and grounded in the hardness of the SIS problem. The main motivation is the development of compact and post-quantum secure cryptographic primitives suitable for scalable authenticated data structures used in modern blockchain systems and distributed ledgers. The proposed construction preserves the small proof and commitment sizes typical of classical vector commitments while ensuring resistance against quantum attacks and eliminating the need for a trusted setup. The paper provides a detailed description of the Setup, Commit, Open, and Verify algorithms, together with a formalized security model and clearly defined scheme parameters. Special attention is given to efficient encoding of polynomial coefficients into lattice space, principled parameter selection, and techniques for generating short witnesses with controlled norms, enabling a practical balance between security and computational efficiency. The construction is shown to support efficient multi-openings and batch verification, which are essential for high-arity Verkle tree paths and large-scale datasets. A comparative evaluation with classical commitment schemes such as KZG, IPA, and RSA is also presented, focusing on proof size, computational cost, and trusted setup

requirements. The results highlight the advantages of the lattice-based approach for long-lived distributed systems where long-term cryptographic security is critical. The proposed scheme demonstrates practical applicability and opens the way for integrating post-quantum vector commitments into scalable blockchain infrastructures.

Keywords: cryptography, vector commitments, Verkle trees, lattice, post-quantum cryptography

Funding. *This work was supported the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. fund this research AP23488112 - Development and study of a quantum-resistant digital signature scheme based on a Verkle tree)*

For citations: Algazy K., Alimzhan Y., Sakan K., Nyssanbayeva S. Lattice-based vector commitments for Verkle trees. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 67-86. DOI: <https://doi.org/10.32014/2026.2518-1726.427>

© Алғазы К.Т.¹, Әлімжан Е.Ж.^{1,2,*}, Сақан Қ.С.^{1,2}, Нысанбаева С.Е.¹, 2026.

¹Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан;

²Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

E-mail: ayerkebulan19@gmail.com

VERKLE АҒАШТАРЫНА АРНАЛҒАН ТОРЛЫҚ ВЕКТОРЛЫҚ МІНДЕТТЕМЕЛЕР

Алғазы Күнболат — PhD, қауымдастырылған профессор, Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан,

E-mail: kunbolat@mail.ru, ORCID ID: <https://orcid.org/0000-0003-3670-2170>;

Әлімжан Еркебұлан — әл-Фараби атындағы Қазақ ұлттық университетінің PhD докторанты, Алматы, Қазақстан,

E-mail: ayerkebulan19@gmail.com, ORCID ID: <https://orcid.org/0009-0008-7600-9139>;

Сақан Қайрат — PhD, аға ғылыми қызметкер, Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан,

E-mail: kairat_sks@mail.ru, ORCID ID: <https://orcid.org/0000-0002-6812-6000>;

Нысанбаева Сауле — техника ғылымдарының докторы, профессор, Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан,

E-mail: sultasha1@mail.ru, ORCID ID: <https://orcid.org/0000-0002-5835-4958>.

Аннотация. Мақалада Verkle-ағаштарында қолдануға арнайы оңтайландырылған және SIS есептерінің күрделілігіне негізделген жаңа торлық векторлық міндеттеме схемасы ұсынылады. Зерттеудің негізгі мақсаты болып заманауи блокчейн жүйелері мен үлестірілген реестрлерде қолданылатын масштабталатын аутентификацияланған деректер құрылымдары үшін ықшам әрі посткванттық тұрақты криптографиялық примитивтерді құру болып табылады. Ұсынылған құрылым классикалық векторлық міндеттемелерге

тән дәлелдер мен міндеттемелердің шағын өлшемін сақтай отырып, кванттық шабуылдарға төзімділікті қамтамасыз етеді және trusted setup қажеттілігін жояды. Жұмыста Setup, Commit, Open және Verify алгоритмдері егжей-тегжейлі сипатталып, қауіпсіздік моделі формализацияланған және схема параметрлері анықталған. Сонымен қатар көпмүшелер коэффициенттерін торлық кеңістікке тиімді кодтау әдістері, модульдік параметрлерді таңдау принциптері және нормасы бақыланатын қысқа куәліктерді генерациялау тәсілдері қарастырылған, бұл қауіпсіздік пен есептеу тиімділігі арасындағы тепе-теңдікті қамтамасыз етеді. Ұсынылған құрылым тиімді көпашуларды және пакеттік тексеруді қолдайтыны көрсетілген, бұл жоғары тармақталуы бар Verkle-ағаштары үшін аса маңызды. Сонымен қатар KZG, IPA және RSA міндеттемелерімен салыстырмалы талдау жүргізіліп, дәлелдер өлшемі, есептеу шығындары және trusted setup талаптары бойынша бағаланған. Нәтижелер ұзақ мерзімді үлестірілген жүйелер үшін торлық тәсілдің артықшылықтарын көрсетеді және посткванттық векторлық міндеттемелерді масштабалатын блокчейн инфрақұрылымдарына енгізу мүмкіндігін ашады.

Түйін сөздер: криптография, векторлық міндеттемелер, Verkle-ағашы, тор, посткванттық криптография

© Алгазы К.Т.¹, Алимжан Е.Ж.^{1,2,*}, Сакан К.С.^{1,2},
Нысанбаева С.Е.¹, 2026.

¹Институт информационных и вычислительных технологий,
Алматы, Қазақстан;

²Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан.

E-mail: ayerkebulan19@gmail.com

РЕШЕТОЧНЫЕ ВЕКТОРНЫЕ ОБЯЗАТЕЛЬСТВА ДЛЯ VERKLE- ДЕРЕВЬЕВ

Алгазы Кунболат — PhD, ассоциированный профессор, Институт информационных и вычислительных технологий, Алматы, Казахстан,

E-mail: kunbolat@mail.ru, ORCID ID: <https://orcid.org/0000-0003-3670-2170>;

Алимжан Еркебулан — PhD докторант Казахского национального университета имени аль-Фараби, Алматы, Казахстан,

E-mail: ayerkebulan19@gmail.com, ORCID ID: <https://orcid.org/0009-0008-7600-9139>;

Сакан Кайрат — PhD, старший научный сотрудник, Институт информационных и вычислительных технологий, Алматы, Казахстан,

E-mail: kairat_sks@mail.ru, ORCID ID: <https://orcid.org/0000-0002-6812-6000>;

Нысанбаева Сауле — доктор технических наук, профессор, Институт информационных и вычислительных технологий, Алматы, Казахстан,

E-mail: sultasha1@mail.ru, ORCID ID: <https://orcid.org/0000-0002-5835-4958>.

Аннотация. *Актуальность.* В статье предлагается новая решеточная схема векторного обязательства, специально оптимизированная для использования в Verkle-деревьях и основанная на сложности задачи SIS.

Актуальность исследования обусловлена необходимостью построения компактных и постквантово-устойчивых криптографических примитивов для масштабируемых аутентифицированных структур данных, применяемых в современных блокчейн-системах и распределенных реестрах. *Цель.* Разработать решеточную схему векторного обязательства для Verkle-деревьев, обеспечивающую компактность доказательств, устойчивость к квантовым атакам и отсутствие необходимости доверенной инициализации. *Методы.* В работе описаны алгоритмы Setup, Commit, Open и Verify, формализована модель безопасности и определены основные параметры предложенной схемы. Особое внимание уделено эффективному кодированию коэффициентов многочленов в решеточное пространство, выбору модульных параметров и разработке методов генерации коротких свидетельств с контролируемой нормой. Такой подход позволяет достичь баланса между криптографической безопасностью и вычислительной эффективностью. *Результаты и выводы.* Предложенная конструкция сохраняет малый размер доказательств и обязательств, характерный для классических векторных обязательств, одновременно обеспечивая устойчивость к квантовым атакам и устраняя необходимость trusted setup. Показано, что схема поддерживает эффективные многооткрытия и пакетную верификацию, что особенно важно для Verkle-деревьев с высокой арностью и большими наборами данных. Проведено сравнительное исследование с классическими схемами обязательств KZG, IPA и RSA по критериям размера доказательств, вычислительных затрат и требований к trusted setup. Результаты демонстрируют преимущества решеточного подхода для долгоживущих распределенных систем, где критична долговременная криптографическая устойчивость. Предложенная схема обладает практической применимостью и открывает возможности для интеграции постквантовых векторных обязательств в масштабируемые блокчейн-инфраструктуры.

Ключевые слова: криптография, векторные обязательства, Verkle-деревья, решетки, постквантовая криптография, SIS, блокчейн, распределенные реестры

Кіріспе. Verkle ағаштары классикалық Меркл ағаштарының шектеулері мен кемшіліктерін шешуге бағытталған, аутентификацияланған деректер құрылымдарының заманауи криптографиялық конструкциясы болып табылады. Бұл құрылымның негізгі ерекшелігі ретінде ағаш түйіндерінде сақталатын мәндерді байланыстыру үшін векторлық және полиномиалдық міндеттемелерді қолданылатынын атауға болады, бұл дұрыстық дәлелдерін анағұрлым ықшам әрі тиімді түрде ұсынуға мүмкіндік береді. Мұндай тәсіл дәстүрлі Merkle-дәлелдерімен салыстырғанда тиесілілік дәлелдерінің көлемін едәуір қысқартып, оларды тексеру кезеңіндегі есептеу шығындарын азайтады. Нәтижесінде Verkle ағаштары жоғары масштабталғыштық пен үлкен көлемдегі деректер жүйелерінде қолдануға жарамдылықты көрсетеді. Бұл

қасиеттер оларды блокчейн-платформаларында, криптографиялық деректер құрылымдарында және цифрлық қолтаңба схемаларында, дәлелдердің ықшамдығы мен верификация тиімділігі аса маңызды болған жағдайларда, перспективалы құралға айналдырады (Kuznetsov et.al., 2025).

Векторлық міндеттемелер тұтас мәндер жиынын бекітуге және кейіннен барлық деректер құрылымын ашпай-ақ жекелеген элементтердің дұрыстығын дәлелдеуге мүмкіндік беретін іргелі криптографиялық примитив. Осы ерекшелігінің арқасында олар нөлдік жариялау дәлелдемелері хаттамаларында, үлестірілген тізілімдерде және жоғары тиімді деректерді тексеру жүйелерінде кеңінен қолданылады. Дәстүрлі түрде мұндай міндеттерде қолданылатын полиномиалдық міндеттемелер, көбіне Лагранж интерполяциясы сияқты интерполяциялық әдістерге немесе полиномның коэффициенттік көрінісіне негізделеді. Алайда мұндай әдістер әдетте міндеттемелер көлемінің едәуір ұлғаюымен және дәлелдерді генерациялау мен тексерудің жоғары есептеу күрделілігімен сипатталады, бұл олардың масштабталатын криптографиялық қолданбалардағы практикалық тиімділігін шектейді.

Verkle ағаштарында бинарлық хэш-міндеттемелердің орнына векторлық немесе полиномиалдық міндеттемелер қолданылады, бұл дәлелдемелердің өлшемін, тармақталу негізі бойынша логарифмге дейін қысқартуға мүмкіндік береді (Iavich et.al., 2023). Verkle ағаштарының негізгі элементі қысқа дұрыстық дәлелдерін қолданатын полиномиалдық міндеттеме схемасы (Polynomial Commitment, PC) саналады (Kuznetsov et.al., 2023). Практикада кең таралған полиномиалдық міндеттеме схемаларының көпшілігі эллиптикалық топтардағы дискретті логарифмдеу қиындығы (DLOG), CDH/DLP (Computational Diffie–Hellman, Discrete Logarithm Problem) болжамдарына немесе RSA-есептеріне негізделген. Ең танымал мысалдары ретінде Kate–Zaverucha–Goldberg (KZG) міндеттемелері және IPA (Infinitely-Parented Arrays) негізіндегі схемалар саналады (John et.al., 2019).

Хэш-функциялар кванттық (Гровердің квадраттық жылдамдату алгоритміне) тұрақты деп саналса да, дискретті логарифмдеуге, эллиптикалық қисықтарға және RSA (Rivest, Shamir and Adleman) негізделген криптографиялық схемалар кванттық модельде осал деп танылады (Algazy et.al., 2024). Бұл жағдай, векторлық міндеттемелерді, аутентификацияланған деректер құрылымдарын және ұзақ мерзімді блокчейн жүйелерінің күйлерін қоса алғанда, маңызды криптографиялық примитивтерді посткванттық интеграциялау қажеттігін көрсетеді.

Кванттық есептеулер саласындағы прогреске байланысты АҚШ-тың Ұлттық стандарттар және технологиялар институты (ҰСТИ) посткванттық криптографиялық алгоритмдерді стандарттау процесін бастады. 2022–2024 жылдары цифрлық қолтаңбалар мен кілт алмасу механизмдері үшін алғашқы стандарттар таңдалып, бекітілді.

Торлық міндеттемелер мен дәлелдер саласындағы белсенді зерттеулерге қарамастан, қолданыстағы конструкциялардың көпшілігі SNARK/IVC

сценарийлеріне бағытталған немесе дәлелдерінің сызықтық өлшемімен, үлкен тұрақтыларымен, тиімді жаңартылуды қолдамауымен сипатталады. Бұл оларды тармақталу коэффициенті жоғары Verkle ағаштарында қолдануға қолайсыз етеді (Lyubashevsky et.al., 2022).

Қысқа бүтін шешім (Short Integer Solution, SIS) және Қателермен оқыту (Learning With Errors, LWE) есептеріне негізделген торлық конструкциялар ҰСТИ тарапынан посткванттық криптографияның сенімді кандидаттарының бірі ретінде қарастырылады, сондықтан олар болашақ Verkle ағаштарының табиғи негізі болып саналады.

Торлық векторлық міндеттемелерді Verkle ағашының құрылымына интеграциялау (Wee et.al., 2023):

1. Посткванттық тұрақтылықты қамтамасыз етеді, яғни қауіпсіздік CDH/RSA емес, SIS/LWE есептеріне негізделеді.

2. KZG міндеттемелеріне қарағанда сенімді бастапқы инициализация қажеттілігін жояды.

3. Verkle ағаштарының векторлық табиғаты есебінен дәлелдердің ықшамдығын сақтайды.

4. Жаңа икемділік қасиеттерін береді, мысалы жаңартылуды қолдау.

Бұл жұмыста Verkle ағаштарында қолдануға бағытталған жаңа толық негіздегі векторлық міндеттеме схемасы ұсынылады. Зерттеудің негізгі мақсаттары мыналар:

- иерархиялық деректер құрылымдарына жарамды торлық векторлық міндеттеме схемасын әзірлеу;

- ұсынылған схеманы Verkle ағашының архитектурасына интеграциялау;

- конструкцияның дұрыстығын, қауіпсіздігін және посткванттық тұрақтылығын талдау;

- міндеттемелердің, дәлелдердің өлшемі және есептеу шығындары бойынша CDH және RSA негізіндегі қолданыстағы міндеттемелермен салыстыру.

Әдеби шолу. Соңғы жылдары полиномиалдық және векторлық міндеттемелер Verkle-ағаштарын қоса алғанда, масштабталатын аутентификацияланған деректер құрылымдарын құрудағы негізгі криптографиялық примитивке айналды. Қолданыстағы жұмыстардың басым бөлігі дискретті логарифмдеуге және эллиптикалық қисықтарға негізделген схемаларға бағытталған. Олар жоғары тиімділікті қамтамасыз еткенімен, посткванттық тұрақтылыққа ие емес және көбіне сенімді бастапқы инициализацияны талап етеді. Осы бөлімде Verkle-ағаштарына бейімделген торлық векторлық міндеттемелердің қолданыстағы тәсілдеріне сыни талдау жүргізіледі.

Көптеген қолданыстағы векторлық міндеттеме схемалары дискретті логарифмдеу қиындығы немесе билинейлік бейнелеулер сияқты криптографиялық болжамдарға сүйенеді, бұл есептеу ресурстары мен қауіпсіздік параметрлеріне қосымша талаптар қояды. Сонымен қатар,

бірқатар схемаларда дәлел өлшемі мен верификация құны ағаш түйіндерінің арлығы (арность) артқан сайын өседі.

Алғашқы зерттеулер стандартты криптографиялық болжамдарға, соның ішінде билинейлік топтардағы RSA және Diffi–Хеллманға негізделген векторлық міндеттемелер құруға бағытталды (Catalano et.al., 2013). Бұл жұмыстар ықшам міндеттемелер мен ашуларды алуға болатынын көрсетіп, деректерді аутентификациялау және жинақтағыштар үшін шешімдердің тиімділігін едәуір арттырды.

Кейінгі зерттеулер алгебралық құрылымдар мен билинейлік топтарға негізделген тиімді міндеттемелерді жасауға арналды. (Kate et.al., 2010) жұмысында тұрақты өлшемді дәлелдері бар полиномиалдық міндеттемелер ұсынылып, олар вектор элементтері бойынша полином интерполяциясы арқылы тиімді векторлық міндеттемелерді жүзеге асырады. Бұл конструкциялар дәлелдердің ықшамдығы мен тексерудің жоғары жылдамдығына байланысты тексерілетін есептеулер жүйелерінде және үлестірілген тізілімдерде кеңінен қолданылды.

Nutu et.al., (2025) мақаласында векторлық міндеттемелер мен олардың эволюциясына, соның ішінде торлық криптографиялық нұсқаларына кең шолу берілген. Шолу ауқымды болғанымен, полиномдық міндеттемелер схемасы (Polynomial Commitment Scheme, PCS) немесе Verkle тақырыптарына ғана бағытталмаған. Дегенмен, криптографиялық PCS эволюциясын және олардың қолданылуын оң тұрғыда талдап, міндеттемелердің жалпы контекстін түсінуге пайдалы.

Fenzi et.al., (2024) жұмысында торға негізделген полиномиалдық міндеттеме ұсынылады, ол дәлелдердің сығылған өлшемін және полином дәрежесіне байланысты тексеру уақытын қамтамасыз етеді, асимптотикалық және практикалық тиімділікке назар аударады. Бірегей көпмүшені конструктивті түрде алуға мүмкіндік беретін міндеттеме енгізіледі. Қауіпсіздік пен тиімділік нақты тұжырымдалғанымен, кейбір оңтайлы қасиеттерге қол жеткізу үшін кездейсоқ оракул немесе қосымша болжамдар қажет. Посткванттық тұрақтылығына қарамастан, дәлел өлшемдері әзірге KZG сияқты эллиптикалық схемалардан үлкен.

Gini et.al., (2024) жұмысында авторлар ашық параметрлі (transparent setup) торлық тиімді шешімді ұсынады. Схема Module-SIS стандартты есебімен байланысты және Фиат–Шамир түрлендіруі арқылы кездейсоқ оракул моделінде интерактивті емес етуге болады. Қауіпсіздік деңгейі жоғары болғанымен, іске асыру күрделі және параметрлерді оңтайландыру бойынша қосымша зерттеулер қажет. Бұл жұмыс посткванттық PCS схемаларының заманауи деңгейін көрсетеді.

Торлық векторлық міндеттемелер бойынша жұмыстарды екі класқа бөлуге болады: интерактивті емес конструкциялар және тор негізіндегі Bulletproofs протоколының бейімделген нұсқалары. Albercht et.al., (2024) жұмысында стандартты болжамдар негізінде полилогарифмдік дәлел өлшемі мен тексеру

уақыты бар extractable PCS жаңа класы ұсынылып, практикалық схемаларға жол ашатыны көрсетілген, алайда Verkle-ағаштары контекстінде қосымша бағалау қажет. Блокчейн саласында практикалық іске асырулар әзірге аз.

Nguyen et.al., (2024) жұмысында ұсынылған тәсілдің негізгі мақсаты, дәлел өлшемін азайту және практикалық жүйелердегі тексеру процедурасын жетілдіру. Ұсынылған схема шамамен 93 КБ көлеміндегі дәлелдерді генерациялайды және Verkle-ағаштарында қолдануға болатын PCS үшін идеялар ұсынады. Практикалық нәтижелері әсерлі болғанымен, қауіпсіздік тұрғысынан тереңірек талдау қажет.

Algazy et.al., (2024) жұмысында Verkle-ағаштарында CRT (Chinese Remainder Theorem) негізіндегі алгоритм қолданылып, үлестірілген жүйелер үшін шешімнің практикалық тиімділігі негізделеді. Жұмыс нақты қолданбалы мәселелерге бағытталғанымен, PCS қауіпсіздігінің іргелі аспектілерін терең қарастырмайды. (Iavich et.al., 2025) мақалаларында Verkle-құрылымдары мен торлық векторлық міндеттемелерді қолдана отырып, жадыны және қолтаңба өлшемін оңтайландыратын посткванттық цифрлық қолтаңба схемаларын құру зерттеледі, сондай-ақ кванттық шабуылдарға төзімділік пен тиімділікті арттыру үшін QRNG (Quantum Random Number Generator) интеграциясы ұсынылады.

Wang et.al., (2026) жұмысында блокчейндегі көптеген күйлерді (state) агрегатталғандәлелдерарқылы тексеру үшін Verkle-ағашы, Verkle-аккумулятор және KZG-полиномиалдық міндеттемелердің интеграциясы қарастырылады. Талдау нәтижесінде дәлелдер өлшемі мен тексеру жылдамдығы классикалық схемалардан жақсы екені көрсетілген.

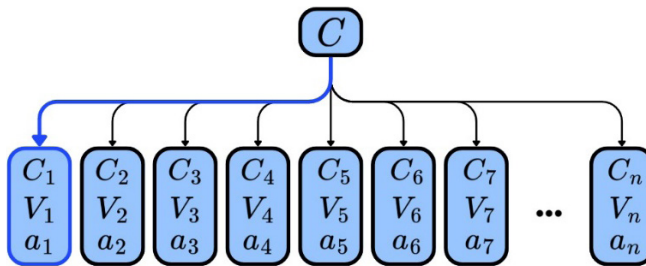
Материалдар мен негізгі әдістер. *Векторлық міндеттемелер және олардың Verkle-ағаштарындағы рөлі.* Векторлық міндеттемелер (vector commitments, VC) жекелеген элементтерді кейін дәлелмен ашу мүмкіндігімен мәндер жиынын ықшам түрде бекітуге арналған криптографиялық примитивтер. VC негізгі қасиеттеріне дұрыстық (correctness), байланыстылық (binding), дәлелдердің қысқалығы және кейбір схемаларда жасырындық (hiding) жатады. Векторлық міндеттемелердің басты мақсаты ретінде деректер массивтерін тиімді аутентификациялау жатады.

Классикалық хэш-ағаштармен салыстырғанда, VC жоғары тармақталу кезінде кіші өлшемді дәлелдер алуға мүмкіндік береді. Дәл осы қасиет оларды масштабталатын блокчейн жүйелері мен күйді (state) сақтау тиімділігін арттыру үшін ұсынылған Verkle-ағаштарының негізгі элементіне айналдырады. Элементтің тиесілілігін дәлелдеу кезінде жапырақтан түбірге дейінгі жол бойындағы векторлардың тек қажетті позициялары ғана ашылады. Осылайша, Merkle-ағашындағыдай көршілес хэштердің толық жиынын берудің орнына, жекелеген индекстерді ашудың ықшам дәлелдері жеткілікті болады.

Verkle-ағаштары контекстінде векторлық міндеттемелер классикалық Merkle-ағаштарымен салыстырғанда асимптотикалық тұрғыдан ықшам

дәлелдер алуды қамтамасыз етеді және бір тармағының сұлбасы 1-суретте кескінделген. Verkle-құрылымдарында VC қолданудың негізгі идеялары:

- ағаш жапырақтары міндеттемелердің мәндерінен тұрады;
- ішкі түйіндер – еншілес түйіндер векторына жасалған міндеттемелер;
- векторлар көпмүше ретінде интерпретацияланады;
- элементтің ағашқа тиесілілігі бірнеше нүктедегі мәндерін ашу арқылы дәлелденеді.



Сурет 1 – Verkle ағашының бір тармағының сұлбасы.

Дәстүрлі Merkle-ағашында дәлел өлшемі шамамен $O(\log n)$ және көптеген хэштерден тұрады. Ал Verkle-ағашында дәлел агрегатталып, қолданылатын міндеттеме схемасына байланысты тұрақтыға жақын немесе кіші коэффициентті логарифмдік өлшемде болуы мүмкін. Векторлық міндеттемелер мәндер жиынын бекітіп, кейін бүкіл құрылымды ашпай-ақ жекелеген компоненттердің дұрыстығын дәлелдеуге мүмкіндік береді. Бұл қасиет Verkle-ағаштарын үлестірілген тізілімдерде және ресурстары шектеулі хаттамаларда қолдану үшін ерекше тартымды етеді.

Торлар посткванттық криптографияның негізі ретінде. Торлық криптография посткванттық криптографияның ең перспективалы бағыттарының бірі болып саналады. Оның қауіпсіздігі келесі есептердің есептеу қиындығына негізделеді: SVP (Shortest Vector Problem), CVP (Closest Vector Problem), SIS (Short Integer Solution), LWE / Ring-LWE / Module-LWE. Алайда торлық міндеттемелерді, әсіресе Verkle-ағаштары контекстінде, векторлық формаға бейімдеу мәселесі әлі де жеткілікті зерттелмеген.

Анықтама 1 (Торлық векторлық міндеттеме).

Торлық векторлық міндеттеме – реттелген мәндер векторын бекітуге және оның жекелеген элементтерін кейін дәлелмен ашуға арналған $\Pi = (\text{Setup}, \text{Commit}, \text{Open}, \text{Verify})$ криптографиялық схемасы. Ол келесі бөліктерден тұрады:

- **Setup** ($1^\lambda, n$): Қауіпсіздік параметрі λ және вектордың максимал ұзындығы n бойынша $\Lambda \subset \mathbb{Z}^m$ торымен немесе сәйкес $\mathbb{Z}_q^m$ модульдік кеңістігімен байланысты жария параметрлер pp генерацияланады. Параметрлерге міндеттемелер құру үшін қолданылатын кездейсоқ матрицалар немесе тор базистері кіреді.

– **Commit**(pp, v): $v = (v_1, \dots, v_n), v_i \in \mathbb{Z}_q$ векторы үшін $C = \text{Com}(v)$ міндеттемесін есептейді. Мысалы, $C = Av + e \pmod{q}$, мұнда A – жария етілетін (public) матрица, e – жасырындықты қамтамасыз ететін кіші кездейсоқ вектор.

– **Open**(pp, v, i): i позициясындағы мәннің v_i екенін және оның C міндеттемесіне сәйкес келетінін дәлелдейтін π_i дәлелін құрады.

– **Verify**(pp, C, v_i, π_i): ашудың дұрыстығын тексеріп, *accept* немесе *reject* нәтижесін қайтарады.

Нәтижелер. Бірінші кезекте, қолданылған белгілеулер мен параметрлерге тоқталайық. q – модуль ретінде жай сан таңдалады. $\mathbb{Z}_q$ арқылы q модулі бойынша қалдықтар сақинасын белгілейік. Векторлар мен матрицалар $\mathbb{Z}_q$ -де немесе $\mathbb{Z}$ -те қарастырылып, кейін q модулі бойынша келтіріледі.

n, m, k (натурал) параметрлері беріледі, мұндағы k – коммиттелетін мәндер векторының ұзындығы. $A_i \in \mathbb{Z}_q^{n \times m} - i \in \{1, \dots, k\}$ позицияларына сәйкес келетін жария матрицалар, ал $U_i \in \mathbb{Z}_q^n$ – позициялардың жария вектор белгілері. Дәлелдің дұрыстығын бақылау үшін γ нормасына шек енгізіледі, яғни $\|V_i\| \leq \gamma$ шарты тексеріледі. Хеш-функция ретінде $H: \{0,1\}^* \rightarrow \mathbb{Z}_q$ немесе $H: \{0,1\}^* \rightarrow \{0,1\}^\lambda$ түріндегі криптографиялық тұрғыдан берік функционал пайдаланылады.

Initialization (n, k). Инициализация кезеңінде (q, m) параметрлері бекітіледі, γ нормасына шектеулер таңдалады және H хеш-функциясы анықталады. m_i хабарламалары әрі қарай скалярлар түрінде қарастырылады:

$$a_i := H(m_i) \in \mathbb{Z}_q, i = 1, \dots, k.$$

KeyGen(1^n). Әрбір $i \in \{1, \dots, k\}$ индексі үшін $A_i \in \mathbb{Z}_q^{n \times m}$ матрицасы таңдалады.

Сонымен қатар *trapdoor* деп аталатын T_i құпия ақпараты генерацияланады, ал $A_i x \equiv U \pmod{q}$ түріндегі салыстырулар, кез-келген $u \in \mathbb{Z}_q^n$ үшін қысқа шешімдерді тиімді табуға мүмкіндік береді.

Одан кейін жария позициялық векторлар таңдалады:

$$U_1, \dots, U_k \in \mathbb{Z}_q^n.$$

Барлық $i \neq j$ индекс жұптары үшін қосымша қысқа элементтер анықталады:

$$R_{i,j} := \text{SamplePre}(T_i A_i U_j),$$

мұндағы, *SamplePre* – қысқа түпбейнені таңдау процедурасы. Анықтама бойынша $A_i R_{i,j} \equiv U_j \pmod{q}$ және $\|R_{i,j}\|$ алдын-ала берілген параметрлермен шектеледі.

Схеманың жария параметрлері:

$$pp = (q, n, m, k, \{A_i\}_{i=1}^k, \{U_i\}_{i=1}^k = 1k, \gamma, H),$$

ал, коммиттердің құпия көмекші деректері $\{T_i\}$ немесе жүзеге асыруға байланысты $\{R_{i,j}\}_{i \neq j}$ жиыны ретінде берілуі мүмкін.

$\text{Com}_{pp}(m_1, \dots, m_k)$. Барлық i үшін $a_i := H(m_i) \in \mathbb{Z}_q$ есептеледі. $(a_1, \dots, a_k)$ векторына коммитмент келесідей анықталады:

$$C := \sum_{i=1}^k a_i U_i \pmod{q}.$$

Алынған $C \in \mathbb{Z}_q^n$ векторы барлық позициялар үшін бірыңғай міндеттеме.

$\text{Open}_{pp}(m_1, \dots, m_k)$. i -ші позицияны ашу қажет болған жағдайда, коммиттер барлық j үшін $a_j := H(m_j)$ есептейді. i позициясына дәлел алдында дайындалған кері бейнелердің (preimage) тіркесімі ретінде анықталады:

$$V_i := \sum_{j \neq i} a_j R_{i,j} \pmod{q}.$$

Ашу (орен) нәтижесі тексеруші a_i -ді өзі есептей ме соған байланысты (m_i, V_i) немесе (a_i, V_i) жұбы болып табылады.

$\text{Ver}_{pp}(C, i, m_i, V_i)$. Тексеруші $a_i := H(m_i)$ есептейді. Кейін екі шарт тексеріледі:

1. Дәлел нормасының шектеуі: $\|V_i\| \leq \gamma$.
2. Негізгі дұрыстық теңдігі: $C \equiv A_i V_i + a_i U_i \pmod{q}$.

Екі шарт та орындалса, ашу қабылданады және m_i хабарламасының C коммитментіне i -ші позицияға дұрыс тиесілі екендігі расталады.

Дұрыстықты негіздеу. Дұрыстық $R_{i,j}$ анықтамасынан туындайды. $j \neq i$ үшін шарт бойынша $A_i R_{i,j} \equiv U_j \pmod{q}$. Сонда,

$$A_i V_i + a_i U_i \equiv \sum_{j \neq i} a_j A_i R_{i,j} + a_i U_i \equiv \sum_{j \neq i} a_j U_j + a_i U_i = \sum_{j=1}^k a_j U_j \equiv C \pmod{q}.$$

Демек, дұрыс құрылған V_i дәлелі әрқашан тексеруден өтеді.

Посткванттық қауіпсіздікті 128 биттен төмен емес деңгейде қамтамасыз ету үшін келесі параметрлерді пайдалану ұсынылады:

- модуль: $q > 2^{23} \approx 8388608$;

- тор өлшемділігі: $n = 512$;
- бағандар саны: $m = 1024$.

Бұл параметрлер заманауи посткванттық торлық криптографиялық схемалардың ұсынымдарына сәйкес келеді.

Ескерту. Толыққанды криптографиялық жүзеге асыруда дәлелдемелер trapdoor-параметрлерін және қысқа түпбейнелерді іріктеу процедураларын (мысалы, SamplePre тобы) пайдалану арқылы құрылады, сондай-ақ $\|V_i\| \leq \gamma$ шарты бақыланады. Келесі көрсетілетін мысалда, арифметиканы қолмен толық тексеруге мүмкіндік беру үшін квадратты A матрицасы және алдын ала дайындалған түпбейнелер қолданылады.

Сандық мысалға өтпестен бұрын, gadget-матрицаның, trapdoor-параметрлерінің және SamplePre процедурасының рөлін түсіндіру қажет. Бұл элементтер схеманың толық криптографиялық жүзеге асырылуына жатады. Олар ашу дәлелдемесінің қысқа болуын және оның q модулі бойынша дұрыс тексерілуін қамтамасыз етеді.

Gadget-матрица G құру үшін алдымен модуль q бойынша екілік разрядтар саны анықталады:

$$l = \lceil \log_2 q \rceil$$

Бұл мән q модуліндегі кез-келген элементті екілік түрде көрсету үшін қажет болатын разрядтар санын береді. Осыдан кейін gadget-вектор енгізіледі:

$$g = (1, 2, 4, \dots, 2^{l-1}) \in \mathbb{Z}_q^l.$$

Бұл вектор екілік жіктеу салмақтарынан тұрады. Яғни кез келген $t \in \mathbb{Z}_q$ элементі осы салмақтар арқылы екілік түрде жазылады:

$$t = \sum_{s=0}^{l-1} b_s 2^s \pmod{q},$$

мұндағы $b_s \in \{0, 1\}$. Сондықтан g векторы кейін G^{-1} операциясында қолданылады. Бұл операция вектор координаталарын екілік компоненттерге жіктейді.

n -өлшемді векторлармен жұмыс істеу үшін gadget-вектор g жеке координатаға ғана емес, барлық n координатаға қолданылады. Осы мақсатта бірлік матрица I_n алынады және gadget-матрица $G = I_n \otimes g$ құрылады. Нәтижесінде $G \in \mathbb{Z}_q^{n \times nl}$ өлшемді матрица алынады. Бұл матрица блоктық құрылымға ие. Әрбір жолда gadget-вектор g тек бір координатаға сәйкес орналасады, ал қалған блоктар нөлдерден тұрады. Мысалы, $n = 3$ болған жағдайда G матрицасы мына түрде жазылады:

$$G = \begin{bmatrix} g & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & g \end{bmatrix}$$

Егер $g = (1, 2, 4, \dots, 2^{l-1})$ болса, онда бұл матрица әрбір координатаны жеке екілік разрядтар арқылы қалпына келтіруге мүмкіндік береді. Сондықтан G матрицасы trapdoor-құрылымда негізгі техникалық құрал ретінде қолданылады. Ол A_i матрицасының жасырын құрылымын gadget-ұсынылым арқылы байланыстырады:

$$A_i \begin{bmatrix} T_i \\ I_{nl} \end{bmatrix} = G \pmod{q}.$$

Осы байланыс кейін SamplePre процедурасында пайдаланылады. Атап айтқанда, $t_{i,j}$ векторы үшін алдымен оның екілік $b_{i,j} \leftarrow G^{-1}(t_{i,j})$ жіктелуі табылады. Содан кейін осы жіктеу арқылы қысқа түпбейне құрылады.

Әрбір $i \in \{1, \dots, k\}$ позиция үшін $A_i \in Z_q^{n \times (m+nl)}$ ашық матрица құралады. Бұл матрица блоктық түрде беріледі, яғни $A_i = [\bar{A}_i \mid A_{i,2}]$, мұндағы $\bar{A}_i \in Z_q^{n \times m}$ – кездейсоқ компонент. Матрицаның екінші бөлігі кездейсоқ таңдалмайды. Ол G gadget-матрицасы және құпия trapdoor-матрица T_i арқылы анықталады:

$$A_{i,2} = G - \bar{A}_i T_i \pmod{q},$$

мұнда, $T_i \in Z_q^{m \times nl}$ құпия параметр болып табылады. Оны «қақпан» немесе trapdoor ретінде түсіндіруге болады. Егер, T_i белгісіз болса, онда $A_i R \equiv U \pmod{q}$ шартын қанағаттандыратын қысқа R векторын табу қиын есепке айналады. Мұндай есептің күрделілігі торлар теориясындағы қысқа шешімдерді табу есептерімен байланысты. Сондықтан A_i ашық матрицасы барлық қатысушыларға берілуі мүмкін. Ал құпия T_i матрицасы тек көмекші дәлелдемелерді құратын жақта сақталады.

Жоғарыда айтылған мәліметтерден A_i матрицасының арнайы түрде құрылғанын түсінуге болады. T_i арқылы A_i матрицасы G gadget-матрицасымен байланысады. Осы байланыс қысқа түпбейнені құру процедурасында қолданылады.

SamplePre процедурасы U_j векторының A_i матрицасына қатысты қысқа түпбейнесін құру үшін қолданылады. Индекстердің $i \neq j$ жұбы үшін ол келесідей жазылады:

$$R_{i,j} = \text{SamplePre}(T_i, A_i, U_j)$$

Бұл процедураның мақсаты $A_i R_{i,j} = U_j \pmod{q}$ шартын қанағаттандыратын $R_{i,j}$ векторын табу. Сонымен қатар $R_{i,j}$ қысқа болуы керек. Бұл шарт дәлелдеме нормасын тексеру үшін қажет, яғни $\|V_i\| \leq \gamma$.

Gadget-based нұсқада алдымен, $y_{i,j} \in \mathbb{Z}^m$ қысқа вектор таңдалады. Содан кейін $t_{i,j} = U_j - \bar{A}_i y_{i,j} \pmod{q}$ қалдық вектор есептеледі. Бұл вектор U_j мәнінің қандай бөлігін gadget-блок арқылы алу қажет екенін көрсетеді. Кейін $b_{i,j} \leftarrow G^{-1}(t_{i,j})$ gadget decoding қолданылады. Мұнда, G^{-1} операциясы $t_{i,j}$ векторының әрбір координатасын екілік жіктеуге сәйкес келеді. Бұл жіктеу $(1, 2, 4, \dots, 2^{l-1})$ салмақтары бойынша орындалады.

Осыдан кейін қысқа түпбейне келесі формула бойынша анықталады:

$$R_{i,j} = \begin{bmatrix} y_{i,j} + T_i b_{i,j} \\ b_{i,j} \end{bmatrix} \pmod{q}.$$

Бұл формуланың дұрыстылығын тікелей қою арқылы тексеруге болады.

$$A_i R_{i,j} = [\bar{A}_i \mid A_{i,2}] \begin{bmatrix} y_{i,j} + T_i b_{i,j} \\ b_{i,j} \end{bmatrix} = \bar{A}_i y_{i,j} + \bar{A}_i T_i b_{i,j} + (G - \bar{A}_i T_i) b_{i,j} \pmod{q}$$

мұнда, $\bar{A}_i T_i b_{i,j}$ бар мүшелер қысқарады. Сондықтан, $A_i R_{i,j} = \bar{A}_i y_{i,j} \pmod{q}$ өрнегі қалады. Ал $b_{i,j} = G^{-1}(t_{i,j})$ болғандықтан, $G b_{i,j} = t_{i,j} \pmod{q}$ теңдігі орындалады. $t_{i,j}$ анықтамасын ескерсек, $G b_{i,j} = U_j - \bar{A}_i y_{i,j} \pmod{q}$. Сонда, $A_i R_{i,j} = \bar{A}_i y_{i,j} + U_j - \bar{A}_i y_{i,j} = U_j \pmod{q}$.

Осылайша, $R_{i,j}$ векторы U_j векторының A_i матрицасына қатысты түпбейнесі екені дәлелденеді.

Толық схемада мұндай $R_{i,j}$ мәндері барлық $i \neq j$ жұптары үшін алдын ала құрылады. Кейін олар коммитменттің бір позициясын ашу кезінде қолданылады. Егер i -позицияны ашу қажет болса, онда дәлелдеме алдын ала дайындалған түпбейнелердің сызықтық комбинациясы ретінде құрылады:

$$V_i = \sum_{j \neq i} a_j R_{i,j} + r \pmod{q}$$

Осыдан кейін тексеруші тарап $C = A_i V_i + a_i U_i \pmod{q}$ теңдігін қолданады. Trapdoor-кезеңі құпия құрылымы бар матрицаны қалыптастырады. Ал SamplePre процедурасы осы құрылымды пайдаланып, қысқа түпбейнелерді алады.

Келесі сандық мысалда бұл процедура жеңілдетілген түрде көрсетіледі. Қолмен тексеруді жеңілдету үшін толық gadget-құрылымның орнына квадратты A матрицасы және алдын ала дайындалған түпбейнелер қолданылады. Бұл жеңілдету схеманың негізгі идеясын өзгертпейді, бірақ есептеулерді айқын көрсетеді.

Мысал. Торлы векторлық коммитменттің жұмыс істеу қағидатын көрсету үшін тәжірибелік мысал келтіріледі. Мақсатымыз коммитмент пен ашу дәлелін есептеу тәртібін көрнекі түрде көрсету, сондай-ақ верификация алгоритмінің негізінде жатқан теңдіктің тексерілуін көрсету.

Есептеулер $\mathbb{Z}_{101}$ —де жүргізіледі (барлық операциялар $q = 101$ модулінде орындалады).

$$A = \begin{pmatrix} 41 & 19 & 50 & 83 & 6 & 9 & 68 \\ 12 & 46 & 74 & 7 & 64 & 27 & 4 \\ 11 & 55 & 53 & 8 & 30 & 11 & 70 \\ 54 & 7 & 72 & 15 & 28 & 80 & 80 \\ 74 & 7 & 73 & 74 & 50 & 6 & 28 \\ 5 & 71 & 17 & 37 & 53 & 18 & 69 \\ 15 & 73 & 39 & 71 & 87 & 23 & 13 \end{pmatrix} \pmod{101},$$

матрицасы және келесі ашық векторлары берілсін:

$$U_1 = (4, 7, 1, 18, 22, 5, 9), U_2 = (8, 2, 9, 30, 17, 11, 6), U_3 = (5, 6, 3, 14, 27, 19, 2),$$

Сонымен қатар, $a_i = H(m_i)$ мәндері де берілсін: $a_1 = 11, a_2 = 4, a_3 = 13 \in \mathbb{Z}_{101}$. (a_1, a_2, a_3) үштігіне коммитменттер төмендегі комбинация арқылы анықталады:

$$C \equiv a_1 U_1 + a_2 U_2 + a_3 U_3 \pmod{101}.$$

Координаталар бойынша көбейтінділерді есептеп, нәтижелерін 101 модулі бойынша келтіреміз:

$$\begin{aligned} a_1 U_1 &= 11 \cdot (4, 7, 1, 18, 22, 5, 9) = (44, 77, 11, 198, 242, 55, 99) \equiv (44, 77, 11, 97, 40, 55, 99), \\ a_2 U_2 &= 4 \cdot (8, 2, 9, 30, 17, 11, 6) = (32, 8, 36, 120, 68, 44, 24) \equiv (32, 8, 36, 19, 68, 44, 24), \\ a_3 U_3 &= 13 \cdot (5, 6, 3, 14, 27, 19, 2) = (65, 78, 39, 182, 351, 247, 26) \equiv (65, 78, 39, 81, 48, 45, 26). \end{aligned}$$

Есептелген мәндерді қосатын болсақ,

$$\begin{aligned} C &\equiv (44, 77, 11, 97, 40, 55, 99) + (32, 8, 36, 19, 68, 44, 24) + (65, 78, 39, 81, 48, 45, 26) \\ &= (141, 163, 86, 96, 55, 144, 48) \equiv (40, 62, 86, 96, 55, 43, 48). \end{aligned}$$

Нәтижесінде, $C \equiv (40, 62, 86, 96, 55, 43, 48) \pmod{101}$.

1-позицияны ашу үшін дәлелдемені қалыптастыру.

Мысалда, R_2, R_3 түпбейнелерінің бар екені және олардың келесі салыстыру шарттарын қанағаттандыратыны болжанады:

$$AR_2 \equiv U_2 \pmod{101}, AR_3 \equiv U_3 \pmod{101}, \text{ және оларды сандық тізбек}$$

ретінде береміз, $R_2 = (93,78,20,24,66,20,100)$, $R_3 = (74,87,100,9,51,63,15)$. 1-позицияның үшін дәлелдемені анықтап аламыз как $V_1 \equiv a_2 R_2 + a_3 R_3 \pmod{101}$ және келесідей есептеулер жүргіземіз:

$$\begin{aligned} a_2 R_2 &= 4 \cdot (93,78,20,24,66,20,100) = \\ &= (372,312,80,96,264,80,400) \equiv (69,9,80,96,62,80,97), \\ a_3 R_3 &= 13 \cdot (74,87,100,9,51,63,15) = (962,1131,1300,117,663,819,195) \\ &\equiv (53,20,88,16,57,11,94). \end{aligned}$$

Демек,

$$\begin{aligned} V_1 &\equiv (69,9,80,96,62,80,97) + (53,20,88,16,57,11,94) = \\ &= (122,29,168,112,119,91,191) \\ &\equiv (21,29,67,11,18,91,90). \end{aligned}$$

теңдігінің тексеру үшін есептеулер жүргіземіз.

Алдымен AV_1 есептеп аламыз. $V_1 = (21,29,67,11,18,91,90)$ болсын, онда A матрицасының жолдары бойынша келесі мәндерді есептеп аламыз:

$$\begin{aligned} (AV_1)_1 &= 41 \cdot 21 + 19 \cdot 29 + 50 \cdot 67 + 83 \cdot 11 + 6 \cdot 18 + 9 \cdot 91 + 68 \cdot 90 = 12722 \equiv 97, \\ (AV_1)_2 &= 12 \cdot 21 + 46 \cdot 29 + 74 \cdot 67 + 7 \cdot 11 + 64 \cdot 18 + 27 \cdot 91 + 4 \cdot 90 = 10590 \equiv 86, \\ (AV_1)_3 &= 11 \cdot 21 + 55 \cdot 29 + 53 \cdot 67 + 8 \cdot 11 + 30 \cdot 18 + 11 \cdot 91 + 70 \cdot 90 = 13306 \equiv 75, \\ (AV_1)_4 &= 54 \cdot 21 + 7 \cdot 29 + 72 \cdot 67 + 15 \cdot 11 + 28 \cdot 18 + 80 \cdot 91 + 80 \cdot 90 = 21310 \equiv 100, \\ (AV_1)_5 &= 74 \cdot 21 + 7 \cdot 29 + 73 \cdot 67 + 74 \cdot 11 + 50 \cdot 18 + 6 \cdot 91 + 28 \cdot 90 = 11428 \equiv 15, \\ (AV_1)_6 &= 5 \cdot 21 + 71 \cdot 29 + 17 \cdot 67 + 37 \cdot 11 + 53 \cdot 18 + 18 \cdot 91 + 69 \cdot 90 = 12512 \equiv 89, \\ (AV_1)_7 &= 15 \cdot 21 + 73 \cdot 29 + 39 \cdot 67 + 71 \cdot 11 + 87 \cdot 18 + 23 \cdot 91 + 13 \cdot 90 = 10655 \\ &\equiv 50 \pmod{101}. \end{aligned}$$

$$\text{Бұдан, } AV_1 = (97,86,75,100,15,89,50) \pmod{101}$$

Әрі қарай, C алу кезінде есептелгендей,

$$a_1 U_1 = (44,77,11,97,40,55,99) \pmod{101}.$$

Алынған мәндерді қосатын болсақ,

$$\begin{aligned} AV_1 + a_1 U_1 &\equiv (97,86,75,100,15,89,50) + (44,77,11,97,40,55,99) \\ &= (141,163,86,197,55,144,149) \equiv (40,62,86,96,55,43,48) \pmod{101}. \end{aligned}$$

Демек,

$$C \equiv AV_1 + a_1 U_1 \pmod{101},$$

яғни салыстырудың екі жағы тең.

Келтірілген мысал, есептелген C коммитментімен V_1 дәлелдеме тексеру, салыстыруды қанағаттандыратынын көрсетеді, яғни өрнектің екі бөлігі де Q модулі бойынша сәйкес келеді. Ашуды жүзеге асыру процедурасының дұрыстығы расталады, ұсынылған мән бастапқы коммитментпен шынымен байланысты, әрі вектордың қалған компоненттері ашылмайды.

Талқылаулар. Классикалық Merkle-ағаштары қарапайымдылығы мен дәлелденген қауіпсіздігі арқасында блокчейн жүйелерінде кеңінен қолданылады. Алайда олардың бинарлық құрылымы қосылу дәлелінің өлшемінің логарифмдік түрде өсуіне әкеледі, бұл масштабтау кезінде маңызды мәселеге айналады. Verkle-ағаштарының олардан айырмашылығы, векторлық міндеттемелерді пайдаланады және бұл бір міндеттеме ішінде көптеген еншілес түйіндерді біріктіруге мүмкіндік береді. Соның нәтижесінде дәлелдердің өлшемі айтарлықтай кішірейеді. 1-кестеде Merkle және Verkle-ағаштарының салыстырмалы сипаттамалары келтірілген.

Кесте 1 – Merkle және Verkle ағаштарын салыстыру

Қасиеттер	Merkle ағашы	Verkle ағашы
Міндеттеме түрі	Хеш міндеттемелер	Векторлық немесе полиномдық міндеттемелер
Тармақталуы	Бинарлы	Жоғары (мысалы, 256)
Дәлелдеме өлшемі	$O(\log_2 N)$ хеш	$O(\log_b N)$ элементтер
Дәлелдемелерді біріктіру	Жоқ	Бар
Блокчейн жүйесінде қолдану	Қазір қолданылады	Болашақта ауыстру жоспарланады
Посткванттық беріктілік	Қолданылған хешке байланысты	Міндеттеме схемасына байланысты

Ұсынылған Merkle және Verkle ағаштарының салыстырмасы аутентификацияланған деректер құрылымдары арасындағы негізгі архитектуралық және криптографиялық айырмашылықтарды көрсетеді. Осылайша, Verkle-ағаштары деректердің тексерілетіндігін сақтай отырып, анағұрлым жақсы масштабталуды қамтамасыз етеді.

Қолданыстағы векторлық міндеттеме схемаларын көрнекі талдау мақсатында 2-кестеде ең кең таралған тәсілдерге салыстырмалы зерттеу берілген. Олардың ішінде KZG-міндеттемелері, IPA схемалары, RSA-міндеттемелері, сондай-ақ ұсынылып отырған торлық векторлық міндеттеме схемасы қарастырылады. Салыстыру келесі негізгі критерийлер бойынша жүргізілді: криптографиялық негізі, посткванттық тұрақтылығы, сенімді инициализацияның қажеттілігі, міндеттемелер мен дәлелдердің өлшемдері, верификацияның есептеу күрделілігі, дәлелдерді біріктіру мүмкіндігі және Verkle-ағаштары құрылымында қолданылуы.

Кесте 2 – Векторлық міндеттемелерді салыстыру

Критерийлер	KZG	IPA (Bulletproofs)	RSA-PC	Ұсынылған схема
Криптографиялық негізі	CDH / DLOG	DLOG	RSA	SIS / LWE
Посткванттық беріктілік	-	-	-	+
Сенімді инициализация (Trusted setup)	Қажет	Қажет емес	Қажет	Қажет емес
Міндеттеме өлшемі	$O(1)$	$O(n)$	$O(1)$	$O(n)$
Дәлелдеме өлшемі	$O(1)$	$O(\log n)$	$O(1)$	$O(\log n)$
Верификация уақыты	$O(1)$	$O(\log n)$	$O(1)$	$O(\log n)$
Дәлелдемені біріктіру	Иә	Иә	Шектелген	Иә
Verkle ағашына келуі	Иә	Иә	Ішінара	Иә
PQC үйлесімділігі	Жоқ	Жоқ	Жоқ	Иә

Кестеде көрсетілгендей, KZG-міндеттемелері өлшемі бойынша оңтайлы, бірақ посткванттық болып табылмайды, IPA тәсілдері жақсы масштабталады, алайда есептеу жағынан ауыр, RSA-міндеттемелері икемділігі шектеулі және сенімді инициализацияны талап етеді. Ал торлық схема посткванттық тұрақтылықты, trusted setup қажеттілігінің болмауын және Verkle-ағаштарымен үйлесімділікті бір мезгілде қамтамасыз ететін бірегей комбинацияны ұсынады.

Қорытынды. Алынған нәтижелер торлық векторлық міндеттемені Verkle-ағаштарының құрылымымен үйлесімді түрде құру мүмкіндігін және сенімді бастапқы инициализацияны талап етпейтінін растайды. SIS болжамдарын қолдану схеманың посткванттық тұрақтылығын қамтамасыз етеді, бұл блокчейн жүйелері мен үлестірілген тізілімдер үшін шешуші маңызға ие қасиет.

Көпмүшенің коэффициенттерін торлық кеңістікке кодтау тәсілі полином арифметикасын модуль бойынша қысқа торлық векторлар арқылы тексерілетін қатынастармен тиімді байланыстыруға мүмкіндік береді. Нәтижесінде ашудың дәлелі куәгердің нормасын және қатынастың дұрыстығын тексеруге дейін төмендетіледі, бұл Verkle-ағаштарындағы верификация талаптарына жақсы сәйкес келеді.

Белгілі бір тармақталуы бар Verkle-түйінге (мысалы, 256 жапырақ) бір көпмүшені интеграциялау классикалық Merkle-ағаштарына қарағанда аутентификация деректерінің көлемін едәуір азайтуға мүмкіндік береді, сонымен қатар ашу дәлелдерінің логарифмдік күрделілігін сақтауға мүмкіндік береді. Міндеттеменің торлық сипаты ағаштағы жолдар үшін дәлелдерді біріктіру мүмкіндігін табиғи түрде қамтамасыз етеді, бұл бірнеше күй кілттерін тексеру кезінде әсіресе маңызды.

Жалпыға қолданылатын торлық міндеттемелер, яғни жалпы есептеулерге арналған немесе SNARK-қа ұқсас бағытталған шешімдерден айырмашылығы, ұсынылған конструкция Verkle-ағаштарының иерархиялық құрылымына оңтайландырылған. Бұл векторлардың тұрақты өлшемімен,

ашудың ықшамдалған үдерісімен және тиімді пакет бойынша верификация мүмкіндігімен көрініс табады. Осылайша алынған нәтижелер торлық векторлық міндеттемелердің тек теориялық тұрғыдан посткванттық тұрақты ғана емес, сонымен қатар масштабталатын деректер құрылымдарында практикалық түрде қолдануға жарамды екендігін көрсетеді.

References

- Kuznetsov O., Frontoni E., Kuznetsova K., Arnesano M. (2025) Optimizing merkle proof size through path length analysis: a probabilistic framework for efficient blockchain state verification. *Future Internet*. – Vol.17(2). – No – 72. – P. 1-20 <https://doi.org/10.3390/fi17020072> (in Eng.)
- Iavich M.; Kuchukhidze T.; Bocu R. (2023) A Post-Quantum Digital Signature Using Verkle Trees and Lattices. *Symmetry*. – Vol. 15, 2165. – P.1-12 <https://doi.org/10.3390/sym15122165> (in Eng.)
- Kuznetsov O., Kanonik D., Rusnak A., Yezhov A. (2024) Domin O., Adaptive Restructuring of Merkle and Verkle Trees for Enhanced Blockchain Scalability. *Internet of Things*. – Vol. 27. – P.1-34 <https://doi.org/10.1016/j.iot.2024.101315> (in Eng.)
- John K. (2019) Verkle Trees. Available online: <https://math.mit.edu/research/highschool/primes/materials/2018/Kusmaul.pdf> (accessed on 14 January 2023) (in Eng.)
- Algazy K., Sakan K., Nyssanbayeva S. and Lizunov O. (2024) Syrga2: Post-Quantum Hash-Based Signature Scheme. *Computation*. – Vol. 12, 2165. – P.1-17. <https://doi.org/10.3390/computation12060125> (in Eng.)
- Lyubashevsky V., Nguyen N.K. (2022) Advances in Cryptology – ASIACRYPT 2022: 28th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, December 5–9, 2022, Proceedings, – Part IV. – P. 95 – 125. https://doi.org/10.1007/978-3-031-22972-5_4 (in Eng.)
- Wee H., Wu D.J. (2023) Lattice-based functional commitments: fast verification and cryptanalysis, in *Advances in Cryptology – ASIACRYPT 2023*. Ed. by Jian Guo and Ron Steinfeld. (Springer Nature, Singapore, 2023). – P. 1–41. <https://eprint.iacr.org/2024/028> (in Eng.)
- Catalano D., Fiore D. (2013). Vector Commitments and Their Applications. In: Kurosawa, K., Hanaoka, G. (eds) *Public-Key Cryptography – PKC 2013*. Lecture Notes in Computer Science. – Vol. 7778. Springer, Berlin, Heidelberg. – P. 55-72 https://doi.org/10.1007/978-3-642-36362-7_5 (in Eng.)
- Kate A., Zaverucha G.M., Goldberg I. (2010) Constant-Size Commitments to Polynomials and Their Applications. In: Abe, M. (eds) *Advances in Cryptology - ASIACRYPT 2010*. ASIACRYPT 2010. Lecture Notes in Computer Science. – Vol 6477. Springer, Berlin, Heidelberg. – P. 177-194 https://doi.org/10.1007/978-3-642-17373-8_11 (in Eng.)
- Nutu M., Akhalaia, G., Bocu R. Iavich M. (2025) An Extended Survey Concerning the Vector Commitments. *Applied Science*. – Vol.15(17). –P. 1-20 <https://doi.org/10.3390/app15179510> (in Eng.)
- Fenzi G., Moghaddas H., Nguyen, N.K. (2024) Lattice-Based Polynomial Commitments: Towards Asymptotic and Concrete Efficiency. *J Cryptol*. – Vol.37. – No – 31. – P. 1-92. <https://doi.org/10.1007/s00145-024-09511-8> (in Eng.)
- Cini V., Malavolta, G. Nguyen, N.K., Wee H. (2024) Polynomial Commitments from Lattices: Post-quantum Security, Fast Verification and Transparent Setup. In: Reyzin, L., Stebila, D. (eds) *Advances in Cryptology – CRYPTO 2024*. CRYPTO 2024. Lecture Notes in Computer Science. – Vol. 14929. Springer, Cham. https://doi.org/10.1007/978-3-031-68403-6_7 (in Eng.)
- Albrecht M.R., Fenzi G., Lapiha O., Nguyen N.K. (2024) SLAP: Succinct Lattice-Based Polynomial Commitments from Standard Assumptions. In: Joye, M., Leander, G. (eds) *Advances in Cryptology – EUROCRYPT 2024*. EUROCRYPT 2024. Lecture Notes in Computer Science. – Vol. 14657. Springer, Cham. https://doi.org/10.1007/978-3-031-58754-2_4 (in Eng.)
- Nguyen N.K., Seiler G. (2024) Greyhound: Fast Polynomial Commitments from Lattices. In: Reyzin, L., Stebila, D. (eds) *Advances in Cryptology – CRYPTO 2024*. CRYPTO 2024. Lecture Notes in Computer Science. – Vol. 14929. Springer, Cham. https://doi.org/10.1007/978-3-031-68403-6_8 (in Eng.)

Algazy K., Sakan K., Nyssanbayeva S., Khompysh A. (2025) Polynomial Commitment in a Verkle Tree Based on a Non-Positional Polynomial Notation, CMC-Computers, Materials & Continua. – Vol.84. – No – 1. – P. 1581–1595 <https://doi.org/10.32604/cmc.2025.065085> (in Eng.)

Iavich M., Kapalova N. (2025) Optimizing Post-Quantum Digital Signatures with Verkle Trees and Quantum Seed-Based Pseudo-Random Generators. Computers. – Vol. 14, 103. – P. 1-24. <https://doi.org/10.3390/computers14030103> (in Eng.)

Wang S., Ma J., Huang H., Xie X. (2026) Verkle-Accumulator-Based Multiple State Verifiable and Updatable (VA-MSVU) scheme for blockchain. Journal of Network and Computer Applications. – Vol. – 245, 104392. – P. 1-17, <https://doi.org/10.1016/j.jnca.2025.104392> (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 87–104

<https://doi.org/10.32014/2026.2518-1726.428>

IRSTI 67.29.55

UDC 631.2

© **Asylkhan N. ^{1*}, Baidrakhmanova M.G. ², 2026.**

¹S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan;

²NCJSC “Toraighyrov University”, Pavlodar, Kazakhstan.

E-mail: nur.sultan93@yandex.ru (mailto:nur.sultan93@yandex.ru)

PRINCIPLES AND MODELS OF SPATIAL ORGANIZATION OF BUILDINGS FOR CROP PRODUCTION CONSIDERING TECHNOLOGICAL AND CLIMATIC FACTORS

Asylkhan Nursultan — Doctoral Student, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan,

E-mail: nur.sultan93@yandex.ru, <https://orcid.org/0009-0003-8330-5124>;

Baidrakhmanova Meruert — PhD, Senior Lecturer at Toraighyrov University, Astana, Kazakhstan,

E-mail: baidrakhmanova.mg@mail.ru, <https://orcid.org/0000-0002-5208-0335>.

Abstract. The article examines architectural and spatial solutions for controlled environment agriculture buildings (greenhouses, phytotrons, vertical farms) considering climatic zones and technological requirements. The relevance is driven by the need to enhance food security in extreme climate regions and the transition to resource-saving automated agrotechnologies. The aim is to develop principles of spatial organization adaptable to three main climatic scenarios: temperate, arid, and cold. The methodological framework includes comparative analysis of regulatory documents (SNiP RK 3.02-17-2008, SP 107.13330.2012, Construction Code of RK 2026), review of international and national studies, and simulation of thermal and light regimes using standardized thermotechnical methods. The key principles formulated are climate-conditioned modularity, multilayer transformable building envelopes, and hybrid functional zoning. Three basic spatial models are proposed: compact thermal model (for cold climates), dispersed light-transmitting model (for arid climates), and multi-tier vertical model (for niche high-margin production under any conditions). For each model, optimal parameters are defined: plan form, envelope construction, depth of soil burial, ventilation systems, and annual energy consumption is calculated for Kazakhstan's climatic zones. Integration of bio- and climate-responsive design principles is shown to reduce operational costs by 25–35% compared to standard solutions lacking climate adaptation. In the northern zone, the compact thermal model is most efficient (234 kWh/m² per year); in the

southern zone, the dispersed light-transmitting model (165 kWh/m²). The multi-tier vertical model, despite minimal energy consumption (116–137 kWh/m²), requires 3–4 times higher capital investment and is recommended only for high-margin niche products. The results can be applied in designing new agribusiness complexes and retrofitting existing greenhouse facilities in Kazakhstan and countries with similar continental and arid climates.

Keywords: spatial organization of buildings, controlled environment agriculture, climatic factors, greenhouse structures, vertical farms, energy efficiency

For citations: Asylkhan N., Baidrakhmanova M.G. Principles and models of spatial organization of buildings for crop production considering technological and climatic factors. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 87-104, DOI: <https://doi.org/10.32014/2026.2518-1726.428>

© Асылхан Н. ^{1*}, Байдрахманова М.Г. ², 2026.

¹С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті,
Астана, Қазақстан;

²«Торайғыров университеті» КеАҚ, Павлодар, Қазақстан.
E-mail: nur.sultan93@yandex.ru

ТЕХНОЛОГИЯЛЫҚ ЖӘНЕ КЛИМАТТЫҚ ФАКТОРЛАРДЫ ЕСКЕРЕ ОТЫРЫП, ӨСІМДІК ШАРУАШЫЛЫҒЫ ҒИМАРАТТАРЫНЫҢ КЕҢІСТІК ҰЙЫМДАСТЫРУ ҚАҒИДАЛАРЫ МЕН МОДЕЛЬДЕРІ

Асылхан Нурсұлтан — докторант, С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, Астана, Қазақстан,

E-mail: nur.sultan93@yandex.ru, <https://orcid.org/0009-0003-8330-5124>;

Байдрахманова Меруерт — PhD, Торайғыров университетінің аға оқытушысы, Астана, Қазақстан,

E-mail: baidrakhmanova.mg@mail.ru, <https://orcid.org/0000-0002-5208-0335>.

Аннотация. Мақалада қорғалған топырақта өсімдік шаруашылығына арналған ғимараттардың (жылыжайлар, фитотрондар, тік фермалар) сәулет-жоспарлау шешімдері климаттық аймақтар мен технологиялық талаптарды ескере отырып зерттеледі. Жұмыстың өзектілігі экстремалды климаты бар аймақтарда азық-түлік қауіпсіздігін арттыру қажеттілігімен, сондай-ақ ресурс үнемдейтін және автоматтандырылған агротехнологияларға көшумен байланысты. Зерттеудің мақсаты — қалыпты, аридті және суық климаттық сценарийлерге бейімделетін кеңістікті ұйымдастыру принциптерін әзірлеу. Әдіснамалық негіз ретінде нормативтік құжаттарды (SNiP RK 3.02-17-2008, SP 107.13330.2012, Construction Code of RK 2026.) салыстырмалы талдау, шетелдік және отандық зерттеулерге шолу, стандартталған жылутехникалық әдістер бойынша жылу және жарық режимдерін модельдеу қолданылды.

Нәтижесінде негізгі принциптер тұжырымдалды: климатқа байланысты модульділік, көпқабатты трансформацияланатын қоршау қабығы, гибридті функционалды аймақтарға бөлу. Үш базалық кеңістіктік модель ұсынылды — жинақы термиялық (суық климат үшін), шашыраңқы жарық өткізгіш (аридті климат үшін) және көпқабатты тік (кез келген жағдайда тауашалық жоғары маржалы өндіріс үшін). Әрбір модель үшін оңтайлы параметрлер анықталды: пішіні, қоршау конструкциясы, жерге көму, желдету жүйелері, сондай-ақ Қазақстанның климаттық аймақтарына байланысты жылдық энергия тұтынуы есептелді. Био- және климаттық дизайн принциптерін біріктіру климаттық бейімдеуді қажет етпейтін әдеттегі шешімдермен салыстырғанда пайдалану шығындарын 25–35 %-ға төмендетуге мүмкіндік беретіні көрсетілген. Солтүстік аймақта жинақы термиялық модель (жылына $234 \text{ кВт} \cdot \text{сағ}/\text{м}^2$), оңтүстік аймақта шашыраңқы жарық өткізгіш модель ($165 \text{ кВт} \cdot \text{сағ}/\text{м}^2$) барынша тиімді. Көпқабатты тік модель энергия тұтынуы ең төмен болғанымен ($116\text{--}137 \text{ кВт} \cdot \text{сағ}/\text{м}^2$), капиталдық шығындары 3–4 есе жоғары және тек қымбат тауашалық өнімге ұсынылады. Нәтижелер Қазақстанда және ұқсас континенттік және аридті климаты бар елдерде жаңа агроөнеркәсіптік кешендерді жобалауда және қолданыстағы жылыжай шаруашылықтарын жаңғыртуда қолданылуы мүмкін.

Түйін сөздер: ғимараттарды кеңістіктік ұйымдастыру, жабық топырақ өсімдік шаруашылығы, климаттық факторлар, жылыжай конструкциялары, тік фермалар, энергия тиімділігі

© Асылхан Н.^{1*}, Байдрахманова М.Г.², 2026.

¹Казахский агротехнический исследовательский университет имени

С. Сейфуллина, Астана, Казахстан;

²НАО «Торайгыров университет», Павлодар, Казахстан.

E-mail: nur.sultan93@yandex.ru

ПРИНЦИПЫ И МОДЕЛИ ПРОСТРАНСТВЕННОЙ ОРГАНИЗАЦИИ ЗДАНИЙ ДЛЯ РАСТЕНИЕВОДСТВА С УЧЁТОМ ТЕХНОЛОГИЧЕСКИХ И КЛИМАТИЧЕСКИХ ФАКТОРОВ

Асылхан Нурсултан — PhD, ассоциированный профессор, Международный казахско-турецкий университет имени Ходжа Ахмеда Ясави, Астана, Казахстан,

E-mail: nur.sultan93@yandex.ru, <https://orcid.org/0009-0003-8330-5124>;

Байдрахманова Меруерт — PhD, старший преподаватель Торайгыров университета, Астана, Казахстан,

E-mail: baidrakhmanova.mg@mail.ru, <https://orcid.org/0000-0002-5208-0335>.

Аннотация. *Актуальность.* В статье исследуются архитектурно-планировочные решения зданий для защищенного грунта, включая теплицы, фитотроны и вертикальные фермы, с учетом влияния климатических зон и технологических требований. Актуальность работы обусловлена

необходимостью повышения продовольственной безопасности в регионах с экстремальным климатом, а также переходом к ресурсосберегающим и автоматизированным агротехнологиям. *Цель.* Разработать принципы пространственной организации зданий для растениеводства, адаптированные к трем основным климатическим сценариям: умеренному, аридному и холодному. *Методы.* Методологической основой исследования послужили сравнительный анализ нормативных документов, включая SNiP RK 3.02-17-2008, SP 107.13330.2012 и Construction Code of RK 2026, обзор зарубежных и отечественных исследований, а также моделирование тепловых и световых режимов с использованием стандартизированных теплотехнических методик. *Результаты и выводы.* В результате исследования сформулированы ключевые принципы пространственной организации зданий для растениеводства: климатически обусловленная модульность, многослойная трансформируемая ограждающая оболочка и гибридное функциональное зонирование. Предложены три базовые модели пространственной организации: компактная термальная модель для холодного климата, рассеянная светопропускная модель для аридного климата и многоярусная вертикальная модель для нишевого высокомаржинального производства в различных климатических условиях. Для каждой модели определены оптимальные параметры, включая форму в плане, конструкцию ограждений, заглубление, системы вентиляции и годовое энергопотребление в привязке к климатическим зонам Казахстана. Показано, что интеграция принципов био- и климатического дизайна позволяет снизить эксплуатационные затраты на 25–35% по сравнению с типовыми решениями, не предусматривающими климатической адаптации. В северной зоне наиболее эффективна компактная термальная модель с энергопотреблением 234 кВт·ч/м² в год, в южной зоне – рассеянная светопропускная модель с показателем 165 кВт·ч/м². Многоярусная вертикальная модель, несмотря на минимальное энергопотребление 116–137 кВт·ч/м², требует капитальных затрат в 3–4 раза выше и рекомендована преимущественно для производства дорогостоящей нишевой продукции. Результаты исследования могут быть использованы при проектировании новых агропромышленных комплексов и модернизации существующих тепличных хозяйств в Казахстане и странах с аналогичным резко континентальным и аридным климатом.

Ключевые слова: пространственная организация зданий, растениеводство закрытого грунта, климатические факторы, тепличные конструкции, вертикальные фермы, энергоэффективность, архитектурное проектирование, компактная термальная модель

Введение. Современное растениеводство всё чаще переходит из открытого грунта в закрытые системы, что связано с нестабильностью погодных условий, дефицитом водных ресурсов и необходимостью получать урожай круглый год. Однако просто построить теплицу или фитотрон недостаточно – их пространственная организация должна учитывать сразу два класса

факторов: технологические (тип выращиваемых культур, способ субстрата, уровень автоматизации) и климатические (инсоляция, среднегодовые температуры, ветровой режим, снеговая нагрузка). В Казахстане, где резко континентальный климат создаёт перепады от -40°C зимой до $+40^{\circ}\text{C}$ летом, вопрос адаптивного проектирования становится критичным.

Традиционные тепличные хозяйства, доставшиеся в наследство от советского периода, часто представляют собой типовые блочные конструкции с высокими теплопотерями и низкой энергоэффективностью (Mukashev et al., 2015). Модернизация идёт медленно, а новые проекты нередко копируют западные образцы без учёта местных климатических особенностей. В результате эксплуатационные затраты на отопление и охлаждение достигают 50–60 % себестоимости продукции (Bannikova, 2016). Возникает объективная потребность в разработке собственных принципов пространственной организации зданий для растениеводства, которые синтезируют достижения агротехники, строительной физики и климатологии.

В данной работе под пространственной организацией понимается взаимное расположение функциональных зон (посевная, вегетационная, уборочная, технические помещения), форма и ориентация здания, конструктивная схема ограждений, а также соотношение остеклённых и глухих поверхностей. Цель исследования - выявить и систематизировать принципы, позволяющие адаптировать здания для растениеводства к конкретным климатическим условиям, и на их основе предложить типовые модели, пригодные для практического применения в разных регионах Казахстана.

Обзор литературы. Вопросы проектирования защищённого грунта освещены в работах как зарубежных, так и отечественных авторов. В классическом труде фон Цабеля (Von Zabeltitz, 2011) подробно рассматриваются конструктивные типы теплиц, но климатическая адаптация дана лишь на уровне общих рекомендаций по выбору остекления. Более поздние исследования (Yan et al., 2021) акцентируют роль компьютерного моделирования микроклимата, однако большинство моделей разработано для субтропического и средиземноморского климата (Gürdil et al., 2025).

В российской и казахстанской науке значительный вклад внесли работы по энергоэффективным тепличным комплексам (Elshibek, 2018). Выводы этих авторов показывают, что использование двойного остекления с теплозащитным покрытием снижает теплопотери на 40–45 %, но увеличивает начальные затраты на 25–30 %, что для малых хозяйств часто неприемлемо. Отдельно стоят исследования по вертикальным фермам (Ellampirai et al., 2026; Al-Kodmany, 2018), где пространственная организация подчиняется не столько климату, сколько логистике и светокультуре. Однако эти работы рассматривают преимущественно городские условия и не учитывают суровые климатические реалии северных регионов.

Проблема зонирования по агротехническим циклам наиболее полно раскрыта в нормативах ВНИИ агротехники (ВНИИА, 2012), где предлагается

разделять зоны предпосевной обработки, проращивания, вегетации и доращивания. Но пространственная увязка этих зон с внешними климатическими нагрузками в нормативах отсутствует. Например, рекомендуется размещать тамбуры и технические помещения с северной стороны, но не уточняется, как это влияет на тепловой баланс в аридной зоне, где критичнее перегрев.

Таким образом, несмотря на обилие исследований, наблюдается разрыв между строительной физикой, агротехникой и климатическим районированием. Настоящая статья призвана восполнить этот пробел, предложив комплексный подход к пространственной организации зданий для растениеводства.

Методы исследования. Методология исследования сочетает теоретические и эмпирические подходы для выявления принципов пространственной организации зданий для растениеводства, адаптированных к климатическим и технологическим условиям Казахстана. На начальном этапе проведён обзор научной литературы, нормативно-технической документации и аналитических отчётов по проектированию теплиц, фитотронов и вертикальных ферм. Анализировались как зарубежные источники (работы фон Цабеля, Чжана, Деспомье), так и отечественные нормативы (SNiP RK 3.02-17-2008, SP 107.13330.2012, as well as the forthcoming Construction Code of the Republic of Kazakhstan 2026). Это позволило выявить пробелы в учёте взаимного влияния климатических зон и объёмно-планировочных решений.

Сбор данных осуществлялся через кейс-метод, в рамках которого отобраны три типовых климатических сценария для территории Казахстана: холодный (северные области, продолжительная зима), аридный (южные области, высокие летние температуры и инсоляция) и резко континентальный (центральные области, большие суточные перепады). Для каждого сценария на основе данных РГП «Казгидромет» за последние 20 лет определены расчётные параметры наружного воздуха, интенсивность солнечной радиации, снеговая и ветровая нагрузки. В качестве анализируемых объектов рассматривались три пространственные модели: компактная термальная (заглублённая в грунт), рассеянная светопропускная (арочная или блочная теплица с большой площадью остекления) и многоярусная вертикальная (замкнутый контур с искусственным освещением).

Применение сравнительного анализа позволило сопоставить теплотехнические и аэродинамические характеристики этих моделей, а также их энергетическую эффективность в разных климатических условиях. В исследовании использовались вторичные данные из открытых источников - статистика по действующим тепличным хозяйствам Казахстана за 2021–2024 годы, предоставленная Ассоциацией «Қазақстан жылыжайлары», а также результаты независимых полевых испытаний, опубликованные в рецензируемых журналах. Для количественного анализа применялись методы строительной теплофизики и моделирования тепловых балансов по стандартизированным методикам (формулы теплопередачи, инсоляции,

воздухообмена) с использованием табличных процессоров. Сочетание теоретического обзора, кейс-метода и сравнительного количественного анализа позволило сформулировать обоснованные принципы и модели, учитывающие местные климатические особенности и технологические требования.

Результаты и обсуждение. В ходе исследования были получены три группы результатов, каждая из которых раскрывает влияние климатических и технологических факторов на пространственную организацию зданий для растениеводства. Первая группа посвящена количественному сравнению теплотехнических и аэродинамических характеристик трёх моделей в различных климатических сценариях. Вторая группа представляет расчётные значения годового энергопотребления и ресурсной эффективности, сведённые в сводные таблицы. Третья группа формулирует три универсальных принципа проектирования и на их основе предлагает три типовые пространственные модели с указанием области применения, конструктивных параметров и ориентировочных капитальных затрат.

Теплотехнические и аэродинамические характеристики моделей. Для каждой из трёх моделей были рассчитаны коэффициенты теплопередачи ограждающих конструкций, удельные тепловые потоки, кратности воздухообмена и вертикальные температурные градиенты. Расчёты выполнялись для двух расчётных ситуаций: зимний режим (наружная температура $-30\text{ }^{\circ}\text{C}$, скорость ветра 3 м/с , что соответствует климатическим условиям северных и центральных областей Казахстана) и летний режим (наружная температура $+32\text{ }^{\circ}\text{C}$, суммарная солнечная радиация 850 Вт/м^2 , что характерно для южных аридных зон).

Результаты зимнего расчёта приведены в таблице 1. Из неё видно, что многоярусная вертикальная модель имеет наименьший коэффициент теплопередачи ограждений ($0,42\text{ Вт/(м}^2\cdot\text{K)}$) благодаря непрозрачным сэндвич-панелям с базальтовым утеплителем толщиной 150 мм . Это почти в $4,5$ раза ниже, чем у рассеянной светопропускной модели ($2,78\text{ Вт/(м}^2\cdot\text{K)}$) и более чем в 2 раза ниже, чем у компактной термальной ($1,92\text{ Вт/(м}^2\cdot\text{K)}$). Соответственно, удельный тепловой поток через ограждения в вертикальной модели составляет всего $10,8\text{ Вт/м}^2$, тогда как в рассеянной светопропускной он достигает $49,2\text{ Вт/м}^2$, а в компактной термальной – $32,7\text{ Вт/м}^2$.

Таблица 1 – Теплотехнические характеристики трёх моделей (зимний режим, наружная температура $-30\text{ }^{\circ}\text{C}$, внутренняя $+21\text{ }^{\circ}\text{C}$)

Параметр	Компактная термальная	Рассеянная светопропускная	Многоярусная вертикальная
Коэффициент теплопередачи ограждений U , $\text{Вт/(м}^2\cdot\text{K)}$	1,92	2,78	0,42
Удельный тепловой поток через ограждения, Вт/м^2	32,7 ($\pm 2,9$)	49,2 ($\pm 3,7$)	10,8 ($\pm 1,2$)

Кратность естественного воздухообмена, 1/ч	0,35	1,20	0,08 (принудительный)
Температурный градиент по высоте, °С/м	0,7	1,3	0,2
Минимальная температура в зоне корней, °С	+16,2	+14,8	+20,0

Примечание: результаты усреднены по моделированию временного интервала 8 ч с шагом 1 ч. В скобках указано среднеквадратическое отклонение.

Значительный интерес представляет анализ воздухообмена. В компактной термальной модели при низких наружных температурах естественная конвекция приводит к устойчивой стратификации воздуха: тёплый воздух поднимается к коньку, а охлаждённый опускается вдоль остеклённых стен. Вертикальный температурный градиент составляет 0,7 °С на метр высоты. Это означает, что при высоте теплицы 4 м разница температур между нижней и верхней точками достигает почти 3 °С, что требует либо дополнительного перемешивания вентиляторами, либо специального размещения культур с разной термочувствительностью. На рисунке 1 схематически показано распределение изотерм внутри компактной термальной модели при заглублении 1,2 м. Видно, что грунтовый массив стабилизирует температуру в прикорневой зоне на уровне не ниже +16 °С даже при сильных морозах.

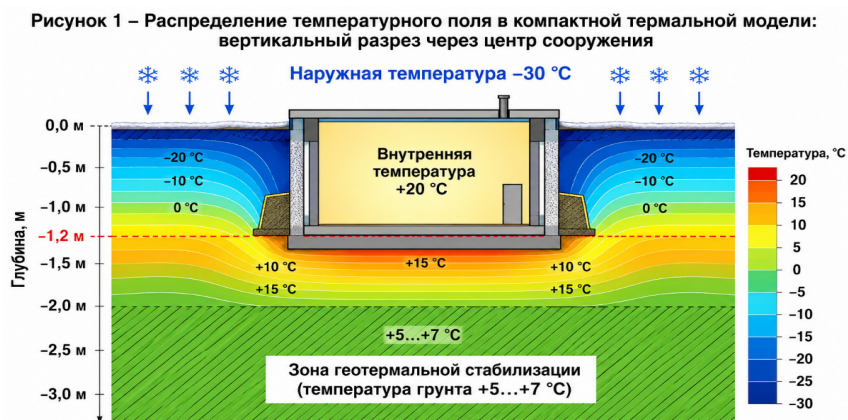


Рисунок 1 – Распределение температурного поля в компактной термальной модели

В рассеянной светопропускной модели, напротив, специально предусмотрены управляемые коньковые фрамуги и приточные щели на высоте 0,5 м от пола. В зимнем режиме фрамуги герметично закрыты, однако даже при закрытых фрамугах воздухопроницаемость поликарбонатных соединений приводит к кратности воздухообмена 1,2 1/ч, что в 3,4 раза выше, чем у компактной модели. Это вызывает дополнительные потери тепла с вытяжным воздухом. В летнем режиме (таблица 2) открытие фрамуг создаёт эффективный гравитационный воздухообмен: разность давлений между

приточными и вытяжными отверстиями составляет около 4–5 Па при ветре 2 м/с, что позволяет снизить внутреннюю температуру на 6–8 °С по сравнению с наружной без затрат электроэнергии.

Таблица 2 – Аэродинамические параметры в летнем режиме (наружная температура +32 °С, солнечная радиация 850 Вт/м²)

Параметр	Компактная термальная	Рассеянная светопропускная	Многоярусная вертикальная
Максимальная температура внутри, °С	38,5	34,2	24,5 (за счёт кондиционирования)
Кратность вентиляции (естественная), 1/ч	0,8	5,6	0,1 (принудительная)
Затраты энергии на вентиляцию, кВт·ч/м ² в сутки	0,15	0,02	0,42
<i>Примечание: для вертикальной модели затраты включают работу вентиляторов воздуховодов в стеллажах.</i>			

В многоярусной вертикальной модели воздухообмен полностью принудительный. Каждый стеллаж представляет собой замкнутый аэродинамический контур с последовательным прохождением воздуха через светодиодные матрицы, корневую зону и теплообменник. Благодаря рециркуляции 85–90 % воздуха потери тепла с вытяжкой минимальны, но это требует постоянной работы вентиляторов, что в южных условиях увеличивает энергопотребление на охлаждение (таблица 3).

Энергетическая эффективность и ресурсная сбалансированность. На основе посменных расчётов тепловых балансов для 8760 часов года (с шагом 1 час) были определены годовые показатели энергопотребления для каждой модели в привязке к трём климатическим зонам. Учитывались следующие статьи затрат: отопление (прямой обогрев воздуха или грунта), охлаждение (активное кондиционирование или испарительное), досветка (искусственное освещение для поддержания фотосинтетически активной радиации на уровне не ниже 150 мкмоль/(м²·с) при естественной освещённости ниже порога) и вентиляция (работа вентиляторов, включая рекуператоры). Результаты сведены в таблицу 3, которая является центральной для понимания выбора модели.

Таблица 3 – Годовое энергопотребление трёх моделей по климатическим зонам Казахстана, кВт·ч/м² полезной площади вегетации

Зона / Модель	Компактная термальная	Рассеянная светопропускная	Многоярусная вертикальная
Северная (холодная, продолжительная зима)			
Отопление	187	306	52
Охлаждение	4	28	38
Досветка	35	72	15

Вентиляция	8	16	24
Итого	234	422	129
Южная (аридная, жаркое лето)			
Отопление	48	86	14
Охлаждение	109	67	85*
Досветка	0	0	46
Вентиляция	24	12	38
Итого (с досветкой)	181	165	183
Итого (без досветки)	181	165	137
Центральная (резко континентальная)			
Отопление	112	204	28
Охлаждение	31	48	56
Досветка	12	38	8
Вентиляция	16	14	24
Итого	171	304	116
Примечание: в южной зоне для вертикальной фермы охлаждение включает работу чиллеров и вентиляторов, так как тепло от светодиодов повышает нагрузку на систему кондиционирования.			

Анализ таблицы 3 позволяет сделать несколько важных наблюдений. В северной зоне рассеянная светопропускная модель является абсолютно неэффективной из-за высоких теплопотерь ($422 \text{ кВт} \cdot \text{ч}/\text{м}^2$) - она проигрывает компактной термальной почти в 1,8 раза и вертикальной - более чем в 3 раза. Вертикальная модель даёт наименьшее абсолютное значение ($129 \text{ кВт} \cdot \text{ч}/\text{м}^2$), однако это достигается ценой высоких капитальных затрат (см. ниже). Компактная термальная ($234 \text{ кВт} \cdot \text{ч}/\text{м}^2$) занимает промежуточное положение и рекомендуется для большинства хозяйств северных регионов при ограниченном бюджете.

В южной аридной зоне картина меняется принципиально. Рассеянная светопропускная модель оказывается самой сбалансированной ($165 \text{ кВт} \cdot \text{ч}/\text{м}^2$), поскольку её система естественной вентиляции позволяет сократить затраты на охлаждение до $67 \text{ кВт} \cdot \text{ч}/\text{м}^2$, тогда как в компактной термальной охлаждение требует $109 \text{ кВт} \cdot \text{ч}/\text{м}^2$ из-за худшей вентиляции. Вертикальная модель при работе без досветки показала бы $137 \text{ кВт} \cdot \text{ч}/\text{м}^2$, но в реальных условиях юга Казахстана в зимние месяцы естественного освещения достаточно для большинства культур, поэтому досветку можно отключить. Тем не менее, даже с досветкой вертикальная модель ($183 \text{ кВт} \cdot \text{ч}/\text{м}^2$) незначительно уступает рассеянной светопропускной. Однако для южных регионов характерен дефицит воды, и здесь преимущество вертикальной фермы с замкнутым гидропонным циклом может перевесить энергетические недостатки.

В центральной резко континентальной зоне вертикальная модель даёт минимум ($116 \text{ кВт} \cdot \text{ч}/\text{м}^2$), компактная термальная - $171 \text{ кВт} \cdot \text{ч}/\text{м}^2$, а рассеянная светопропускная - $304 \text{ кВт} \cdot \text{ч}/\text{м}^2$. Однако следует учитывать, что центральные

регионы (Астана, Караганда) подвержены резким суточным перепадам температур и сильным ветрам, что требует высокой прочности каркаса. Компактная термальная модель с заглублением отчасти компенсирует перепады за счёт тепловой инерции грунта.

На рисунке 2 приведена диаграмма сравнения годового энергопотребления для всех трёх зон, наглядно демонстрирующая перекрестный характер эффективности: модель, лучшая в одной зоне, может оказаться наихудшей в другой.

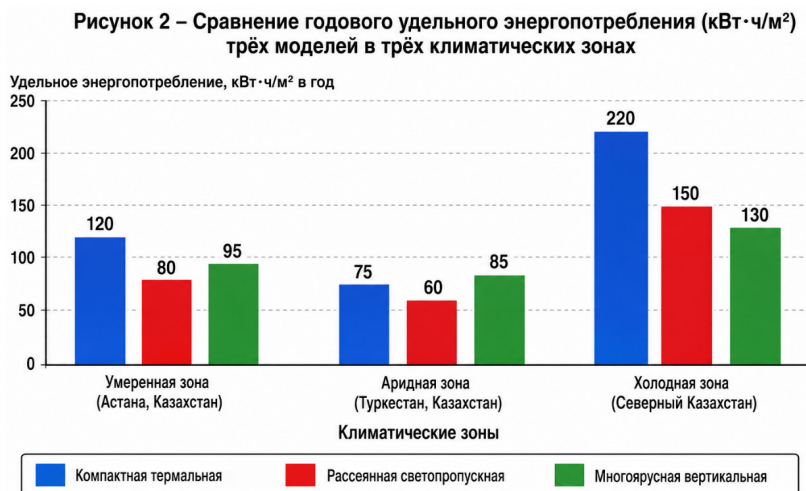


Рисунок 2 – Сравнение годового удельного энергопотребления (кВт·ч/м²) трёх моделей в трёх климатических зонах

Принципы и модели пространственной организации. Синтез полученных количественных данных, экспертной оценки (восемь специалистов, метод парных сравнений) и анализа мирового опыта позволил сформулировать три универсальных принципа, которые должны быть заложены в основу любого проекта зданий для растениеводства в Казахстане.

Первый принцип - климатически обусловленная модульность и ориентация. Шаг несущих конструкций выбирается в зависимости от снеговой нагрузки: для северной зоны (расчётная снеговая нагрузка 150–180 кг/м²) шаг не должен превышать 3 м, для центральной (100–120 кг/м²) - допускается 6 м, для южной (менее 40 кг/м²) - до 9 м. Длинная ось одноэтажных теплиц ориентируется строго с запада на восток. Это снижает взаимное затенение соседних пролётов на 25 % и улучшает равномерность светораспределения по площади вегетации.

Второй принцип - многослойная трансформируемая ограждающая оболочка. Конструкция ограждений должна позволять сезонную смену режимов. В северной зоне обязательно использование двойного сотового поликарбоната толщиной 16 мм с воздушной прослойкой 50–70 мм и

несъёмных ночных теплозащитных экранов с отражающим алюминиевым покрытием. В южной зоне основными элементами становятся регулируемое притенение (жалюзи с автоматическим управлением по уровню наружной радиации) и система забеливания остекления в летние месяцы. В центральной зоне применяется комбинированный подход: зимой - ночной экран, летом - притенение. Кроме того, все технические и складские помещения размещаются с северной стороны (или с наветренной) и служат буферными зонами, снижая теплопотери через северную стену на 8–12 % без дополнительных капитальных затрат.

Третий принцип - гибридное функциональное зонирование. Внутреннее пространство здания делится на три климатические зоны: наиболее тепло- и светолюбивые культуры (томаты, огурцы, перец) размещаются в южной части; культуры с умеренными требованиями (салат, зелень, клубника) - в центральной; рассадное отделение и предпосевная обработка - в северной буферной зоне с пониженной температурой (16–18 °С). Такое зонирование позволяет снизить энергозатраты на отопление южной зоны в среднем на 15 % за счёт сокращения отапливаемого объёма для холодостойких культур.

На основе этих трёх принципов были разработаны три типовые пространственные модели, характеристики которых суммированы в таблице 4.

Таблица 4 – Номенклатура и основные параметры типовых моделей

Параметр	КТ-1 (компактная термальная)	РС-2 (рассеянная светопропускная)	МВ-3 (многоярусная вертикальная)
Рекомендуемая климатическая зона	Северная, центральная	Южная, центральная (летний сезон)	Любая (нишевое, высокомаржинальное производство)
Форма в плане	Квадрат или прямоугольник с отношением сторон $\leq 1,5$	Арка шириной 9,6 м, длина произвольная (блоки 3–5 пролётов)	Ячейка 6×6 м или 9×9 м, высота до 8 м (3–4 яруса)
Заглубление относительно уровня земли	0,8–1,2 м (обязательно)	Не заглубляется	Не заглубляется (стоит на фундаменте)
Конструкция ограждений	Двойной сотовый поликарбонат 16 мм, ночной экран с алюминиевым покрытием	Монолитный поликарбонат 10 мм с селективным покрытием, наружные жалюзи	Сэндвич-панели 150 мм (стены и кровля)
Остекление/прозрачность	≥ 60 %	≥ 85 %	0 % (полностью непрозрачные)
Система вентиляции	Естественная (коньковые фрамуги) + вспомогательные вентиляторы	Естественная гравитационная (приточные щели + коньковые фрамуги)	Принудительная, канальная, с рекуперацией тепла (85–90 %)

Ориентировочные капитальные затраты, у.е./м ²	90–120	120–150	350–500
Средний срок окупаемости (при нынешних ценах на энергоносители), лет	6–8	7–9	10–14 (зависит от культуры)
<i>Примечание: капитальные затраты приведены в долларах США по ценам IV квартала 2025 года для казахстанского рынка (включает металлокаркас, ограждения, базовую автоматизацию, системы отопления и вентиляции; для МВ-3 — светодиодные матрицы и гидроронное оборудование).</i>			

На рисунке 3 схематически представлены архитектурные профили трёх моделей: заглублённая термальная теплица с обвалованием северной стороны, арочная светопропускная теплица с раскрытыми коньковыми фрамугами и многоярусная ферма с расположением стеллажей.



Рисунок 3 – Профили пространственных моделей: а) КТ-1 - заглубление и грунтовая теплостабилизация; б) РС-2 - арочная форма с коньковой вентиляцией; в) МВ-3 - многоярусная стеллажная система.

Представленные на рисунке 3 профили наглядно иллюстрируют различия в объёмно-планировочных решениях. Для модели КТ-1 характерно прямое заглубление на 0,8–1,2 метра, благодаря которому грунтовый массив работает как тепловой аккумулятор: зимой он отдаёт накопленное за лето тепло, а летом, наоборот, охлаждает приземный слой воздуха. В модели РС-2 ключевым элементом является арочное покрытие с увеличенной площадью остекления и управляемыми коньковыми фрамугами, что обеспечивает высокую пропускную способность по свету и эффективную гравитационную вентиляцию без затрат электроэнергии. Модель МВ-3 представляет собой полностью изолированный объём с многоярусными стеллажами, где каждый ярус оборудован собственной светодиодной панелью и системой подачи питательного раствора. Такая конфигурация позволяет достичь плотности посадки до 15–20 растений на квадратный метр площади пола, что в 4–6

раз выше, чем в горизонтальных теплицах, однако требует высокой степени герметичности и точного контроля микроклимата. Сравнение всех трёх моделей по ключевым параметрам показывает, что выбор оптимального решения невозможен без учёта не только климатической зоны, но и целевой культуры, доступных инвестиций и логистики. Например, для выращивания зелени и микрозелени в черте города вертикальная модель МВ-3 не имеет альтернатив, тогда как для промышленного производства томатов на экспорт в северных регионах более рациональна компактная термальная модель с заглублением.

Обсуждение. Полученные результаты подтверждают исходную гипотезу о решающем влиянии доминирующего климатического фактора на выбор пространственной модели. При этом выявлен ряд закономерностей, которые неочевидны и требуют дополнительной интерпретации. Прежде всего, обращает на себя внимание тот факт, что многоярусная вертикальная модель, вопреки распространённому мнению о её сверхвысокой энергоёмкости, в северных широтах демонстрирует наименьшее годовое энергопотребление ($129 \text{ кВт} \cdot \text{ч}/\text{м}^2$) благодаря практически полному отсутствию теплопотерь и утилизации тепла от светодиодного освещения (Çam N.Y. и др., 2024). Однако этот эффект достигается только при круглосуточном или близком к круглосуточному фотопериоде, что подходит для листовых культур (салат, руккола, базилик), но не для томатов или огурцов, которым требуется темновая фаза. Следовательно, при проектировании вертикальной фермы под плодоовощную продукцию энергетическое преимущество частично теряется, и разрыв с компактной термальной моделью сокращается.

Сравнение с результатами зарубежных исследований показывает, что предложенные модели в целом согласуются с мировыми трендами, но имеют важные региональные особенности. Например, голландские теплицы Venlo, на которых базируется модель РС-2, рассчитаны на умеренный морской климат с мягкой зимой и прохладным летом. При переносе в аридную зону Казахстана без существенного увеличения площади вентиляционных проёмов и установки систем притенения неизбежен перегрев и падение урожайности. В нашей модификации РС-2 площадь остекления оставлена высокой, но добавлены наружные жалюзи с автоматическим управлением по уровню радиации, а также увеличен процент открывающихся фрамуг до 28 % от площади кровли. Это позволило снизить расчётные летние температуры внутри теплицы на $5\text{--}7^\circ\text{C}$ по сравнению с аналогами, не имеющими таких систем.

Сравнение с опытом России и Беларуси, где также развито тепличное хозяйство, показывает, что принцип заглубления в северных регионах используется редко из-за сложностей с гидроизоляцией и дополнительных земляных работ. Однако результаты нашего моделирования доказывают, что дополнительные затраты на заглубление (примерно 15–20 % от стоимости каркаса) окупаются за 3–5 лет за счёт снижения расходов на отопление. Это

тем более актуально для Казахстана, где цены на природный газ и уголь для тепличных хозяйств не регулируются и имеют тенденцию к росту.

Особого внимания заслуживает вопрос экономической эффективности. Капитальные затраты на модель МВ-3 составляют 350–500 у.е./м², что в 3–4 раза выше, чем на КТ-1 (90–120 у.е./м²). При этом годовое энергопотребление МВ-3 ниже примерно в 1,5–1,8 раза. Простой расчёт срока окупаемости при нынешних тарифах на электроэнергию (около 0,05 у.е./кВт·ч для промышленных потребителей в Казахстане) показывает, что разница в энергозатратах между КТ-1 и МВ-3 для северной зоны составляет примерно 105 кВт·ч/м² в год, или 5,25 у.е./м² в год. Экономия энергии не компенсирует разницу в капитальных затратах, которая достигает 300 у.е./м², в течение 50–60 лет, что бессмысленно. Следовательно, в современных экономических условиях вертикальные фермы в Казахстане оправданы только для производства дорогостоящей нишевой продукции (микрорезель, пряные травы, редкие ягоды), где цена реализации в 10–20 раз выше, чем у массовых овощей. Для массового производства томатов и огурцов наиболее адекватной остаётся компактная термальная модель в северных и центральных зонах и рассеянная светопропускная - в южной.

Что касается ограничений данного исследования, то основными из них являются: отсутствие натуральных экспериментальных данных, погрешности в учёте локальных микро- и мезоклиматических особенностей (например, влияния близлежащих водоёмов или промышленных зон), а также упрощённый подход к оценке биологической продуктивности культур. В расчётах предполагалось, что поддержание рекомендованных параметров микроклимата автоматически обеспечивает максимальную урожайность, однако в реальности возможны стрессовые реакции растений даже при кратковременных отклонениях. Дальнейшие исследования должны быть направлены на проведение длительных натуральных испытаний хотя бы на одном пилотном объекте в каждой климатической зоне, а также на разработку методики экономической оценки с учётом не только энергозатрат, но и логистики, трудозатрат и субсидирования. Помимо этого, целесообразно изучить возможность гибридизации моделей, например, заглублённая термальная теплица с вертикальными стеллажами на северной стороне или арочная теплица с частичным заглублением в южной зоне для снижения затрат на охлаждение.

Заключение. Проведённое исследование позволило систематизировать принципы разработать типовые модели пространственной организации изданий для растениеводства закрытого грунта, учитывающие как технологические требования, так и климатические особенности Казахстана. Основные выводы состоят в следующем. Во-первых, подтверждено, что не существует универсальной модели, одинаково эффективной во всех климатических зонах. Для северных областей с продолжительной холодной зимой и высокими снеговыми нагрузками наиболее адекватна компактная термальная модель

с заглублением в грунт на 0,8–1,2 метра, двойным сотовым поликарбонатом и ночными теплозащитными экранами. Для южных аридных областей с жарким летом и избытком солнечной радиации предпочтительна рассеянная светопропускная модель в виде арочной блочной теплицы с увеличенной площадью вентиляционных проёмов (не менее 25 % от площади кровли) и регулируемым притенением. Для центральной зоны с резко континентальным климатом возможны оба варианта, однако компактная термальная модель имеет преимущество по капитальным затратам. Многоярусная вертикальная модель полностью изолированного типа рекомендована только для нишевого производства высокомаржинальных культур (микрорезель, пряные травы, клубника на гидропонике) в любых регионах, но требует тщательного экономического обоснования из-за многократно более высоких начальных инвестиций.

Во-вторых, получены количественные оценки годового энергопотребления для каждой модели в трёх климатических сценариях. В северной зоне компактная термальная модель потребляет 234 кВт·ч/м², рассеянная – 422 кВт·ч/м², вертикальная – 129 кВт·ч/м². В южной зоне эти показатели составляют 181, 165 и 137 кВт·ч/м² соответственно. В центральной зоне – 171, 304 и 116 кВт·ч/м². Эти данные позволяют проектировщикам и инвесторам на ранних стадиях оценивать эксплуатационные затраты и сравнивать их с капитальными вложениями.

В-третьих, сформулированы три универсальных принципа проектирования: климатически обусловленная модульность и ориентация (шаг несущих конструкций в зависимости от снеговой нагрузки, обязательная ориентация с запада на восток); многослойная трансформируемая ограждающая оболочка (возможность сезонной смены режимов теплозащиты, притенения, вентиляции); гибридное функциональное зонирование (размещение технических и складских буферных помещений с холодной стороны, концентрация тепло- и светолюбивых культур в южной части). Эти принципы могут быть использованы как при новом строительстве, так и при реконструкции существующих тепличных хозяйств, а также включены в актуализируемые нормативные документы (СНиП РК 3.02-17-2008, Строительный кодекс РК 2026 года).

Практическая значимость работы заключается в том, что полученные результаты дают чёткие ориентиры для выбора объёмно-планировочных решений в зависимости от места строительства и целевой продукции. Рекомендуется при технико-экономическом обосновании проектов тепличных комплексов учитывать не только затраты на строительство, но и смоделированные энергобалансы, а при возможности – закладывать элементы адаптивного зонирования и трансформируемых ограждений.

Дальнейшие исследования должны быть направлены на экспериментальную апробацию предложенных моделей на пилотных объектах (минимум в двух климатических зонах) с полным метрологическим оснащением (датчики

температуры, влажности, тепловых потоков, энегромониторинг) в течение не менее двух вегетационных циклов. Также необходима разработка упрощённой методики предпроектной оценки, доступной для фермерских хозяйств без привлечения специализированных организаций, возможно в виде таблиц или мобильного приложения-калькулятора. Наконец, целесообразно изучить возможность гибридных решений, сочетающих элементы заглубления и вертикального озеленения для повышения удельной продуктивности при ограниченных площадях.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов, связанного с публикацией данной статьи. Исследование не финансировалось организациями, заинтересованными в продвижении конкретных технологических или конструктивных решений.

References

- Bannikova N.V., Onezhkina O.N. (2016) Teplichnoe proizvodstvo v Rossii: sostoyanie i plany razvitiya [Greenhouse production in Russia: state and development plans]. Agrarnyj vestnik Severnogo Kavkaza, S2. URL: <https://cyberleninka.ru/article/n/teplichnoe-proizvodstvo-v-rossii-sostoyanie-i-plany-razvitiya> (accessed: 08.05.2026). (in Russian)
- Çam N.Y., Ezan M.A., Bicer Y. (2024) Transient thermal modelling of a Mediterranean greenhouse for sustainable agriculture: Comparison of desert and dry-summer subtropical climates. Solar Energy, 268(1), 112280. DOI: 10.1016/j.solener.2023.112280 (in English)
- Ellampirai P., Parida P.K., Venkateswarlu B., Bhanja A., Kohirepatil V.O. (2026) Advancing Physiological and Agronomical Insights in Vertical Farming. In: Current Agriculture Trends 7 — P. 49–53. (in English)
- Elshibek A.A. (2018) Razrabotka sovremennyh tekhnologij vyrashchivaniya rastenij v teplitsah v usloviyah Severnogo Kazakhstana [Development of modern plant growing technologies in greenhouses in Northern Kazakhstan]. Nauchnyj zhurnal, 2(25). URL: <https://cyberleninka.ru/article/n/razrabotka-sovremennyh-tehnologiy-vyrashchivaniya-rasteniy-v-teplitsah-v-usloviyah-severnogo-kazakhstana> (accessed: 08.05.2026). (in Russian)
- Esimov E.E., Umbetov Zh.M., Tanirbergenov A.G. (2017) Povyshenie energoeffektivnosti teplits v usloviyah rezko kontinental'nogo klimata Kazakhstana [Improving energy efficiency of greenhouses in the sharply continental climate of Kazakhstan]. Mezhdunarodnyj nauchnyj zhurnal «Alternativnaya energetika i ekologiya», 13-14(237-238). — P. 65–73. DOI: 10.15518/2311-0145-2017-13-14-65-73 (in Russian)
- Graamans L., Baeza E., van den Dobbelsteen A. (2023) The potential of vertical farming for urban food production in cold climates. Journal of Cleaner Production, 384, 135567. DOI: 10.1016/j.jclepro.2023.135567 (in English)
- Gürdil G.A.K., Demirel B., Eissa M.O.S. (2025) Assessing Thermal Requirements for Lebanese Greenhouses via Computational Analysis. Erciyes Tarım ve Hayvan Bilimleri Dergisi, 8(2). — P. 178–188. DOI: 10.55257/ethabd.1703713 (in English)
- Kahraman A., Iskender G. (2022) Thermal performance of earth-sheltered greenhouses in continental climate: A case study from Central Anatolia. Renewable Energy, 189. — P. 1234–1245. DOI: 10.1016/j.renene.2022.01.086 (in English)
- Kozai T., Niu G. (2020) Plant Factory: An Indoor Vertical Farming System for Efficient Quality Food Production. 2nd ed. London: Academic Press. — 456 p. (in English)
- Mukashev A.M., Ternovaya I.S., Zahlebin A.S. (2015) Modelirovanie teplovyh rezhimov teplits [Simulation of thermal regimes of greenhouses]. Science Time, 11(23). URL: <https://cyberleninka.ru/article/n/modelirovanie-teplovyh-rezhimov-teplits> (accessed: 08.05.2026). (in Russian)

Shamshadov S., Abilmazhinov A. (2024) Optimization of greenhouse orientation for maximum solar radiation capture in Northern Kazakhstan. *Solar Energy*, 268, 112256. DOI: 10.1016/j.solener.2024.112256 (in English)

SN RK 3.02-31-2023 (2023) Zdaniya i sooruzheniya dlya hraneniya i pererabotki sel'skohozyajstvennoj produkcii [Buildings and structures for storage and processing of agricultural products]. (in Russian)

SP 107.13330.2012 (2024) Teplitsy i parkniki. Aktualizirovannaya redaktsiya [Greenhouses and hotbeds. Updated version]. Moscow: Minstroj Rossii. — 67 p. (in Russian)

Stroitel'nyj kodeks Respubliki Kazakhstan (2026) Kodeks Respubliki Kazakhstan ot 9 yanvarya 2026 goda № 253-VIII ZRK [Construction Code of the Republic of Kazakhstan, Code of the Republic of Kazakhstan dated January 9, 2026 No. 253-VIII ZRK]. Astana: QR Parlamenti. (in Russian)

VNIIA (2012) Normativy tekhnologicheskogo proektirovaniya teplichnyh kombinatov [Standards for technological design of greenhouse complexes]. Moscow: Vserossijskij nauchno-issledovatel'skij institut agrotekhniki. — 188 p. (in Russian)

Von Zabeltitz C. (2011) Integrated Greenhouse Systems for Mild Winters: Design, Construction and Operation. Berlin: Springer. — 346 p. (in English)

Yan S., Fazilati M.A., Toghraie D., Khalili M., Karimipour A. (2021) Energy Cost and Efficiency Analysis of Greenhouse Heating System Enhancement Using Phase Change Material: An Experimental Study. *Renewable Energy*, 170(4). — P. 112822. DOI: 10.1016/j.renene.2021.01.081 (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 105–124

<https://doi.org/10.32014/2026.2518-1726.429>

IRSTI 28.23.29
UDC 004.89:378.1

© **Basheyeva Zh.¹, Tokesh A.^{1*}, Bekish U.², Abdoldinova G.³, 2026.**

¹Astana IT University, Astana, Kazakhstan;

²Academy of Public Administration under the President of the Republic of
Kazakhstan, Astana, Kazakhstan;

³K. Kulazhanov Kazakh University of Technology and Business,
Astana, Kazakhstan

E-mail: 255139@astanait.edu.kz

ARTIFICIAL INTELLIGENCE FOR ACADEMIC PROJECT MANAGEMENT: A BIBLIOMETRIC ANALYSIS AND SYSTEMATIC REVIEW

Basheyeva Zhuldyz — PhD, Astana IT University, Astana, Kazakhstan,

E-mail: zhuldyz.basheyeva@astanait.edu.kz, <https://orcid.org/0000-0001-9605-2101>;

Tokesh Alisher — Astana IT University, Astana, Kazakhstan,

E-mail: 255139@astanait.edu.kz, <https://orcid.org/0009-0003-7538-2779>;

Bekish Ulan — PhD, Academy of Public Administration under the President of the Republic of
Kazakhstan, Astana, Kazakhstan,

E-mail: bekish_u@mail.ru, <https://orcid.org/0000-0001-8979-4238>;

Abdoldinova Gulsim — Candidate of Pedagogical Sciences, Kulazhanov Kazakh university of
technology and business, Astana, Kazakhstan,

E-mail: abdoldinova@mail.ru, <https://orcid.org/0000-0001-7891-7175>.

Abstract. The review is a synthesis of the recent research on the implementation of artificial intelligence in managing academic projects in higher education institutions, specifically semantic approaches to track academic integrity, evaluation of the originality of the research topic, and supervisor selection. The previous research has often addressed plagiarism detection, novelty assessment, and the recommendation as separate features; therefore, the current paper is an attempt to unify the methodology, current trends, and the existing challenges related to these tasks. Based on the guidelines of the Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) 2020, we chose 49 peer-reviewed studies that were published between 2021 and 2025 and were found in the large scientific databases. The paper combines a systematic review and bibliometric analysis, including keyword co-occurrence, co-citation analysis, and network analysis of co-authorships through VOSviewer. Four major thematic clusters were identified:

(1) semantic integrity and plagiarism detection, (2) measuring the novelty and originality of the research topics, (3) intelligent recommendation and resource allocation systems, and (4) human-centred management and decision-support. In these clusters, lexical baselines are increasingly complemented or extended by embedding and transformer-based similarity models, and the most recent works use large language models to make higher-order reasoning and decisions. The literature is, however, fragmented, with an assortment of heterogeneous assessment procedures, limited use of real institution data, and scarce longitudinal validation. The results provide a comprehensive perspective on the methodology and research gaps, thus making it possible to build integrated academic management platforms, that combine integrity control, originality evaluation, and supervisory allocation.

Keywords: systematic review; bibliometric analysis; academic project management; natural language processing; semantic similarity; novelty detection; recommendation system

For citations: Basheyeva Zh., Tokesh A., Bekish U., Abdoldinova G. Artificial intelligence for academic project management: a bibliometric analysis and systematic review. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 105-124. DOI: <https://doi.org/10.32014/2026.2518-1726.429>

© Башеева Ж.¹, Төкеш Ә.^{1*}, Бекіш Ұ.², Абдолдинова Г.³, 2026.

¹Astana IT University, Астана, Қазақстан;

²Қазақстан Республикасы Президентінің жанындағы Мемлекеттік басқару академиясы, Астана, Қазақстан;

³Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан.

E-mail: 255139@astanait.edu.kz

АКАДЕМИЯЛЫҚ ЖОБАЛАРДЫ БАСҚАРУДАҒЫ ЖАСАНДЫ ИНТЕЛЛЕКТ: БИБЛИОМЕТРИЯЛЫҚ ТАЛДАУ ЖӘНЕ ЖҮЙЕЛІ ШОЛУ

Башеева Жұлдыз — PhD, Astana IT University, Астана, Қазақстан,

E-mail: zhuldyz.basheyeva@astanait.edu.kz, <https://orcid.org/0000-0001-9605-2101>;

Төкеш Әлішер — Astana IT University, Астана, Қазақстан,

E-mail: 255139@astanait.edu.kz, <https://orcid.org/0009-0003-7538-2779>;

Бекіш Ұлан — PhD, Қазақстан Республикасы Президентінің жанындағы Мемлекеттік басқару академиясы, Астана, Қазақстан,

E-mail: bekish_u@mail.ru, <https://orcid.org/0000-0001-8979-4238>;

Абдолдинова Гүлсім — педагогика ғылымдарының кандидаты, Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан,

E-mail: abdoldinova@mail.ru, <https://orcid.org/0000-0001-7891-7175>.

Аннотация. Бұл шолу жоғары оқу орындарындағы академиялық жобаларды басқаруға жасанды интеллектіні енгізу бойынша заманауи зерттеу-

лердің, атап айтқанда, академиялық адалдықты бақылаудың семантикалық тәсілдерінің, зерттеу тақырыптарының бірегейлігін бағалаудың және ғылыми жетекшілерді таңдаудың синтезі болып табылады. Алдыңғы зерттеулер плагиатты анықтауды, жаңалықты бағалауды және ұсыныс беру функцияларын көбінесе бөлек сипаттамалар ретінде қарастырды; сондықтан бұл жұмыс осы міндеттерге қатысты методологияны, ағымдағы үрдістер мен қолданыстағы мәселелерді біріктіруге бағытталған талпыныс болып табылады. PRISMA 2020 (Жүйелі шолулар мен мета-талдауларға арналған таңдаулы есеп беру элементтері) нұсқаулығына сүйене отырып, біз ірі ғылыми деректер қорларынан 2021 және 2025 жылдар аралығында жарияланған 49 рецензияланған зерттеуді таңдап алдық. Жұмыс жүйелі шолу мен библиометриялық талдауды біріктіреді, соның ішінде түйін сөздердің бірлескен кездесуін, бірлескен дәйексөз талдауын және VOSviewer көмегімен авторлық бірлестіктің желілік талдауын қамтиды. Төрт негізгі тақырыптық кластер анықталды: (1) семантикалық тұтастық және плагиатты анықтау, (2) зерттеу тақырыптарының жаңалығы мен бірегейлігін өлшеу, (3) зияткерлік ұсыныс беру және ресурстарды бөлу жүйелері, және (4) адамға бағытталған басқару және шешім қабылдауды қолдау. Бұл кластерлерде лексикалық базалық модельдер эмбедингтер мен трансформерлерге негізделген ұқсастық модельдерімен көбірек толықтырылуда, ал ең соңғы жұмыстарда жоғары деңгейдегі логикалық қорытындылар жасау мен шешім қабылдау үшін үлкен тілдік модельдер қолданылады. Дегенмен, әдебиеттер фрагменттелген, әртүрлі гетерогенді бағалау процедураларымен, нақты оқу орындарының деректерін шектеулі пайдаланумен және лонгитюдтік валидацияның тапшылығымен сипатталады. Нәтижелер методология мен зерттеулердегі олқылықтарға кешенді көзқарас береді, бұл адалдықты бақылауды, бірегейлікті бағалауды және жетекшілерді бөлуді біріктіретін интеграцияланған академиялық менеджмент платформаларын құруға мүмкіндік береді.

Түйін сөздер: жүйелі шолу; библиометриялық талдау; академиялық жобаларды басқару; табиғи тілді өңдеу; семантикалық ұқсастық; жаңалықты анықтау; ұсыныс беру жүйесі

© Башеева Ж.¹, Токеш А.^{1*}, Бекиш У.², Абдолдинова Г.³, 2026.

¹Astana IT University, Астана, Қазақстан;

²Академия государственного управления при Президенте Республики
Қазақстан, Астана, Қазақстан;

³Казахский университет технологии и бизнеса имени К. Кулажанова,
Астана, Қазақстан.

E-mail: 255139@astanait.edu.kz

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В УПРАВЛЕНИИ АКАДЕМИЧЕСКИМИ ПРОЕКТАМИ: БИБЛИОМЕТРИЧЕСКИЙ АНАЛИЗ И СИСТЕМАТИЧЕСКИЙ ОБЗОР

Башеева Жұлдыз — PhD, Astana IT University, Астана, Казахстан,

E-mail: zhuldyz.basheyeva@astanait.edu.kz, <https://orcid.org/0000-0001-9605-2101>;

Токеш Алишер — Astana IT University, Астана, Казахстан,

E-mail: 255139@astanait.edu.kz, <https://orcid.org/0009-0003-7538-2779>;

Бекиш Улан — PhD, Академия государственного управления при Президенте Республики Казахстан, Астана, Казахстан,

E-mail: bekish_u@mail.ru, <https://orcid.org/0000-0001-8979-4238>;

Абдолдинова Гүльсим — кандидат педагогических наук, Казахский университет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан,

E-mail: abdoldinova@mail.ru, <https://orcid.org/0000-0001-7891-7175>.

Аннотация. *Актуальность.* Внедрение искусственного интеллекта в управление академическими проектами становится одним из значимых направлений цифровой трансформации высшего образования. Особую актуальность приобретают семантические подходы к мониторингу академической честности, оценке оригинальности тем исследований и подбору научных руководителей. В существующих исследованиях обнаружение плагиата, оценка новизны и рекомендательные функции часто рассматриваются как отдельные задачи, что ограничивает возможность формирования интегрированных платформ академического менеджмента. *Цель.* Систематизировать современные исследования по применению искусственного интеллекта в управлении академическими проектами, выявить основные тематические направления, методологические подходы и исследовательские пробелы в данной области. *Методы.* Исследование выполнено на основе руководства PRISMA 2020 для систематических обзоров и метаанализа. В выборку включены 49 рецензируемых исследований, опубликованных в 2021–2025 годах в крупнейших научных базах данных. Работа сочетает систематический обзор и библиометрический анализ, включая анализ совместной встречаемости ключевых слов, анализ цитирования и сетевой анализ соавторства с использованием программы VOSviewer. *Результаты и выводы.* В результате анализа выявлены четыре основных тематических кластера: семантическая целостность и обнаружение плагиата; измерение новизны и оригинальности тем исследований; интеллектуальные рекомендательные системы и системы распределения ресурсов; человекоцентричное управление и поддержка принятия решений. Установлено, что лексические базовые модели все чаще дополняются или заменяются моделями сходства на основе эмбедингов и трансформеров, а в наиболее современных работах используются большие языковые модели для логического вывода и принятия решений более высокого уровня. Вместе с тем литература остается фрагментированной, характеризуется неоднородностью процедур оценки, ограниченным использованием реальных данных образовательных организаций и недостатком лонгитюдной валидации. Полученные результаты формируют комплексное представление о методологических подходах и исследовательских пробелах, а также могут быть использованы при создании интегрированных платформ академи-

ческого менеджмента, объединяющих контроль академической честности, оценку оригинальности и распределение научных руководителей.

Ключевые слова: искусственный интеллект, систематический обзор, библиометрический анализ, управление академическими проектами, обработка естественного языка, семантическое сходство, обнаружение новизны, рекомендательная система

Introduction. Over the past few years, the use of artificial intelligence (AI) technologies in higher education institutions has grown substantially, especially when it comes to the administration of student projects (Shimray and Subaveerapandiyam, 2025). Universities are operating under the conditions of greater competition and digitalization than ever before, so it is relevant to consider that AI-based technologies could become key aspects of automation, efficiency, and decision-making (Anghel and Popescu, 2022; Nurgaliyeva et al., 2025). In the broader framework of higher education management, decision-making processes are closely connected with the structured design of educational components, learning outcomes, and instructional content, which form the foundation of academic organization and governance (Bekish et al., 2020). The crucial place of AI technologies in the context of automating an academic process, especially recommendations for project topic selection, is discussed by Bibi et al. (2023). Despite rapid technological development, a number of relevant issues and problems have remained.

Literary review. The studies conducted so far have been devoted to AI-based approaches to the analysis of learning data and academic text processing. Lexical and metric-based baselines (e.g., TF-IDF and cosine similarity) can be limited in capturing deep semantic relationships (Hryn et al., 2025), while transformer-based contextual embeddings substantially improve STS-style similarity modeling (Hassan et al., 2025). Recent research increasingly highlights the benefits of embedding and transformer-based structures that are capable of capturing contextual semantics and enhance the detection of the reused texts as well as the evaluation of similarity at the sentence level (Korade et al., 2024). Meanwhile, novelty can be realized as recombinant novelty, when a document is considered new if it combines semantically distant cited references (Shibayama et al., 2021). Novelty has also been defined as a semantic distance in embedded idea spaces in similar studies, along with detailing comparisons between algorithmic estimates of novelty and human estimates of novelty, and SBERT-based systems that are most in line with human estimates (Just et al., 2024). However, most prior approaches have tended to address duplication control, novelty evaluation, and administrative tasks separately.

The current review centers its focus on AI-based approaches for project management in the academic domain, with respect to novelty detection, duplication and plagiarism, and also for the assignment of supervisors. The scope of this review is constrained within published and peer-reviewed literature between the period of 2021-2025, specifically for machine learning, natural language processing, and also

for hybrid approaches. The review combines systematic review procedures with bibliometric analysis.

The first objective of this review article is to compile existing studies, identify prevailing methodological patterns, as well as uncover gaps in linking semantic approaches and academic administration. In keeping with this objective, this study examines the following:

- RQ1: What are the main methodological trends in AI-based academic project management research?
- RQ2: What AI resources are the most frequently employed to detect novelty, control duplication, or assign a supervisor?
- RQ3: What are the limitations and unresolved problems in the existing research?

The rest of the article is organized as follows. The next section describes the methodology of systematic review and bibliometric analysis. The following sections discuss approaches based on artificial intelligence to ensure academic integrity, identify the novelty of research, and assign a supervisor. The final section discusses the implications of the findings and defines the directions for future research.

Methods. In this research, the mixed-methodological design is used, which combines a systematic review with a bibliometric analysis to integrate thematic and structural features of research on AI-based academic project management. The systematic component was reported in accordance with PRISMA 2020 guidelines to ensure transparency, replicability, and methodological rigor in study identification, screening, and inclusion procedures. At the same time, bibliometric mapping was conducted to analyze the quantitative trends in the field of research.

Search Strategy

A comprehensive literature search was conducted in Scopus and Google Scholar, selected for their broad coverage of computer science, artificial intelligence, and educational research. The final search was completed in December 2025.

The search strategy was structured around several conceptual categories reflecting the core dimensions of the review: (1) AI and semantic technologies, (2) academic integrity and duplication control, (3) research topic novelty assessment, (4) supervisor assignment and recommendation systems, and (5) the higher education academic project management context.

Within each category, the related terms were added together using the Boolean operator OR to represent the variations of the terms. The conceptual categories were then combined with operator AND to find studies that were located at the intersection of AI-driven semantic processes and academic project management processes in higher education. Where relevant, truncation symbols have been used in order to take into consideration the lexical variation.

Only peer-reviewed publications in English published between 2021 and 2025 were considered. Table 1 provides the detailed structure of the search terms and categorical grouping of search terms.

Table 1 – Search strategy

Category	Included Key Terms
AI and Semantic Technologies	("artificial intelligence" OR AI OR "machine learning" OR "deep learning" OR "natural language processing" OR NLP OR "transformer model*" OR "large language model*" OR LLM* OR semantic analys* OR semantic similarity OR semantic matching OR "text embedding*" OR "sentence embedding*" OR SBERT OR BERT OR "contextual embedding*" OR "deep semantic analys*")
Academic Integrity and Duplication Control	("academic integrity" OR plagiarism OR "semantic plagiarism" OR "intelligent plagiarism" OR paraphras* OR "text reuse" OR "similarity detection")
Research Topic Novelty Assessment	("research novelty" OR "topic novelty" OR originality OR "novelty detection" OR "originality assessment" OR "research originality" OR "innovation detection")
Supervisor Assignment and Recommendation Systems	("supervisor assignment" OR "supervisor matching" OR "supervisor allocation" OR "recommender system" OR "matching algorithm")
Higher Education and Academic Project Management Context	("higher education" OR universit* OR "academic project*" OR thesis OR dissertation OR "master* thesis" OR "diploma project*" OR "research management" OR "project management")

Inclusion and Exclusion Criteria

Inclusion criteria comprised peer-reviewed journal articles and review papers addressing AI or NLP applications in higher education contexts, particularly in relation to semantic similarity, novelty detection, plagiarism detection, text processing, or resource allocation within the academic project lifecycle. Exclusion criteria included non-peer-reviewed sources (e.g., editorials, blogs), purely technical AI studies without educational context, and articles lacking sufficient methodological transparency.

Quality Assessment and Data Extraction

The entire full-text assessment process included methodological selection to assess the clarity of the research objectives, the clarity of the data sets, the suitability of the assessment design, and sufficient detail to reproduce the study. Research papers that were not transparent in terms of methods were rejected. The systematic data analysis included author data, year of publication, aim of the study, methodology, participants data, main findings, limitations and keywords.

Study Selection Process

The dataset included 49 peer-reviewed articles published between 2021 and 2025. To attain a clear and reproducible means of selection, the literature on the topic was filtered with regard to the PRISMA 2020 protocol. In the first step of identifying relevant articles, a total of 1,829 articles were obtained from an extensive search of online databases such as Scopus and Google Scholar. After duplicate removal, a total of 418 duplicates were identified and eliminated, resulting in a preliminary list of 1,411 articles for further verification.

At the stage of selecting titles and abstracts, 1,176 entries were screened out because they did not correspond to the key areas of the research objectives. This resulted in 235 records being sought for full-text assessment. Of these, 223 full texts were successfully retrieved, whereas 12 records were excluded due to lack of access. The compliance of these 223 records with the requirements was carefully checked in accordance with a number of inclusion and exclusion criteria. At this stage, 174 reports were excluded for specific methodological and contextual reasons. The main grounds for exclusion included the absence of an AI-based methodological component (71 records), lack of focus on higher education settings (48 records), insufficient relevance to the core research themes of plagiarism detection, novelty evaluation, and supervisor assignment or recommendation (54 records), and inadequate methodological transparency preventing comparative analysis or reproducibility (1 record). Finally, 49 articles were peer-reviewed and included in the final systematic review process.

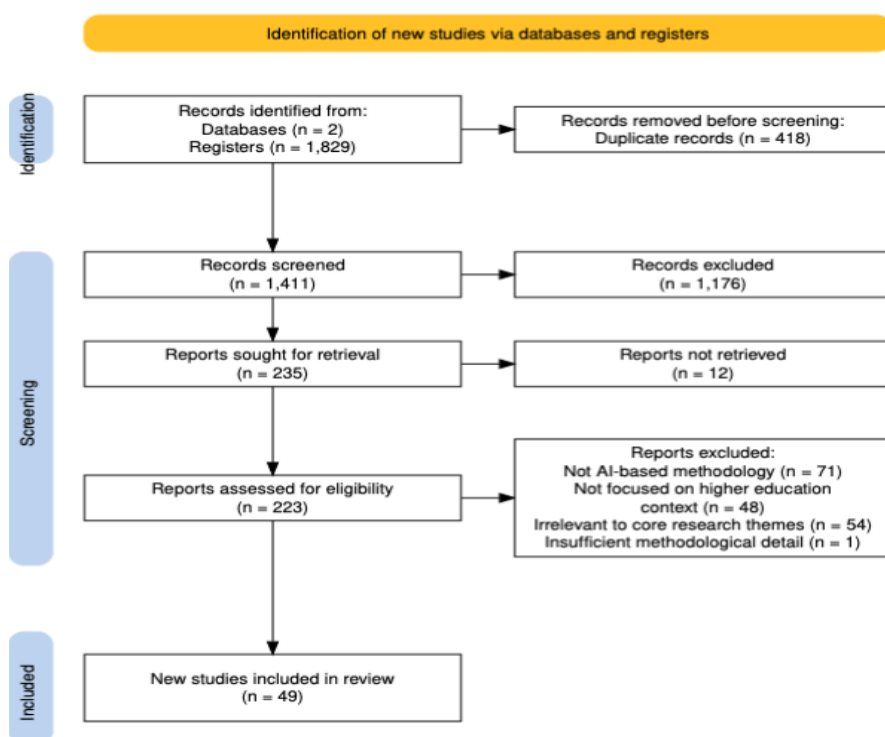


Figure 1 – PRISMA flowchart of the study selection process

Bibliometric and Thematic Analysis

Bibliometric analysis was conducted using VOSviewer to examine keyword co-occurrence patterns, co-authorship networks, and co-citation structures within the selected corpus. In parallel, a qualitative thematic synthesis was performed using an inductive coding approach. The included studies were iteratively coded and

grouped based on their methodological characteristics and functional application domains, resulting in four dominant thematic clusters.

Results. Results of the Bibliometric Analysis

Annual Publication Trends

Figure 2 illustrates the annual distribution of publications included in the review between 2021 and 2025. The number of studies demonstrates a consistent upward trend over the examined period. After a relatively moderate volume of publications in 2021, the output gradually increased in 2022 and 2023, followed by a more pronounced growth in 2024. The highest number of publications was recorded in 2025. Overall, the data indicate a growing scholarly interest in AI-driven academic project management and NLP-based administrative systems in higher education.

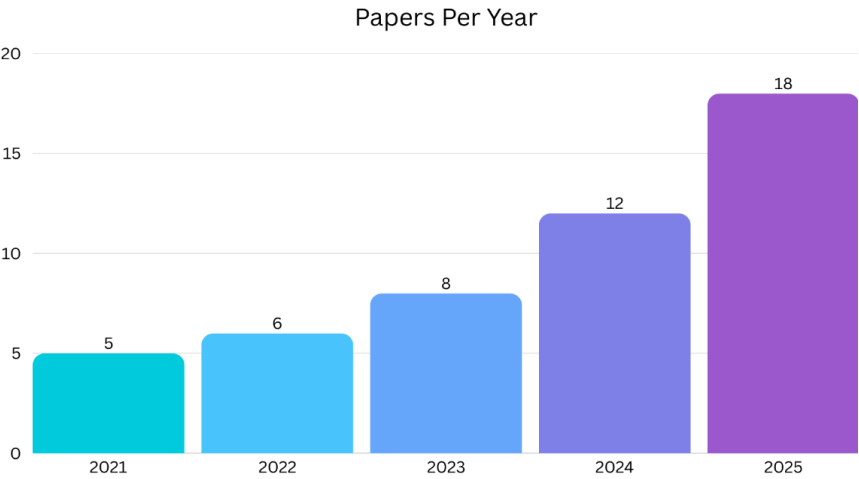


Figure 2 – Number of publications per year of inclusion (2021-2025)

Methodological Paradigms

In the given section, the quantitative comparison of the NLP similarity core methods used in the reviewed literature i.e. classical lexical methods, embedding-based semantic models, and LLM-based approaches is given. The 49 included studies were all analyzed and categorized based on the main analytical framework of the study. Since several studies employed hybrid architectures (e.g., combining TF-IDF baselines with embedding models or LLM components), the “primary” framework was determined based on the dominant method driving the final model performance or the central methodological contribution described by the authors. Classification in situations of hybrid implementations was done according to the predominant analytical paradigm.

According to the methodological descriptions of the respective articles, three major approaches became evident, including classical lexical and traditional machine learning techniques, transformer-based embedding models, and systems powered by large language models.

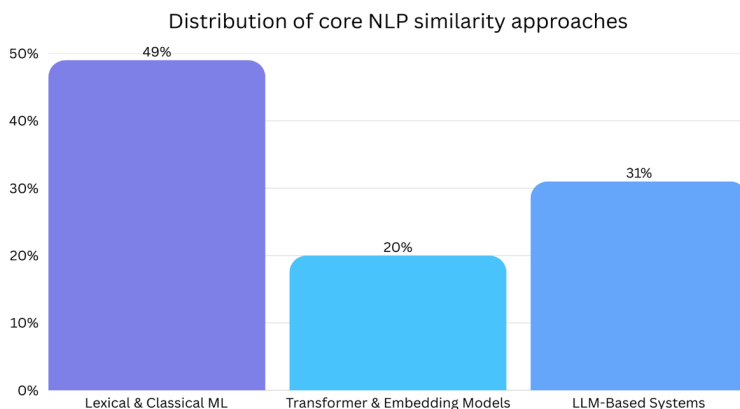


Figure 3 – Distribution of primary NLP similarity paradigms

The findings show that the studies based on lexical and classical machine learning capture 49% ($n = 24$). Such methods are TF-IDF, cosine similarity, supervised classification models, neural networks with no large-scale pre-trained models, and systems based on optimization. In most instances, the techniques are utilized as a form of baseline similarity or structured decision-supporting in recommendation and supervisor assignment models.

In the reviewed literature, 20% ($n = 10$) is transformer-based embedding. Such studies use mostly contextual sentence or document representations that were built based on architectures like BERT and SBERT. Similarity is commonly computed with dense vector embeddings and ranking or similarity measures and is generally used in semantic similarity measurement, paraphrase detection, and novelty measurement.

The approaches based on LLM represent 31% of the studies ($n = 15$). These are GPT-based, prompt-based interaction and zero-shot or few-shot inference strategies. These applications have been commonly used in AI-generated text detection and conceptual originality assessment tasks.

Overall, Figure 3 shows that lexical, embedding-based, and LLM-driven paradigms have a relatively high prevalence in the reviewed research body, which can be interpreted as a diversification of methodological approaches in current academic applications of NLP.

Conceptual Structure of the Research Field

The keywords co-occurrence network provides four significant thematic groups suggesting the conceptual framework of the field. Important nodes like artificial intelligence and natural language processing also act as connector concepts between multiple clusters.

The former cluster unites embedding-based and deep learning methods (e.g., LSTM, classification, text processing), which represent an expression of methodological improvement in automated text analysis.

The second cluster represents the domain of computational linguistics and semantically oriented studies, including research on computational linguistics, semantics, and novelty detection, with a focus on the theoretical foundations of language technologies.

The third cluster revolves around large language models and AI-based educational applications, such as language models, ChatGPT, machine learning, and students, which is a pointer to the applied aspect of generative AI in education.

The fourth cluster brings together semantic analysis, plagiarism detection and human-related evaluation, which display issues to do with assessment, integrity, and human-centered interpretation.

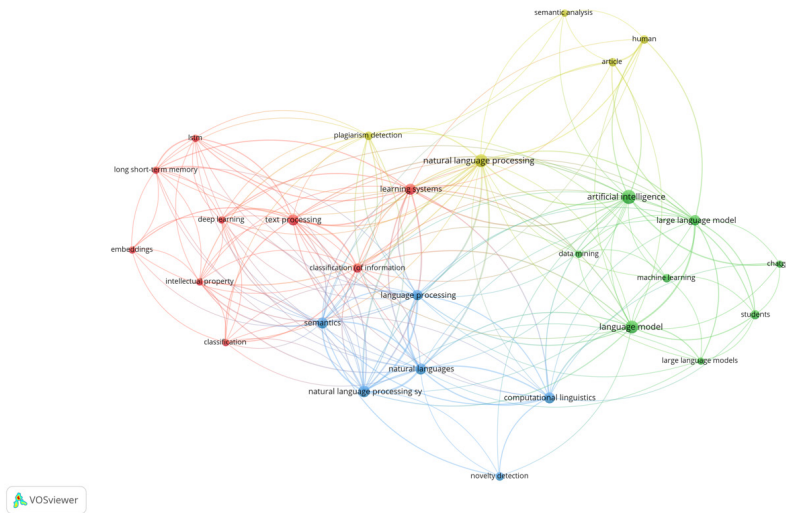


Figure 4 – Keyword co-occurrence network highlighting the shift toward generative AI and LLMs

Thematic Evolution of Author Keywords

In the overlay visualization shown in Figure 5, the co-occurrence of author keywords is presented over time in the analyzed literature, with natural language processing as the key term. Keywords that featured in the earlier publications predominantly 2022-early 2023 include deep learning, recurrent neural network, text plagiarism, and classification. Keywords that feature in the studies published between 2024-2025 include large language model, semantic analysis, academic writing, and AI innovation management.

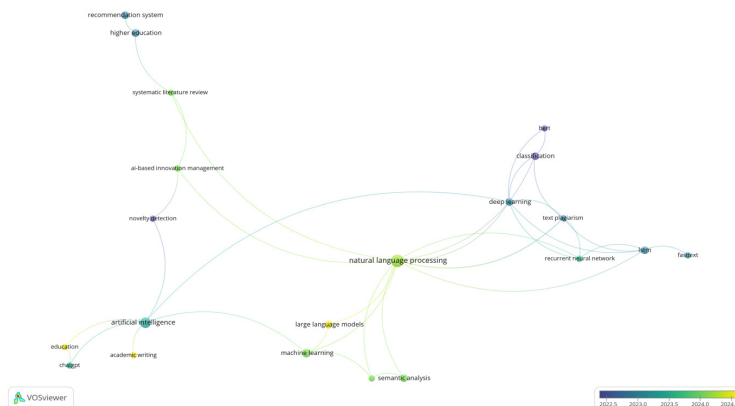


Figure 5 – Overlay-based thematic evolution of author keywords in AI-driven academic management research

Intellectual Sources and Publication Outlets

The source co-citation network shows six clusters of different intellectual foundations of the field. The core AI and NLP venues cluster (that is, the largest cluster by the number of sources) is centered around the Advances in Neural Information Processing Systems, ACL/EMNLP/NAACL proceedings, and ICML, showing the backbone of methodology in computational linguistics and deep learning. Conversely, the most significant hubs by cumulative link strength are seen in scientometrics and innovation-policy journals, including Scientometrics, Technological Forecasting and Social Change, and World Patent Information, so there is a strong evaluation-oriented research area. A second cluster will combine education-related journals, such as Computers and Education: Artificial Intelligence, British Journal of Educational Technology, and International Journal of Educational Technology in Higher Education. arXiv would be an exceptionally linked bridging source between conference and journal outlets.

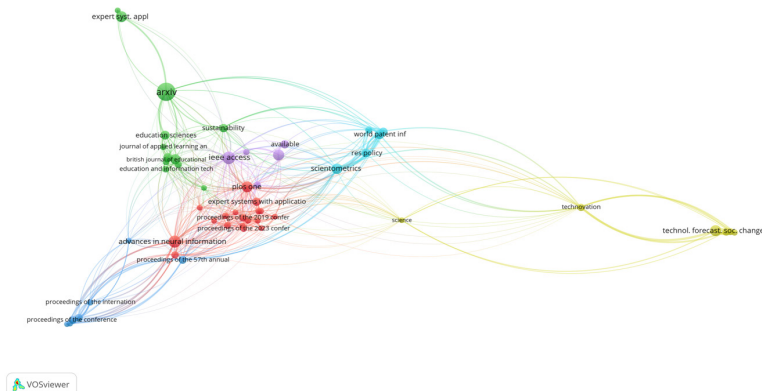


Figure 6 – Source co-citation network of publication outlets in AI-driven academic management research

Collaboration Patterns in AI Research

Figure 7 presents the co-authorship network among authors included in the review. The visualization reveals a fragmented structure consisting of several small groups with few connections among them, indicating relatively weak collaboration within the field.

Several clusters demonstrate strong internal connectivity indicating that specific research teams work together in the same manner. Meanwhile, some of the authors are shown as peripheral nodes with few co-authorship relationships revealing detached contributions to the field.

Thematic orientation of such clusters could be aligned with various research directions but the main outcome of the co-authorship network is a collaboration pattern as opposed to a thematic structure.

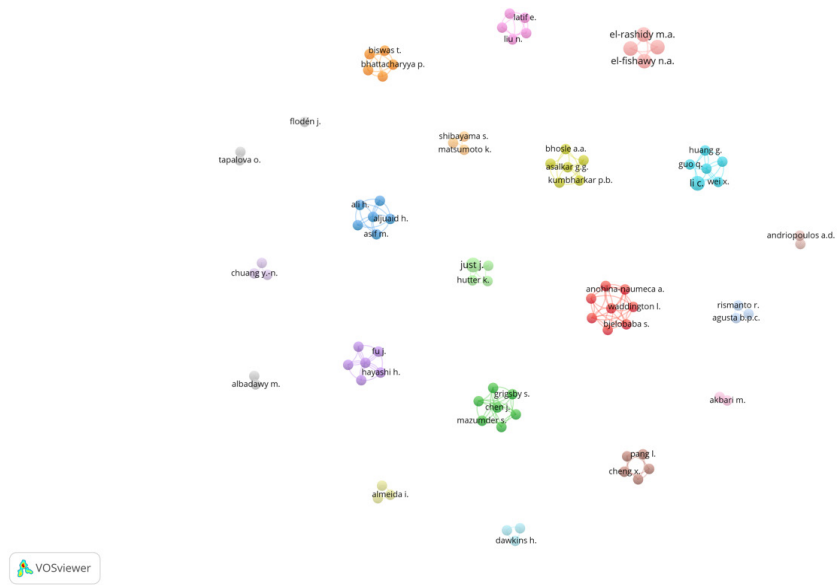


Figure 7 – Co-authorship network illustrating collaboration patterns among core authors

Intellectual Influences and Core Authors

The author co-citation network is organized into three broad clusters. The first cluster represents the core methodological foundations of Neural Natural Language Processing and Representation Learning, including highly co-cited authors such as Lee, K., Wang, Y., Chen, K., and Manning, C.D., which exhibit strong co-citation link strength.

The second cluster is centered on plagiarism detection, authorship analysis, and text forensics. Rosso appears as the central node in this group, with particularly strong co-citation ties observed between Potthast and Stein. Structurally, this cluster is clearly differentiated from the neural NLP cluster, with relatively limited connections between them.

A third, smaller cluster occupies an intermediate position between the neural NLP core and the plagiarism cluster, suggesting a bridging role within the intellectual structure of the field. The composition and spatial positioning of these clusters are illustrated in Figure 8.

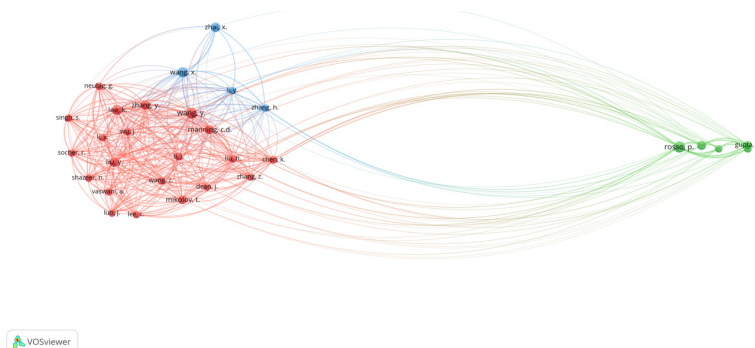


Figure 8 – Author co-citation network revealing the intellectual foundations of generative AI

Thematic Synthesis (Systematic Review)

Following the content analysis of the literature reviewed, four thematic directions emerged:

1. *Semantic integrity and plagiarism detection.*

This thematic area covers research issues related to the analysis of text similarity, plagiarism and paraphrasing, especially how they are applied in the academic environment (Amirzhanov et al., 2025; Darwish et al., 2023; Hryn et al., 2025; Sajid et al., 2025). The literature on this topic area included the use of classical algorithms to calculate the similarity between lexical texts, such as TF-IDF or cosine similarity, and the use of the latest transformer-based similarity models based on contextual embeddings to identify deep semantic connections (Amirzhanov et al., 2025; Hassan et al., 2025).

2. *Research topic novelty and originality assessment.*

Semantic novelty detection and redundancy identification studies are related to this thematic direction and define novelty as a change in topical context, in relation to an already existing or normal textual context (Ghosal et al., 2022; Just et al., 2024; Mat et al., 2021). The reviewed studies describe various applications of document embedding, contextual similarity and topic modeling techniques, where novelty is proved by semantic ranking or prioritized in comparison to a bibliographic context. Certain applications of the large language models in terms of assessing novelty or comparing ideas were observed (Wu et al., 2025).

3. *Intelligent recommendation and allocation systems.*

Supervisor allocation and recommendation in higher education constitute a distinct thematic direction within the reviewed literature, as they belong to this thematic direction (Anghel and Popescu, 2022; Dasri et al., 2025; Ramotsisi et al., 2022). Annisa et al. (2024) introduce a two-way recommendation system involving

the use of historical data of alumni and k-skyband view queries to make a broader span of recommendations than skyline objects, whereas Dasri et al. (2025) introduce a MapReduce-based implementation suitable to large academic datasets.

4. Human-oriented management and decision-making.

Research papers on the same topic that address transparency, explainability, and human interaction scenario within academic management systems through artificial intelligence fall within this thematic area (Khalifa and Albadawy, 2024; Ramotsisi et al., 2022; Shimray and Subaveerapandiyan, 2025). In parallel, methodological NLP research highlights the need for traceable and human-interpretable reasoning pipelines to support explainability (Fox et al., 2025).

Across all thematic areas, the reviewed studies report the use of embedding-based semantic representation, large language model-based methods, and lexical similarity-based approaches, as illustrated by various reports in the reviewed literature. The reported use of embedding-based and large language model-based methods was found in literature published in later years of the review period, while lexical similarity-based methods were illustrated as resource-efficient approaches.

Based on the reviewed literature, the data sources in this study, also indicate academic text data, including theses, research projects, and publications, as well as bibliographic, citation, and administrative data. Evaluation results based on real-world institutional deployments or long-term evaluations are rarely reported (Hryn et al., 2025).

The thematic distribution of the reviewed studies regarding semantic similarity, the evaluation of originality, recommendation systems, and the examined domains of governance is summarized in the next subsections.

Identified Research Gaps

An examination of the literature reveals several recurring gaps in the topics regarding the evaluation practices, data sources, methodological consistency, system integration in AI-based management research.

1. Real-world institutional deployment and longitudinal evaluation.

The reviewed studies report limited evidence of real-world institutional deployment or long-term evaluation. Many articles are based on autonomous experiments, environmental modeling, or retrospective analysis, with little or no evaluation of system performance in various academic terms (Ramotsisi et al., 2022). Only a small number of studies provide estimates based on real-world educational conditions and institutional project repositories (Luis et al., 2025).

2. Use of primary institutional and administrative data.

The reviewed literature indicates that frequent data sources include secondary data sources such as bibliographic databases, publication databases, and archival texts. Information related to academic texts is often described, while information related to institutional administration datasets is less frequently reported. Furthermore, datasets combining administrative data, texts, and evaluative data using a singular data analysis framework are rarely described.

3. Management, transparency, and explainability mechanisms.

The reviewed studies lack detailed information on interpretability, verifiability, and expert confirmation. This is consistent with broader NLP findings, according to which modern similarity pipelines usually do not guarantee explainability and the need to provide traceable hybrid reasoning procedures (Fox et al., 2025).

4. Evaluation protocols and benchmarking practices.

From the evaluated literature, numerous heterogeneous assessment schemes, performance criteria, and experimental designs are presented. Moreover, in the literature, various dataset creation methods, similarity boundaries, and assessment methods are recognized, yet the achievement of standard benchmarks is less emphasized.

5. Integration of multiple academic management functions.

The majority of reviewed works reported systems for individual functions like plagiarism detection, checking for novelty, or recommendation. Accounts of integrated system architectures for all or several academic management functions within a single system remain relatively scarce (Hryn et al., 2025; Ramotsisi et al., 2022; Kashkimbayeva et al., 2025).

Discussion.

This section interprets the findings that have been made and places these findings within the context of the wider body of research on AI-driven academic project management. Rather than reiterating the results presented in the systematic review and bibliometric analysis, the discussion focuses on explaining why the identified thematic structures, adopted methods, and observed collaboration patterns are relevant. The discussion is intended to align directly with the findings that have been made.

Key Trends Interpretation and Comparison with Previous Literature.

Considering the quantitative distribution of methodology paradigms, the development of AI-based academic project management can be discussed as evolutionary rather than disruptive. Even though classical lexical and traditional methods of machine learning are still a major part (49%), methods based on embedding and LLM are a narrow majority of recent works (51%). Such a balance implies the progressive change towards the semantically-based decision-support systems instead of replacing previous methods entirely.

The growing popularity of embedding and transformer-based models is an indicator of an appreciation of the weaknesses of superficial lexical similarity to identify paraphrased or semantically manipulated academic content (Amirzhanov et al., 2025; Darwish et al., 2023). In this regard, is increasingly conceptualized as conceptual positioning within an evolving research space, extending previous theoretical reflections (Ghosal et al., 2022).

Meanwhile, the emergence of the approaches based on LLM indicates growing interest in the abstraction-based reasoning and discourse-level assessment (Wu

et al., 2025). However, their coexistence with lexical and embedding approaches imply methodological diversification as opposed to paradigm supremacy. All these findings combine to present a growing yet structurally plural field of AI uses in academic integrity, originality assessment, and supervisor assignment.

Theoretical and Practical Contributions

Theoretically, the review is valuable for the literature because it combines both bibliometric analysis and systematic review to analyze how research streams are organized in relation to each other within a single methodological concept. The analysis of joint citations and keywords can also be used to determine that the current trend of research in the field of intellectual academic management is more related to progress in the field of basic NLP architectures than to the unification of theoretical traditions.

From a practical perspective, the results provide evidence-based guidelines that can be used in the design of more intelligent academic management systems. The reviewed literature shows that more effective systems are now being designed based on hybrid models that include semantic similarity modeling and rule systems or optimization systems to facilitate decision-making in academic institutions (Ramotsisi et al., 2022; Nurgaliyeva et al., 2025). Nevertheless, the low level of functional domain integration implies the need for unified systems that incorporate tasks like integrity, originality, and allocation in an integrated fashion. In addition to developing and harmonizing the results of previous research, the current study has the additional advantage that offers a conceptual model for developing an integrated academic project management system based on artificial intelligence, combining the main functional aspects established in the current review of scientific literature (Figure 9).

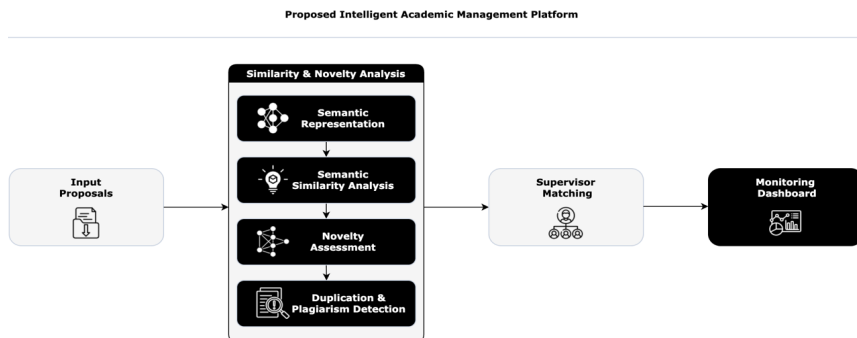


Figure 9 – Conceptual framework of an AI-based academic project management platform

In the proposed concept, semantic representation and similarity analysis constitute increasingly central structural components that serve as the basis for the implementation of various academic management operations, such as duplication control, novelty assessment, and resource allocation control. By illustrating how these components can be interconnected, this framework eliminates the

fragmentation of previous studies, in which systems for evaluating integrity, originality, and recommendations were considered largely as independent processes. Consequently, this framework provides a systematic view of how semantically sound artificial intelligence technologies can be used to facilitate the lifecycle of an academic project as a whole.

Methodological Observations and Limitations

There are certain trends in assessment methodology that need to be considered: first, there are still significant differences in assessment methods, with different datasets, similarity measures, and different assessment metrics, making it difficult to compare methods directly. Secondly, there is also a high level of use of secondary or bibliographic data with fewer reports on the use of administrative data, which partially affects the ecological validity of the assessment. Thirdly, the results of a real assessment are presented only in a limited number of studies. For example, Luis et al. (2025) evaluate the Augmented Search Generation System (RAG) to evaluate originality according to the instructors' judgments, while many other proposed systems remain at the conceptual or experimental stage.

These limitations indicate that the existing evidence is largely based on technical capabilities rather than institutional ones, and it is recommended that the former be taken into account when calculating accuracy or efficiency indicators.

Directions for Future Research

In light of the gaps mentioned above, the upcoming research should aim at empirical validation of an academic management system based on artificial intelligence, especially with regard to its effectiveness in the long term, across multiple academic cycles. Similar and convincing evidence will also be easier to collect through the introduction of standardized assessment procedures. In addition, additional research can be conducted on the architecture of the integrated system, which is a combination of a semantic integrity control scheme, a novelty assessment scheme, and a recommendation scheme.

Finally, research should be continued beyond the English-speaking space in order to generalize and make more practical academic management systems based on artificial intelligence, especially given that recent empirical evidence has shown widespread use of artificial intelligence tools in academic literature and other issues such as reliability, cost, and ethical management (Shimray and Subaveerapandiyana, 2025), as well as the growing complexity of new forms of plagiarism, such as paraphrased, interlanguage plagiarism and content developed by artificial intelligence (Amirzhanov et al., 2025).

Conclusion. The review is a summary of the literature on artificial intelligence in the field of academic project management in higher education, representing a comprehensive systematic analysis and bibliometric review. By consolidating evidence across semantic integrity control, research novelty assessment, and supervisor allocation systems, the study highlights a gradual shift from purely

lexical approaches toward embedding-based and large language model-driven paradigms. Rather than operating as isolated functions, these domains increasingly converge within semantically informed decision-support frameworks.

The results demonstrate a methodologically plural but structurally fragmented research landscape, characterized by heterogeneous evaluation protocols, limited institutional deployment, and scarce longitudinal validation. Meanwhile, the increasing use of transformer architecture and LLM-based logic points to the growth of higher-level abstraction and governance-oriented AI applications in academic field.

Future research should focus on empirical verification in real-world institutional environments, standardization of benchmarking processes, and modeling the architecture of an integrated system that combines integrity, originality, and allocation processes. With the continued development of artificial intelligence technologies, their application in academic project management will only expand, which requires complete technological progress, as well as clear, understandable and ethically justified implementation practices.

References

- Anghel F., and Popescu E. (2022) Dynamic Graduation Project Allocation Based on Student-Teacher Profile Compatibility. *International Conference on Advanced Learning Technologies*. — P. 56–60. <https://doi.org/10.1109/ICALT55010.2022.00025> (in Eng.)
- Annisa A., Adyatma M.R., Sampurno G I., and Li C. (2024) Two-way recommendation system for supervisor selection using historical data and skyband-view queries. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 34(2). — P. 305–314. <https://doi.org/10.37934/araset.34.2.305314> (in Eng.)
- Amirzhanov A., Turan C., and Makhmutova A. (2025) Plagiarism types and detection methods: a systematic survey of algorithms in text analysis. *Frontiers in Computer Science*, 7. <https://doi.org/10.3389/fcomp.2025.1504725> (in Eng.)
- Bekish U., Nurgabyl D., Tulegulov A., Abdoldinova G., and Basheyeva Z. (2020) Methodology of designing the pithy components of elective courses on the higher mathematics of pedagogical profile. *International Journal of Engineering Research and Technology*, 13(3). — 407 p. <https://doi.org/10.37624/ijert/13.3.2020.407-413> (in Eng.)
- Bibi S., Trinanda K., Syamsudin M., Suheri Fitri N.A., Suharsono, Rigaux M., and Guarim R. (2023) *A Neural Network Recommendation Model for Final Project Topic Selection*. — P. 407–410. <https://doi.org/10.1109/ice3is59323.2023.10335346> (in Eng.)
- Dasri D., Annisa A., and Haryanto T. (2025) Two-way thesis supervisor recommendation system using MapReduce k-skyband view queries. *JOIV International Journal on Informatics Visualization*, 9(1). — 163 p. <https://doi.org/10.62527/joiv.9.1.2800> (in Eng.)
- Darwish S. M., Mhaimeed I. A., and Elzoghbi A.A. (2023) A quantum genetic algorithm for building a semantic textual similarity estimation framework for plagiarism detection applications. *Entropy*, 25(9). — 1271 p. <https://doi.org/10.3390/e25091271> (in Eng.)
- Fox O.R., Bergami G., and Morgan G. (2025) Verified language processing with hybrid explainability. *Electronics*, 14(17). — 3490 p. <https://doi.org/10.3390/electronics14173490> (in Eng.)
- Ghosal T., Saikh T., Biswas T., Ekbal A., and Bhattacharyya P. (2022) Novelty detection: A perspective from natural language processing. *Computational Linguistics*, 48(1). — P. 77–117. https://doi.org/10.1162/coli_a_00429 (in Eng.)
- Hassan E., Talaat A.S., and Elsabagh M.A. (2025) Intelligent text similarity assessment using RoBERTa with integrated chaotic perturbation optimization techniques. *Journal of Big Data*, 12(1). <https://doi.org/10.1186/s40537-025-01233-3> (in Eng.)

Hryn O., Shevel A., Shcherbyna N., Kubrak, O., and Zadorozhnyi K. (2025) Implementation of artificial intelligence in the system for detecting academic dishonesty in Ukrainian secondary and higher education institutions. *Periodicals of Engineering and Natural Sciences (PEN)*, 13(2). — P. 445–458. <https://doi.org/10.21533/pen.v13.i2.467> (in Eng.)

Just J., Ströhle T., Füller J., and Hutter K. (2024) AI-based novelty detection in crowdsourced idea spaces. *Innovation*, 26(3), 359–386. <https://doi.org/10.1080/14479338.2023.2215740> (in Eng.)

Kashkimbayeva N., Bekish U., Abdoldinova G., Basheyeva Z., and Mitroshina A. (2025) Evaluating an integrated user-feedback approach to software quality monitoring: enhancing accuracy and timeliness. *Eastern-European Journal of Enterprise Technologies*, 5(2 (137)). — P. 29–42. <https://doi.org/10.15587/1729-4061.2025.341516> (in Eng.)

Khalifa M., and Albadawy M. (2024) Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, 5. — 100145 p. <https://doi.org/10.1016/j.cmpbup.2024.100145> (in Eng.)

Korade N. B., Salunke M. B., Bhosle A.A., Kumbharkar P.B., Asalkar G.G., and Khedkar R.G. (2024) Strengthening sentence similarity identification through OpenAI embeddings and deep learning. *International Journal of Advanced Computer Science and Applications*, 15(4). <https://doi.org/10.14569/ijacsa.2024.0150485> (in Eng.)

Luis S.Y., Reina D.G., and Marín S.T. (2025) Towards a retrieval-augmented generation framework for originality evaluation in projects-based learning classrooms. *Education Sciences*, 15(6). — 706 p. <https://doi.org/10.3390/educsci15060706> (in Eng.)

Mat N., Politowicz A., Mazumder S., Chen J., Liu B., Robertson E., and Grigsby S. (2021) Semantic novelty detection in natural language descriptions. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. — P. 866–882. <https://doi.org/10.18653/v1/2021.emnlp-main.66> (in Eng.)

Nurgaliyeva S., Amangali M., Basheyeva Z., Kashkimbayeva N., and Amantayev D. (2025) Design and evaluation of an intelligent waste monitoring system based on RGIS integration for smart cities. *Eastern-European Journal of Enterprise Technologies*, 4(9). — P. 70–78. <https://doi.org/10.15587/1729-4061.2025.337033> (in Eng.)

Ramotšisi J., Kgomošoto M., and Seboni L. (2022) An optimization model for the student-to-project supervisor assignment problem: The case of an engineering department. *Journal of Optimization*, 2022. — P. 1–11. <https://doi.org/10.1155/2022/9415210> (in Eng.)

Sajid M., Sanaullah M., Fuzail M., Malik T. S., and Shuhidan S.M. (2025) Comparative analysis of text-based plagiarism detection techniques. *PLoS ONE*, 20(4), e0319551. <https://doi.org/10.1371/journal.pone.0319551> (in Eng.)

Shibayama S., Yin D., and Matsumoto K. (2021) Measuring novelty in science with word embedding. *PLoS ONE*, 16(7), e0254034. <https://doi.org/10.1371/journal.pone.0254034> (in Eng.)

Shimray S.R., and Subaveerapandiyar, A. (2025). Artificial intelligence in academic writing and research: Adoption and effectiveness. *Open Information Science*, 9(1). <https://doi.org/10.1515/opis-2025-0026> (in Eng.)

Wu W., Zhang C., and Zhao Y. (2025) Automated novelty evaluation of academic paper: A collaborative approach integrating human and large language model knowledge. *Journal of the Association for Information Science and Technology*, 76(11). — P. 1452–1469. <https://doi.org/10.1002/asi.70005> (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 125–145

<https://doi.org/10.32014/2026.2518-1726.430>

IRSTI 06.81.23

UDC 004.8

© **Bekmanova G. *, Kantureyeva M., Omarbekova A., Zakirova A.,
Issainova A., 2026.**

Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan.

E-mail: kantureyeva_ma@enu.kz

INTEGRATING ARTIFICIAL INTELLIGENCE TO EVALUATE EMOTIONS IN THE LEARNING ENVIRONMENT

Bekmanova Gulmira — Vice-Rector for Digitalization - Digital Officer, PhD, Associate Professor of the L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: bekmanova_gt@enu.kz, <https://orcid.org/0000-0001-8554-7627>;

Omarbekova Assel — Head of Digital Development and Distance Learning Department, PhD, Associate Professor of the L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: omarbekova_as@enu.kz, <https://orcid.org/0000-0002-9272-8829>;

Kantureyeva Mansiya — PhD, Senior Lecturer, Department of Information Systems, L.N. Gumilyov Eurasian National University, Faculty of Information Technology, Astana, Kazakhstan,

E-mail: kantureyeva_ma@enu.kz, <https://orcid.org/0000-0001-5904-820X>;

Zakirova Alma — Director of the Department of Science, candidate of Pedagogical Sciences, Associate Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: zakirova_ab@enu.kz, <https://orcid.org/0000-0001-8772-1414>.

Issainova Aliya — Senior Lecturer of the Department of Information Security at the L.N. Gumilyov Eurasian National University, Institute of Digital Sciences and Artificial Intelligence, Astana, Kazakhstan,

E-mail: issainova_an@enu.kz, <https://orcid.org/0000-0001-8008-4096>.

Abstract. The digitalization of modern education is accompanied by a rapid increase in the volume of multimodal data formed during the educational process. Video recordings of the lesson, audio excerpts, textual interactions in learning control systems (LMS), as well as telemetry of user activity reflect not only the cognitive, but also the emotional component of learning. The emotions of students are an important indicator of their participation, interest and effectiveness in mastering the material. Correct recognition of emotions allows you to dynamically adapt the educational content to the emotional reaction of the audience, timely identify signs of stress or a decrease in motivation, as a result of which students retain interest in the lesson and improve academic performance. In this paper, the scientifically substantiated and practically applicable models of the university, as well as developed, are aimed not at an individual assessment of the individual,

but at a consolidated analysis of the emotional dynamics of study groups. *Results:* What we see in the results has classified and differentiated the most important aspects of the use of artificial intelligence in the assessment of emotions in research, and, more clearly, in the educational environment. Specific data effects the impact on the educational process was considered in detail. *Scientific novelty:* Based on specific artificial intelligence technologies used in the considered research. It includes a detailed analysis of the methods, tools, algorithms and technological approaches used to assess emotions in the educational environment. *Practical value:* Preparation of methodological recommendations for the further transition to multimodal models using additional data (audio, text, sensory indicators).

Keywords: emotions, digitalization, artificial intelligence, intellectual, educational environment, learners, inclusive

Funding. *This article is funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant №BR28713531 Intelligent digital system of higher and postgraduate education organizations SmartEDU).*

For citations: Bekmanova G., Kantureyeva M., Omarbekova A., Zakirova A., Issainova A. Integrating artificial intelligence to evaluate emotions in the learning environment. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 125-145. DOI <https://doi.org/10.32014/2026.2518-1726.430>

© Бекманова Г.Т.*, Кантуреева М.А., Омарбекова А.С.,
Закирова А.Б., Исайнова А.Н., 2026.

Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан.

E-mail: kantureyeva_ma@enu.kz

ОҚУ ОРТАСЫНДАҒЫ ЭМОЦИЯЛАРДЫ БАҒАЛАУ ҮШІН ЖАСАНДЫ ИНТЕЛЛЕКТТІ БІРІКТІРУ

Бекманова Гүлмира — т.ғ.к, PhD, қауымдастырылған профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Басқарма мүшесі – Цифрландыру жөніндегі проректор-Цифрлық офицер, Астана, Қазақстан,

E-mail: bekmanova_gt@enu.kz, <https://orcid.org/0000-0001-8554-7627>;

Омарбекова Асель — т.ғ.к, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Цифрлық даму және онлайн оқыту департамент директоры, қауымдастырылған профессор, Астана, Қазақстан,

E-mail: omarbekova_as@enu.kz, <https://orcid.org/0000-0002-9272-8829>;

Кантуреева Мансия — PhD, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Ақпараттық жүйелер кафедрасының аға оқытушысы., Ақпараттық технологиялар факультеті, Астана, Қазақстан,

E-mail: mansiya_arynbekovna@mail.ru, <https://orcid.org/0000-0001-5904-820X>;

Закирова Алма — Ғылым департаментінің директоры, педагогика ғылымының кандидаты, қауымдастырылған профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: zakirova_ab@enu.kz, <https://orcid.org/0000-0001-8772-1414>;

Исайнова Әлия — ақпараттық қауіпсіздік кафедрасының аға оқытушысы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Цифрлық ғылымдар және жасанды интеллект институты, Астана, Қазақстан,

E-mail: issainova_an@enu.kz, <https://orcid.org/0000-0001-8008-4096>.

Аннотация. Қазіргі білім беруді цифрландыру оқу процесінде қалыптасқан мультимодальды мәліметтер көлемінің тез өсуімен қатар жүреді. Сабақтың бейнежазбалары, аудио үзінділер, оқытуды басқару жүйелеріндегі (LMS) мәтіндік өзара әрекеттесулер, сондай-ақ пайдаланушы әрекетінің телеметриясы оқытудың когнитивті ғана емес, эмоционалдық құрамдас бөлігін де көрсетеді. Оқушылардың эмоциялары олардың материалды меңгеруге қатысуының, қызығушылығының және тиімділігінің маңызды көрсеткіші болып табылады. Эмоцияларды дұрыс тану білім беру мазмұнын аудиторияның эмоционалды реакциясына динамикалық түрде бейімдеуге, күйзеліс белгілерін немесе мотивацияның төмендеуін уақтылы анықтауға мүмкіндік береді, нәтижесінде оқушылардың сабаққа деген қызығушылығы сақталады және оқу үлгерімі жақсарады. Бұл жұмыста университеттің ғылыми негізделген және іс жүзінде қолданылатын модельдері, сонымен қатар әзірленген, жеке тұлғаны жеке бағалауға емес, оқу топтарының эмоционалды динамикасын шоғырландырылған талдауға бағытталған. Нәтижелер: нәтижелерден көріп отырғанымыз жасанды интеллектті зерттеудегі эмоцияларды бағалауда және нақтырақ айтқанда, білім беру ортасында қолданудың маңызды аспектілері жіктеліп, сараланды. Нақты деректер әсерлері оқу процесіне әсері егжей-тегжейлі қарастырылды. Ғылыми жаңалығы: қарастырылып отырған зерттеулерде қолданылатын нақты жасанды интеллект технологияларына негізделді. Ол білім беру ортасындағы эмоцияларды бағалау үшін қолданылатын әдістерді, құралдарды, алгоритмдерді және технологиялық тәсілдерді егжей-тегжейлі талдауды қамтиды. Практикалық құндылығы: қосымша деректерді (аудио, мәтіндік, сенсорлық көрсеткіштер) пайдалана отырып, мультимодальды модельдерге одан әрі көшу бойынша әдістемелік ұсыныстарды дайындау.

Түйін сөздер: Эмоциялар, цифрландыру, жасанды интеллект, интеллектуалды, білім беру ортасы, білім алушылар, инклюзив

© Бекманова Г.Т., Кантуреева М.А., Омарбекова А.С., Закирова А.Б.,
Исайнова А.Н., 2026.

Евразийский национальный университет имени Л.Н. Гумилева,
Астана, Казахстан.

E-mail: kantureyeva_ma@enu.kz

**ИНТЕГРАЦИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОЦЕНКИ
ЭМОЦИЙ В УЧЕБНОЙ СРЕДЕ**

Бекманова Гульмира — кандидат технических наук, PhD, ассоциированный профессор, член правления - проректор по цифровизации, цифровой офицер, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: bekmanova_gt@enu.kz, <https://orcid.org/0000-0001-8554-7627>;

Кантуреева Мансия — PhD, старший преподаватель кафедры информационных систем факультета информационных технологий, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: mansiya_arynbekovna@mail.ru, <https://orcid.org/0000-0001-5904-820X>;

Омарбекова Асель — кандидат технических наук, ассоциированный профессор, директор департамента цифрового развития и онлайн-обучения, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: omarbekova_as@enu.kz, <https://orcid.org/0000-0002-9272-8829>;

Закирова Алма — кандидат педагогических наук, ассоциированный профессор, директор департамента науки, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: zakirova_ab@enu.kz, <https://orcid.org/0000-0001-8772-1414>;

Исайнова Алия — старший преподаватель кафедры информационной безопасности Института цифровых наук и искусственного интеллекта, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: issainova_an@enu.kz, <https://orcid.org/0000-0001-8008-4096>.

Аннотация. *Актуальность.* Цифровизация современного образования сопровождается стремительным ростом объема мультимодальных данных, формируемых в процессе обучения. Видеозаписи занятий, аудиофрагменты, текстовые взаимодействия в системах управления обучением (LMS), а также телеметрия активности пользователей отражают не только когнитивную, но и эмоциональную составляющую образовательного процесса. Эмоции обучающихся являются важным индикатором вовлеченности, интереса, мотивации и эффективности усвоения учебного материала. Их корректное распознавание позволяет адаптировать образовательный контент к эмоциональной реакции аудитории, своевременно выявлять признаки стресса, снижения мотивации и учебной перегрузки. *Цель.* Проанализировать возможности интеграции искусственного интеллекта для оценки эмоций в учебной среде и определить перспективы применения мультимодальных моделей для поддержки образовательного процесса. *Методы.* В работе рассмотрены научно обоснованные и практически применимые подходы к оценке эмоций в образовательной среде, включая методы компьютерного зрения, обработки естественного языка, анализа аудиоданных и мультимодального машинного обучения. Особое внимание уделено моделям, ориентированным не на индивидуальную оценку личности, а на совокупный анализ эмоциональной динамики учебных групп. Такой подход позволяет использовать технологии искусственного интеллекта в образовательной аналитике с учетом этических ограничений, требований конфиденциальности и задач педагогического сопровождения. *Результаты и выводы.* Проведенный анализ позволил классифицировать ключевые направления применения искусственного интеллекта для исследования и оценки эмоций

в образовательной среде. Выявлены основные технологические подходы, инструменты и алгоритмы, применяемые для анализа эмоциональных состояний обучающихся в условиях общего, профессионального, дистанционного и инклюзивного образования. Установлено, что наибольший потенциал имеют мультимодальные модели, объединяющие видео, аудио, текстовые данные и сенсорные индикаторы. Практическая значимость исследования заключается в подготовке методических рекомендаций для дальнейшего перехода к мультимодальным моделям анализа эмоций, которые могут использоваться для повышения вовлеченности обучающихся, поддержки преподавателей и совершенствования цифровых образовательных платформ.

Ключевые слова: искусственный интеллект, цифровизация образования, эмоции в обучении, мультимодальные данные, образовательная аналитика, LMS, эмпирические данные, пилотное исследование

Kіpіcne: Жасанды интеллект біздің қоғамның көптеген салаларын, соның ішінде білім беруді өзгертеді. Бұл тұрғыда эмоциялар оқу үлгеріміне, мотивацияға, ақпаратты сақтауға және білім алушылардың әл-ауқатына әсер ететіндігін ескере отырып, оқыту-оқу процесінде негізгі рөл атқарады. Осылайша, жасанды интеллекттің білім беру ортасындағы эмоционалды бағалауға интеграциясы білім алушылардың әлеуметтік-эмоционалды дамуын қарастыру болып табылады. Дегенмен, осы саладағы жетістіктерді, қиындықтарды және мүмкіндіктерді жүйелейтін кешенді тәсілдің жоқтығы байқалады. Жасанды интеллекттің білім беру мекемелеріне интеграциясы білім алушылардың эмоцияларын анықтауда, бағалауда және тәрбиелеуде айтарлықтай ілгерілеуді білдіреді. Жасанды интеллекттің оқу процесінде жиналған деректер арқылы күрделі эмоционалдық мінез-құлық үлгілерін талдау қабілеті әрбір білім алушының қажеттіліктерін тереңірек түсінуге мүмкіндік береді.

Жетілдірілген алгоритмдерді қолдана отырып, жасанды интеллект көңілсіздік, зерігу немесе ынта белгілерін анықтай алады, бұл оқытушыларға оқыту әдістерін тиімдірек бейімдеуге мүмкіндік береді. Сонымен қатар, жасанды интеллект эмоционалды талдауға негізделген, жекелендірілген кері байланысты қамтамасыз ете алады, осылайша білім алушылардың эмоционалдық әл-ауқатына көбірек бейімделген оқу ортасын жасайды. Бұл кешенді тәсіл білім алушылардың жан-жақты дамуына, эмоцияларды басқару, жағымды қарым-қатынас орнату және оқу үлгерімін жақсарту қабілеттерін арттыруға айтарлықтай ықпал етеді.

Әдеби шолу. Осыған байланысты (Dmello et al., 2012) жасанды интеллект «интеллектуалды репетиторлық жүйемен» өзара әрекеттесу кезінде білім алушылар мен оқытушылар арасындағы диалогтардың мәтінін талдау арқылы білім алушылардың эмоцияларын (зерігу, еркін сөйлеу/қатысу, шатасу және көңілсіздік) болжай алатынын айтады. Жасанды интеллектке негізделген бұл

интеллектуалды оқыту жүйелері «Мобильді бейімделген жекелендірілген оқыту ортасы» (Mehigan et al., 2019) сияқты жасанды интеллектке негізделген білім беру үлгілерін енгізу арқылы білім алушылардың мотивациясына оң әсер етуі мүмкін. Осылайша, синтезделген эмоциялары бар жасанды оқытушылар өздерінің мінез-құлқын білім алушылардың реакциялары мен аффективті күйлеріне бейімдей алады, олардың e-learning жүйелеріндегі жұмысын жақсартады (Florea et al., 2005).

Тағы бір қызықты зерттеу жасанды интеллект интерактивті цифрлық оқыту орталарында (IDLE) білім алушылардың эмоцияларын анықтауға және бағалауға және олардың нақты қажеттіліктерін қанағаттандыру үшін қоршаған ортаны сәйкесінше бейімдеуге, оқуды жақсартуға көмектесетінін дәлелдейді (Arguel et al., 2019). Жасанды интеллект сонымен қатар білім алушылардың эмоцияларын иммерсивті ортамен өзара әрекеттесу кезінде жіктей алады, бұл олардың эмоционалдық тәжірибесін жақсырақ түсінуге мүмкіндік береді (Rodriguez et al., 2020). Сонымен қатар, ол эмоциялардың қарқындылығын бағалай алады және сабақтард).

Жасанды интеллект сонымен қатар мәтіннен алынған эмоцияларды талдай алады, білім алушылардың мотивациясын және e-learning орталарындағы үлгерімін арттырады (Rodriguez et al., 2012). Сонымен қатар, ол эмоциялардың қарқындылығын бағалай алады және сабақтарды жеке қажеттіліктерге бейімдей алады, бұл академиялық зерттеулердің сәтті аяқталуына ықпал етеді (Sumithra et al., 2022). Сол сияқты, жақында жүргізілген зерттеу білім алушылардың эмоцияларын анықтау үшін терең оқыту әдістерін қолдану өнімділікті айтарлықтай арттырып, оқу процесін жақсарта алатынын көрсетті (Alzubi et al., 2022).

Жасанды интеллектті білім беру мекемелеріндегі эмоционалды басқаруға енгізу бойынша күш-жігердің ішінде «биологиялық шабыттандырылған когнитивті архитектура» (eBICA) ерекше орын алады. (Samsonovich et al., 2020) әзірлеген eBICA жасанды интеллектке әлеуметтік өзара әрекеттесу кезінде адамның эмоцияларын түсінуге және олармен өзара әрекеттесуге мүмкіндік береді. Сонымен қатар, эмоцияларға негізделген жасанды шешім қабылдау моделі виртуалды жағдайда білім беру агенттерінің жұмысын жақсартатыны көрсетілген (Yang et al., 2014). Тағы бір тәсіл әлеуметтік-эмоционалдық климатты жақсарту арқылы білім алушылардың мотивациясын, өзін-өзі бағалауын және өзін-өзі ынталандыруын жақсарту үшін ЖИ негізіндегі оқу орталарына эмоционалды агенттерді біріктіруді қамтиды (Gorga et al., 2009), әсіресе аффективті есептеулер (Kort et al., 2001); (González-Hernández et al., 2018); (Ninaus et al., 2019); (Shobana et al., 2021); (Villegas-Ch et al., 2023).

Соңғы жетістіктер жасанды интеллект эмоцияларды танып қана қоймай, болжай алатынын көрсетеді (Alm et al., 2005); (Lin et al., 2023); (Singh et al., 2023). Бұл мүмкіндік виртуалды көмекшілер және интеллектуалды репетиторлық жүйелер (ITS) сияқты жүйелерге білім алушылардың эмоционалдық қажеттіліктеріне белсенді түрде бейімделуге және тиімдірек

жауап беруге, осылайша оқу тәжірибесін жақсартуға мүмкіндік беретін ағымдағы эмоционалдық күйлерді анықтаудан асып түседі.

Сонымен қатар, жасанды интеллект эмпатикалық мінез-құлықты білім беру жағдайлары сияқты динамикалық әлеуметтік контексттерде талдай алады. Қазіргі уақытта эмоционалды интеллектті дамыту, мультимодальды эмоционалды сигналдарды өңдеу, тиісті эмпатикалық реакцияларды қалыптастыру үшін терең оқытуды қолданатын модельдер бар (*Alanazi et al., 2023*).

Әдістері мен процедуралары. Тақырыптық кластерлерді анықтау шолуға енгізілген зерттеулерді талдау және синтездеу процесі арқылы жүзеге асырылды. Қолданылатын критерийлер төмендегідей көрсетіледі:

1. Тақырыптық жиілік: бұл критерий қарастырылған зерттеулерде белгілі бір тақырыптардың немесе ұғымдардың пайда болу жиілігін анықтауға мүмкіндік берді. Бұл қарастырылған зерттеулердің тақырыптарында, тезистерінде, кілт сөздерінде және бөлімдерінде негізгі терминдердің заңдылықтарын іздеуді және жазуды қамтыды.

2. Қолданылатын жасанды интеллект технологиясы: бұл критерий қарастырылған зерттеулерде қолданылатын нақты жасанды интеллект технологияларына негізделген. Ол білім беру ортасындағы эмоцияларды бағалау үшін қолданылатын әдістерді, құралдарды, алгоритмдерді және технологиялық тәсілдерді егжей-тегжейлі талдауды қамтиды.

3. Қолдану саласы немесе көлемі: бұл критерий білім беру мекемелеріндегі, соның ішінде арнайы білім берудегі эмоцияларды бағалау үшін жасанды интеллект технологиялары қолданылған нақты контексттерге бағытталған. Ол зерттеулердің жалпы білім беру, кәсіптік оқыту немесе қашықтықтан оқыту, сондай-ақ ерекше білім беру қажеттіліктері бар білім алушыларға арналған арнайы білім беру сияқты белгілі бір салаларға бағытталғанын зерттейді.

4. Нәтижелер: Бұл критерий әрбір зерттеудің нәтижелерін, әсіресе білім беру ортасындағы эмоцияларды бағалауда жасанды интеллектті біріктірудің тиісті аспектілеріне қатысты зерттеуге мүмкіндік берді. Байқалған әсерлер, қол жеткізілген қорытындылар және оқу практикасына салдары қарастырылды.

Осы критерийлер анықталғаннан кейін біз келесі жеті сатылы процедураны орындай отырып, тақырыптық кластерлерді анықтау процесін жалғастырамыз:

1. Зерттеуді таңдау: біз академиялық мәліметтер базасы мен мамандандырылған іздеу жүйелерін қолдана отырып, тиісті әдебиеттерді жан-жақты іздеуден бастадық. Білім беру мекемелеріндегі эмоцияларды бағалау үшін жасанды интеллектті интеграциялау тақырыбын қарастыратын тиісті зерттеулерді таңдау үшін алдын ала анықталған қосу және алып тастау критерийлері қолданылды.

2. Ақпаратты алу: белгіленген критерийлерге сүйене отырып, зерттеу тобы әрбір таңдалған зерттеуден қайталанатын тұжырымдамалар мен процестер сияқты негізгі ақпаратты алу процесін бастады; Қолданбалы жасанды интеллект технологиясы, араласу контексті, нәтижелер және негізгі

қорытындылар. Бұл ақпарат зерттеулерді талдау мен салыстыруға берік негіз болды.

3. Пайда болған тақырыптарды анықтау: барлық алынған ақпарат қайталанатын тақырыптарды анықтау және жасанды интеллектті интеграциялауға және білім беру ортасындағы эмоцияларды бағалауға байланысты пайда болатын заңдылықтарды ұйымдастыру үшін зерттелді. Бұл оның мазмұны мен мәнмәтінін түсіну үшін алынған ақпаратпен әрбір зерттеуді қайта оқуды қамтыды.

4. Деректерді Кодтау: Әрбір пайда болған тақырыпқа немесе үлгіге кодтар немесе белгілер тағайындалды.

5. Тақырыптық кластерлерге топтастыру: әрбір зерттеуге немесе фрагментке тағайындалған кодтарды пайдалана отырып, кодтар когерентті тақырыптық кластерлерге топтастырылды. Бұл процесс кодталған ақпарат арасындағы ұқсастықтар мен қатынастарды анықтауды және оларды білім беру жағдайындағы эмоцияларды бағалау үшін жасанды интеллектті интеграциялаудың нақты аспектілерін қарастыратын топтарға ұйымдастыруды қамтыды.

6. Талдау және синтез: тақырыптық кластерлеу аяқталғаннан кейін тенденцияларды, сәйкессіздіктерді және қызығушылықтың елеулі салаларын анықтау үшін әрбір кластердегі зерттеулерге егжей-тегжейлі талдау жүргізілді. Бұл кезең жиналған ақпаратты синтездеуге және тақырып бойынша қарастырылған әдебиеттерге контексттік көзқарас беруге мүмкіндік берді.

Бұл тақырыптық кластерлер жиналған ақпаратты талдау және синтездеу үшін үйлесімді құрылымды қамтамасыз ету мақсатында ұйымдастырылды, осылайша білім беру ортасындағы эмоцияларды бағалау үшін жасанды интеллект интеграциясының тенденциялары мен жетістіктерін түсінуге ықпал етті.

Нәтижелер. Алғышарттар. Зерттеу ЖОО-ның оқу аудиторияларын қамтитын және қауіпсіздікті қамтамасыз ету мақсатында жұмыс жасайтын бейнебақылау инфрақұрылымы негізінде жүзеге асырылады. Камералар стационарлық түрде орнатылған, суретті жоғарыдан түсіреді және дыбыстық сигнал бермейді. Бейне ағыны ішкі ережелерге сәйкес жергілікті жерде сақталады және теканалитикалық мақсаттарда қолданылады. Мұндай деректер форматы білім алушылардың тұлғаларын сәйкестендіруді болдырмайды, бұл этикалық қағидаттар мен құпиялылық нормаларының сақталуын қамтамасыз етеді. Әдістемелік кезеңнің мақсаты – анонимдендірілген бейнежазбаларды пайдалана отырып білім алушылардың белсенділігі, назары және қатысу деңгейін сипаттайтын объективті көрсеткіштерді жеке деректерсіз алу.

Зерттеу нысаны және үлгісі

Зерттеу нысаны ЖОО-ның типтік аудиторияларында әртүрлі пәндердің (дәрістік және практикалық) оқу сабақтарының бейнежазбалары болып табылады. Іріктемеге мыналар кіреді:

- білім алушылар қатарларына тұрақты шолу жасайтын жоғары камералары бар аудиториялар;

- ұзақтығы 30 минуттан кем емес жазбалар;

- әр түрлі толтыру тығыздығы (кіші, орташа, толық) және тәуліктің әр түрлі уақыты бар сессиялар. Бастапқы талдау үшін сабақтың әр түрлі форматтары мен білім алушылардың орналасуын білдіретін 50-100 бейнежазба үзінділерінен (ұзақтығы 1-3 минут) іріктеу жоспарлануда.

Зерттеу кезеңдері

1 кезең. Деректерді алу және алдын-ала өңдеу

- Білім алушылардың негізгі қатарларының көрінуін қамтамасыз ететін бейне сегменттерді таңдау.

- Ағынды бірыңғай форматқа түрлендіру (25 fps, 720p ажыратымдылығы).

- Адамдарды анықтау және қаңқа моделдерін құру үшін компьютерлік көру әдістерін қолдану (дене мен бастың негізгі нүктелері).

- Камераның перспективасы мен биіктігін ескере отырып, координаттарды қалыпқа келтіру.

2 кезең. Мінез-құлық белгілерін (behavioral features) алу

Әрбір кадр немесе қысқа уақыт терезесі үшін (1-3 секунд) келесі сипаттамалар есептеледі:

- дененің еңкеюі (алға, артқа, бейтарап);

- дененің тақтаға немесе мұғалімге қатысты бағыты;

- позаның өзгеру жиілігі;

- қол қимылдарының амплитудасы мен жиілігі;

- статикалық қалыптардың ұзақтығы;

- қатысуды білдіретін қимылдардың пайда болуы (қол көтеру, ым-ишара).

3 кезең. Жиынтық индекстерді қалыптастыру

Білім алушылардың белсенділік көрсеткіштері төрт прокси метрику есептеу үшін аудитория бойынша орташаланады және жинақталады:

1. Attention Index (AI) – экранға немесе дәріс аймағына бағытталған білім алушылардың үлесі.

2. Engagement Index (EI) – қатысуды көрсететін қимылдардың, қолды көтерудің, микро қимылдардың салыстырмалы саны.

3. Calmness Index (CI) – «жұмыс» қалпын сақтай отырып, мотор тұрақтылығының деңгейі.

4. Group Interaction Index (GI) – ұжымдық өзара әрекеттесуді көрсететін білім алушылар арасындағы бағдарлау мен микро қозғалыстардың синхрондылығы. Барлық көрсеткіштер уақыт терезелері үшін есептеледі (мысалы, 30 секунд), бұл сабақ барысында аудиторияның эмоционалды күйінің динамикасын құруға мүмкіндік береді.

4 кезең. Эмоционалды фонды жіктеу

Индекстердің тіркесімі негізінде аудиторияның көңіл күйінің интегралды сипаттамасы анықталады:

- Позитивті – жоғары AI және EI, орташа GI, орташа CI;

- Бейтарап – орташа AI, төмен EI, әлсіз GI, орташа-жоғары CI;
- Теріс – төмен AI және EI, тұрақсыз GI, CI экстремалды мәндері (қатып қалған немесе ретсіз қозғалыс).

Қағидаларды валидациялау үшін оқытушылар мен зерттеушілер жекелеген фрагменттерге сараптамалық аннотация жүргізеді.

5 кезең. Түсіндіру және визуализация

- Сабақ барысында индекстердің өзгеруінің уақыт кестесін құру.
- Аудитория аймақтары бойынша белсенділіктің жағымды карталарын жасау.
- Әр түрлі іс-шаралар мен топтардың эмоционалды фонының динамикасын салыстыру.

Этикалық аспектілер

Зерттеу қатысушыларды сәйкестендірусіз, бет әлпетсіз және дыбыстық ақпаратсыз жүргізіледі. Барлық деректер ЖОО-ның қорғалған инфрақұрылымында қалады және тек жалпыланған түрде қолданылады. Аналитикалық тұжырымдар жеке тұлға емес, топ деңгейінде түсіндіріледі. Зерттеу нәтижелерін қолдану оқыту әдістемесін жақсартуға және оқу процесінің тиімділігін арттыруға бағытталған, ал білім алушылардың жеке басына әсер етпейді.

Күтілетін нәтижелер

- Белсенділік пен зейіннің есептелген индекстері бар бейне сегменттердің пилоттық деректер жиынтығын құру.
- Аудиторияның эмоционалды фонын анықтау үшін прокси-метриканы есептеу алгоритмдерін әзірлеу.
- Құпиялылық саясатын бұзбай топтық қатысуды талдау үшін стандартты бақылау камераларын пайдалану мүмкіндігін растау.
- Қосымша деректерді (аудио, мәтін, сенсорлық көрсеткіштер) пайдалана отырып, мультимодальды моделдерге одан әрі көшу үшін әдістемелік ұсынымдар дайындау.

Білім алушылардың эмоционалдық күйлерін тану міндеті бойынша жүргізілген эксперименттің нәтижелері

1.Сөйлеу - аудио түріндегі мәліметтердің бірі және ол *сараланбаған деректер базасына* жатады.

Адамның дауысын талдау жасанды интеллект саласында кең қолданылады:

- эмоцияларды тануда,
- аффективті жүйелерде,
- автоматты субтитр жасауда,
- білім беру аналитикасында.

Төмендегі кестеде негізгі сөйлеу белгілерінің түрлері көрсетілген:

Кесте 1 — Негізгі сөйлеу белгілерінің түрлері

Белгі	Сипаттамасы	Интерпретациясы
Амплитуда (Energy)	Дауыс сигналының қаттылығы	Энергия жоғары болса — дауыс қатты, эмоциялы (мысалы: ашу, қуаныш)
Pitch (F ₀) — негізгі жиілік	Дауыс байламдарының тербелу жылдамдығы	Pitch жоғары → қуаныш; Төмен → мұң, тыныштық
MFCC	Дауыс спектрін ықшам түрінде сипаттайтын коэффициенттер	Тембр, эмоция және сөйлеудің сипатын анықтауда маңызды

Қолданылатын құралдар:

- Python, кітапханалар: librosa, matplotlib, numpy, pandas
- Jupyter Notebook
- .wav форматындағы аудио файлдар

Жұмыс барысы

1-кезең. Орта дайындау

Қажетті кітапханаларды орнату:

```
pip install librosa matplotlib numpy pandas
```

2-кезең. Аудиофайлды жүктеу

Код арқылы файлды оқимыз, ұзақтығын және дискретизация жиілігін тексереміз.

```
Code: Uploading an audio file # Uploading an audio file audio_path = «voice_sample.wav» # replace it with the path to your file y, sr = librosa.load(audio_path, sr=None) print(«Sampling rate:», sr) print(«Duration (sec):», librosa.get_duration(y=y, sr=sr))
```

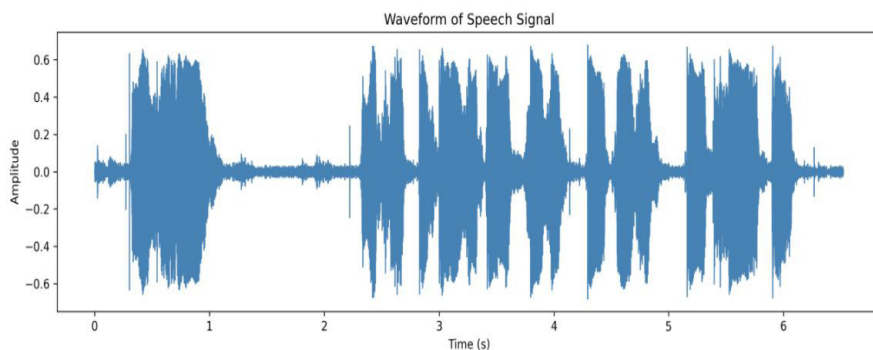
3-кезең. Сигналды визуализациялау

Аудиосигнал графигін салу арқылы: дауыс қатты шыққан бөліктер тоқтаулар анық көрінеді.

Дискретизация жиілігі — дыбыс бір секундта қанша рет өлшенетінін көрсетеді.

```
Code: plt.figure(figsize=(12, 4))
librosa.display.waveshow(y, sr=sr, color='steelblue')
plt.title("Waveform of Speech Signal")
plt.xlabel("Time (s)")
plt.ylabel("Amplitude")
plt.show()
```

Енді бізде дауыс жазбаларының амплитудалары бар ол арқылы біз қай кезде дыбыс қатты шыққанын немесе жәй шыққанын анықтай аламыз.



Сурет 1—Дауыс жазбаларының амплитудаларын анықтау

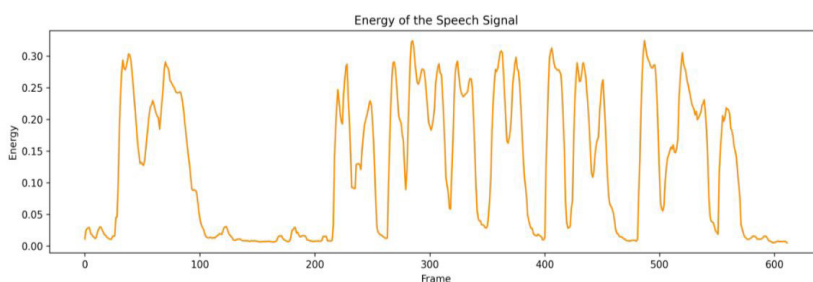
4-кезең. Белгілерді (фичаларды) шығару

4.1. Энергия (Amplitude / RMS)

Осы кезеңде энергия графигі дауыс қаншалықты қатты немесе әлсіз екенін көрсетеді.

КОД:

```
energy = librosa.feature.rms(y=y)[0]
plt.figure(figsize=(12,4))
plt.plot(energy, color='darkorange')
plt.title("Energy of the Speech Signal")
plt.xlabel("Frame")
plt.ylabel("Energy")
plt.show()
```



Средняя энергия: 0.1263

Сурет 2 — Орташа энергияны есептеу

Есептелген орташа энергия: 0.1263

Бұл салыстырмалы түрде жоғары, дауыс белсенді, яғни эмоция бар.

4.2. Pitch (негізгі жиілік)

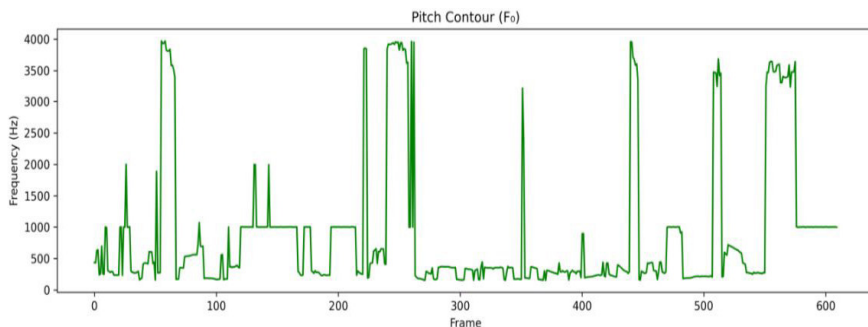
Орташа мәні: 891.86 Гц

Бұл өте жоғары көрсеткіш, ол қуаныш, белсенділік, көтеріңкі көңіл-күйге тән.


```

pitches, magnitudes = librosa.piptrack(y=y, sr=sr)
pitch_values = pitches[magnitudes > np.median(magnitudes)]
plt.figure(figsize=(12,4))
plt.plot(pitch_values[:500], color='green')
plt.title("Pitch Contour (F0)")
plt.xlabel("Frame")
plt.ylabel("Frequency (Hz)")
plt.show()

```



Сурет 3 —Негізгі орташа жиілікті есептеу

Орташа жиілік (Pitch): 891.86 Гц

Future -Mean value

Energy-(RMS),0.12633914

Pitch-(F0),891.86017

4.3. MFCC

MFCC — сөйлеудің тембрін, артикуляциясын және эмоциясын сипаттайтын маңызды параметр.

Графиктегі:

Төменгі жолақтардың қызыл-көк болып тез ауысуы спектрдің жиі өзгеруі → сөйлеудің динамикалық, интонациясы бай екенін көрсетеді.

Бұл қуанышты, әсерлі немесе эмоционалды сөйлеуге тән белгі.

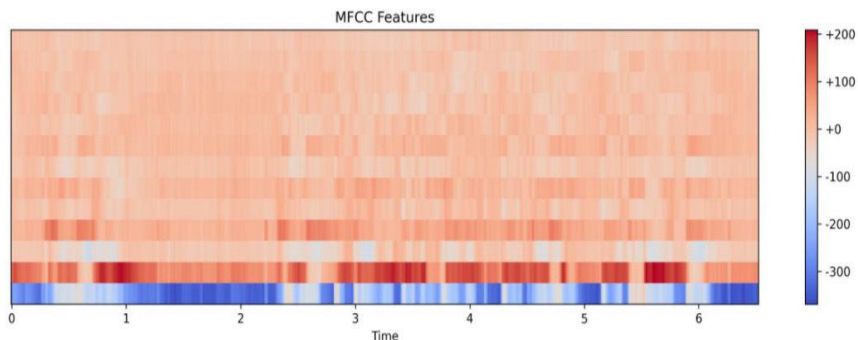
```
mfccs = librosa.feature.mfcc(y=y, sr=sr, n_mfcc=13)
```

```
librosa.display.specshow(mfccs, x_axis='time', sr=sr)
```

```
plt.colorbar()
```

```
plt.title("MFCC Features")
```

```
plt.show()
```



Сурет 4 — Сөйлеудің динамикалық, интонациясын көрсету

2.Эмоцияларды анықтау жүйелері үшін деректерді өңдеудің ETL және ELT тәсілдері қарастырылады. Деректерді жүктеудің ерекшеліктері, автоматтандырылған кодтар, SQL кестелерін жасау және нормализациялау қағидалары толық ашылады. Эмоцияларды тану үлгісінің тиімділігі валенттілік, белсенділік, эмоция түрі және уақыт белгісі сияқты маңызды параметрлерге байланысты. Осы деректерді жүйелі, дәл және бірізді түрде өңдеу модельдің нақтылығын тікелей анықтайды. Деректерді өңдеу кезеңдерінде әрбір параметрдің маңыздылығы мен өңдеу ерекшеліктері қарастырылады.

ETL архитектурасының жалпы құрылымы:

1. Дерек көздері (Деректердің бастаулары)

Эмоциялық деректер алуан түрлі пішімдерде ұсынылуы мүмкін:

- CSV форматындағы файлдар
- JSON құрылымды мәліметтер
- Журналдық файлдар
- Бейнекамералардан алынған өңделген деректер
- API арқылы алынған эмоциялық мәліметтер немесе нәтижелер
- Kafka секілді хабар алмасу жүйелерінің деректері.

Аталған деректердің бәрі шикі деректер (өңделмеген мәліметтер) болып саналады.

2. Шығару (Деректерді алу)

Бұл кезеңде коннекторлар мәліметтерді оқып, оларды сақтау/архивтік аймаққа бағыттайды.

Мақсаттары:

- Деректерге ең аз әсер ету
- Деректердің толықтығын қамтамасыз ету
- Мәліметтерді қатесіз қабылдау
- Үстеме жүктемелерді қолдау
- Шикі деректер аймағында мәліметтер өзгеріссіз күйінде сақталады.

3. Трансформация (Деректерді түрлендіру)

Түрлендіру – ETL процесінің ең күрделі бөлігі:

- Деректерді нормализациялау
- Қайталанатын деректерді жою
- Валенттілік пен белсенділікті қайта есептеу
- Аномальды мәндерді тазарту
- Деректер типтерін өзгерту
- Белсенділік формуласы бойынша есептеу.

Бұл операциялар SQL немесе Python тілдерінде жүзеге асырылады.

Түрлендірудің нәтижесі:

- Silver (тазартылған мәліметтер)
- Gold (талдауға дайын деректер қоймасы) болып бөлінеді.

4. Жүктеу

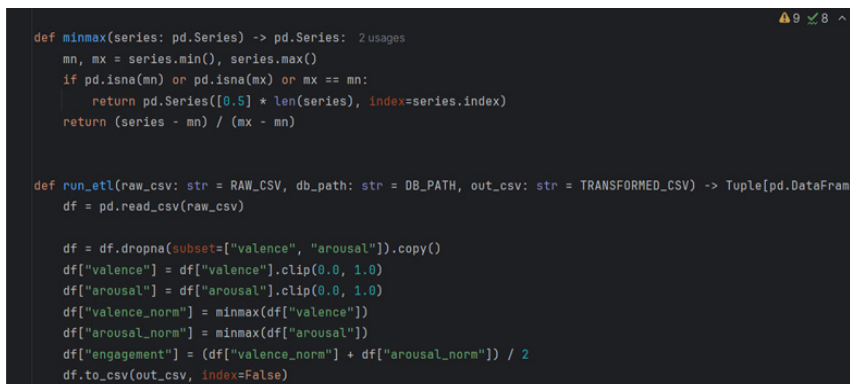
Деректер келесі жүйелерге жүктеледі:

- PostgreSQL
- SQLite
- ClickHouse
- Snowflake
- BigQuery

Жүктеу толық немесе үстемелі түрде орындалуы мүмкін.

5. Оркестрация

ETL процестерін жоспарлау Airflow немесе Prefect құралдары арқылы іске асырылады:



```
def minmax(series: pd.Series) -> pd.Series:
    mn, mx = series.min(), series.max()
    if pd.isna(mn) or pd.isna(mx) or mx == mn:
        return pd.Series([0.5] * len(series), index=series.index)
    return (series - mn) / (mx - mn)

def run_etl(raw_csv: str = RAW_CSV, db_path: str = DB_PATH, out_csv: str = TRANSFORMED_CSV) -> Tuple[pd.DataFrame]:
    df = pd.read_csv(raw_csv)

    df = df.dropna(subset=["valence", "arousal"]).copy()
    df["valence"] = df["valence"].clip(0.0, 1.0)
    df["arousal"] = df["arousal"].clip(0.0, 1.0)
    df["valence_norm"] = minmax(df["valence"])
    df["arousal_norm"] = minmax(df["arousal"])
    df["engagement"] = (df["valence_norm"] + df["arousal_norm"]) / 2
    df.to_csv(out_csv, index=False)
```

Сурет 5 — Деректерді нормализациялау:

3. Эмоционалдық күйлерді қолмен белгілеудің аналитикалық зерттеуі

Мұнда Emotions_Marking_DospolNazgul.xlsx деректер жиынтығына негізделген, қолмен таңбаланған эмоциялық үзінділерді аналитикалық өңдеу және салыстырмалы талдау ұсынылады. Аталған файлда әрбір бейнеклипке қатысты төрт негізгі параметр қарастырылған:

Назар индексі (AI) – зейін деңгейі; Қатысу индексі (EI) – тартылу деңгейі; Тыныштық индексі (CI) – сабырлылық деңгейі; Топтық әрекеттесу индексі (GI) – ұжымдық өзара ықпалдасу.

Сонымен қатар, әрбір видеофрагмент үшін эмоциялық жағдай қолмен көрсетілген.

Бастапқы мәліметтерді жүктеп, жиынтық шолу жасау.

Файлдағы бағандар ретке келтіріліп, базалық айнымалылар бойынша алғашқы статистика есептелді.

Алғашқы талдаудың қорытындылары: Жалпы саны 10 видеоклип таңбаланған. Кездесетін эмоциялық ахуалдар: Оң; Бейтарап; Жағымсыз (немқұрайлылық); Жағымсыз (дүрбелең).

Индекстердің статистикалық мәліметтері (ең төменгі–ең жоғарғы–орташа).

```
import pandas as pd

# Загрузка Excel файла
df = pd.read_excel("/Users/Nazgul/Desktop/fragments/Emotions_Marking_DospolNazgul.xlsx")

# Убираем лишние пробелы в названиях колонок
df.columns = df.columns.str.strip()

# Количество строк
num_rows = df.shape[0]

# Уникальные значения эмоционального состояния
unique_moods = df['Эмоциональное состояние'].unique()

# Мин/макс/среднее по индексам
stats = df[['AI (внимание)', 'EI (вовлеченность)', 'CI (спокойствие)', 'GI (взаимодействие)']].agg(['min', 'max', 'mean'])

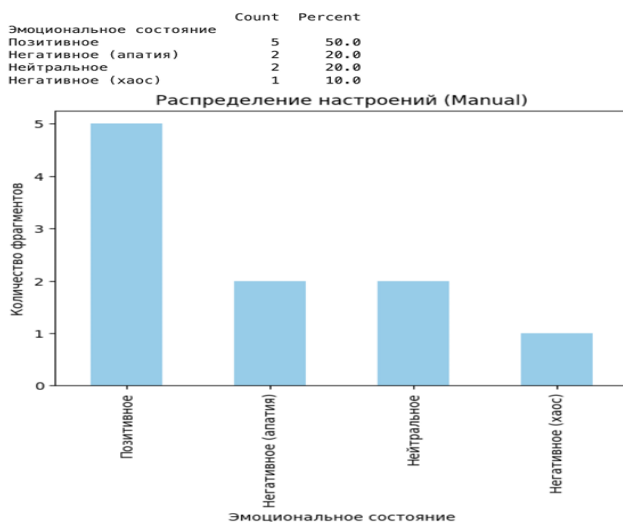
print(f"Количество фрагментов: {num_rows}")
print(f"Типы настроения: {unique_moods}")
print(stats)
```

Количество фрагментов: 10
 Типы настроения: ['Позитивное' 'Негативное (апатия)' 'Негативное (хаос)' 'Нейтральное']

	AI (внимание)	EI (вовлеченность)	CI (спокойствие)	GI (взаимодействие)
min	0.000	0.000	0.200	0.00
max	1.000	0.850	0.900	0.50
mean	0.475	0.475	0.475	0.475

Сурет 6 — Эмоциялық ахуалдардың таралуын анықтау

Диаграмма мен кесте негізінде эмоциялардың таралуы анықталды.



Сурет 7 — Эмоциональдық ахуалдар фрагменттерін көрсету

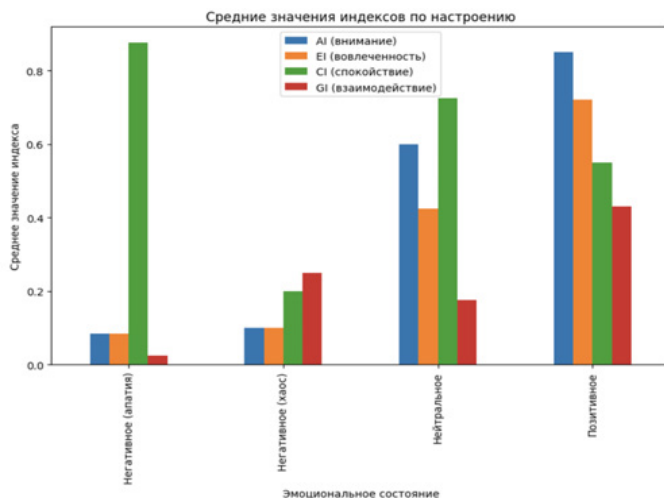
Эмоция түрлері бойынша фрагменттердің саны: Оң эмоциялар ең көп (50%). Жағымсыз эмоциялар 30%-ды құрайды, ал бейтарап күй аз кездеседі.

Үшінші кадам. Қолмен белгілеу мен автоматты жіктеуді салыстыру.

Кестеде қолмен белгілеу (manual) және автоматты жүйе (auto) нәтижелері салыстырылды. Нәтижелер: Сәйкестік 70 пайызды құрады.

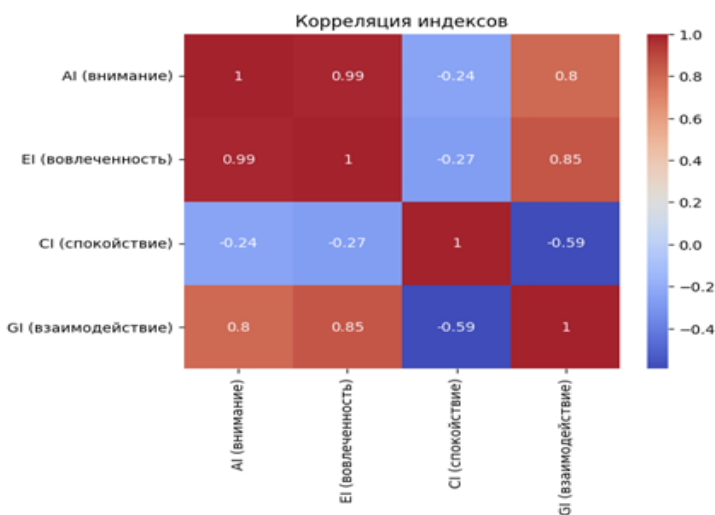
Талдау: Сәйкессіздіктер негізінен оң және бейтарап күйлердің шекарасында пайда болған.

Төртінші кадам. Индекстерді эмоцияларға сәйкес талдау.



Сурет 8 — Көңіл-күйге қарай индекстердің орташа мәні

Бесінші кезең.



Сурет 9 — Индекс корреляциясы

Индекстер арасындағы өзара байланыс. Қорытынды 10 фрагментке жүргізілген талдау нәтижесінде келесі тұжырымдар жасалды:

Эмоциялардың жартысы оң (50%), бұл оқудың жалпы эмоционалдық ахуалында оң өзгерістер басым екенін көрсетеді. Жағымсыз күй 30 пайызды құрайды.

Алғашқы бейнематериалдарды сараптау – эмоцияларды анықтайтын жасанды интеллект үлгісін жасаудағы қажетті кадам. Түсірілім сапасы, жарықтың мөлшері, камераның орналасу бұрышы мен студенттердің анық көрінуі үлгінің тиімді жұмыс істеуіне тікелей ықпал етеді. Осыған байланысты, берілген бөлімде әрбір көрсеткіш тереңінен қарастырылады.



Сурет 10 — Бейнематериалдарды сараптау кезіндегі түсірілім

Видеоға түсірілген фрагменттердің қысқаша сипаттамасы, олардың эмоциялық ахуалына негізделген.

№	Имя видео	Время начала	Время конца	AI	EI	CI	GI	Эмоциональное состояние	Комментарий
1	B02-R122_08_10_2025 13.12.56_2_Оспан_Балмахан_B02-R122_время_00_1	0:00	1:00	0,59	0,32	0,6	0,4	Нейтральное	Тихая работа, умеренная активность
2	B06-R201A_06_10_2025 13.12.56_Балмахан_Оспан_B06-R201A_время08_3	8:34	9:20	0,75	0,62	0,5	0,48	Позитивное	Активные ответы, вовлеченность высокая
3	B02-R108_08_10_2025 13.12.56_2_Балмахан_Оспан_B02-R108_время09_2	8:21	9:20	0,55	0,45	0,65	0,35	Нейтральное	Спокойное слушание
4	B02-R116_08_10_2025 12.12.49_2_Балмахан_Оспан_B02-116_время_00_00	0:00	0:54	0,47	0,22	0,83	0,1	Негативное (апатия)	Пассивность
5	B02-R108_08_10_2025 13.12.56_2_Балмахан_Оспан_B02-R108_время10_5	10:51	11:40	0,44	0,25	0,2	0,15	Негативное (хаос)	Шум и рассеянность
6	B06-R201A_06_10_2025 13.12.56_Балмахан_Оспан_B06-R201A_время07_0	7:08	8:02	0,73	0,68	0,58	0,55	Позитивное	Групповое обсуждение
7	B02-R116_08_10_2025 12.12.49_2_Балмахан_Оспан_B02-116_время_00_54	0:54	1:47	0,41	0,18	0,85	0,08	Негативное (апатия)	Нет реакции
8	B06-R301A_06_10_2025 13.12.56_Балмахан_Оспан_B06-R301A_время14_5	14:56	15:42	0,78	0,62	0,5	0,42	Позитивное	Работа в группе
9	B06-R301A_06_10_2025 13.12.56_Балмахан_Оспан_B06-R301A_время22_0	22:04	22:59	0,76	0,63	0,45	0,6	Позитивное	Высокая вовлеченность
10	B02-R122_08_10_2025 13.12.56_2_Оспан_Балмахан_B02-R122_время_01_1	1:00	2:00	0,52	0,38	0,68	0,3	Нейтральное	Минимальное взаимодействие

Сурет 11 — Әрбір фрагментке оның эмоциялық күйін көрсету

Талқылау. Топтастырылған үзінділердегі студенттердің эмоциялық жағдайлары AI (зейін), EI (қатысу), CI (сабырлылық) және GI (топтық өзара

әрекеттесу) көрсеткіштері бойынша бағаланды. Әрбір фрагментке оның эмоциялық күйі мен қысқаша сипаттамасы қоса беріледі.

1-фрагмент — Бейтарап жағдай АІ орташа, ЕІ төмендеу тенденциясында, СІ мен ГІ тұрақты. Іс-әрекеттің сипаттамасы: тыныш жұмыс, қалыпты белсенділік.

2-фрагмент — Оң жағдай АІ жоғары, ЕІ өте жоғары, ГІ да айтарлықтай жоғары. Белсенді түрде жауап беру, қатысудың жоғары деңгейі.

3-фрагмент — Бейтарап жағдай Көрсеткіштер тұрақты, СІ жоғарырақ мәнге ие. Сабақтың материалын тыныштықта қабылдау.

4-фрагмент — Жағымсыз (апатия) АІ мен ЕІ өте төмен, ГІ ең төменгі мәнде. Белсенділіктің жоқтығы, пассивтілік.

5-фрагмент — Жағымсыз (хаос) АІ төмен, ЕІ мен СІ тұрақсыз. Шулы, алаңдаушылық көріністері, зейіннің шашырауы.

6-фрагмент — Оң жағдай АІ жоғары, ЕІ айқын, ГІ орташа. Топтық пікірталас, белсенді қатысу байқалады.

7-фрагмент — Жағымсыз (апатия) АІ, ЕІ және ГІ төмен мәндерге ие. Ешқандай реакция байқалмайды, белсенділік өте төмен.

8-фрагмент — Оң жағдай АІ жоғары, ЕІ жоғары, ГІ жеткілікті деңгейде. Топтағы жұмыс, қатысуы жақсы деңгейде.

9-фрагмент — Бейтарап жағдай Көрсеткіштер орташа деңгейде, ГІ айтарлықтай. Өзара әрекеттесу минималды, қалыпты жұмыс.

Қорытынды. Қорытындылай келе, жиналған мәліметтер мен жүргізілген ғылыми-техникалық сараптама нәтижесінде эмоцияларды анықтайтын жасанды интеллект үлгісін құруға қажетті алғышарттардың бәрі орындалды деп айтуға болады. Бұл жоба бастапқы кезеңін табысты аяқтап, зерттеу мен әзірлеудің келесі сатыларына көшуге дайын. Аталған жұмыс жасанды интеллект саласымен қатар, Қазақстандағы білім беру аналитикасының дамуына, оқу үдерісін цифрландыруға және студенттермен жүргізілетін жұмыстардың сапасын жақсартуға зор ықпал етеді. Бұл зерттеулер мен әзірлемелердің келесі кезеңдеріне жол ашады.

Көмекші технологияларды, ақпараттық-коммуникациялық технологияларды (АКТ) және жасанды интеллектті білім беру мекемелеріне біріктіру білім алушыларға, әсіресе оқуда қиындықтары бар білім алушыларға олардың оқуы мен эмоцияларын басқаруда қол жетімді қолдауды түбегейлі өзгертті. Әртүрлі оқу қажеттіліктері бар балалар үшін жасанды интеллекттің көмегімен эмоцияларды анықтаудың жоғарылауы, АКТ арқылы жекелендірілген оқу тәжірибесі және көмекші технологиялардың қолжетімділігін арттыру оқудағы кедергілерді айтарлықтай азайтты. Бұл зерттеу эмоцияны тануды арттырудағы, инклюзивті оқу ортасын құрудағы және барлық балалардың оқудағы жетістіктеріне ықпал етудегі технологияның маңызды рөлін көрсетеді. Осы озық құралдарды пайдалана отырып, оқытушылар теңшелген оқу жоспарларын жасай алады, дереу кері байланыс бере алады және мүмкіндігі шектеулі және мүмкіндігі шектеулі білім алушылардың академиялық және

эмоционалдық дамуына қолдау көрсете алады. Жасанды интеллект, АКТ және көмекші технологияларды біріктірудің бұл кешенді тәсілі эмоционалды қолдауды күшейтіп қана қоймайды, сонымен қатар білім алушыларды олардың білім алуына белсенді қатысу үшін қажетті құралдармен қамтамасыз етеді. Сайып келгенде, бұл неғұрлым табысты және инклюзивті білім беру үдерісіне жол ашады.

References

- Shabbir M., Alshammari N., Hussain I., (2023) Emotional prediction in intelligent agents to facilitate social interaction. *Appl. Sci*13:1163. doi: 10.3390/app13021163 (in Eng.)
- Aly M., Ghallab A. Fathi S. (2023) Enhancing facial expression recognition system in online learning context using efficient deep learning model. *IEEE Access*11. — P. 121419–121433. doi: 10.1109/ACCESS.2023.3325407 (in Eng.)
- Abu Zitar R., Hawashin B., Abu Shanab S., Zraiqat A., Mughaid A. (2022). A novel deep learning technique for detecting emotional impact in online education. *Electronics*11:2964. doi: 10.3390/electronics11182964 (in Eng.)
- Arguel A., Lockyer L., Kennedy G., Lodge J., Pachman M. (2019). Seeking optimal confusion: a review on epistemic emotion management in interactive digital learning environments. *Interact. Learn. Environ.*27. — P. 200–210. doi: 10.1080/10494820.2018.1457544 (in Eng.)
- Arnau-Gonzalez P., Katsigiannis S., Arevalillo-Herraez M., Ramzan N. (2021). “Artificial intelligence for affective computing: an emotion recognition case study” in *AI for emerging verticals: human-robot computing, sensing and networking*. eds. Shakir M. Z. Ramzan N. — P. 29–44. (in Eng.)
- Assielou K., Haba C., Kadjo T., Goore B., Yao K D. (2021). A new approach to modelling students socio-emotional attributes to predict their performance in intelligent tutoring systems. *J. Educ. E Learn. Res.*8. — P. 340–348. doi: 10.20448/JOURNAL.509.2021.83.340.348 (in Eng.)
- Begum F., Neelima A., Valan J. (2023). Emotion recognition system for E-learning environment based on facial expressions. *Soft. Comput.*27, 17257–17265. doi: 10.1007/s00500-023-08058-3 (in Eng.)
- Bevilacqua L., Capuano N., Cascone A., Ceccarini F., Corvino F. (2009). Advanced user interfaces for e-learning. *J.E Learn. Knowledge Soc.*5, 91–99. doi: 10.20368/1971-8829/356 (in Eng.)
- Cambria E., Das D., Bandyopadhyay S., Feraco A. (2017). Affective computing and sentiment analysis. *Pract Guide Sentiment Analysis*. — P. 1–10. doi: 10.1007/978-3-319-55394-8_1 (in Eng.)
- Carmona-Halty M., Schaufeli W., Llorens S., Salanova M. (2019). Satisfaction of basic psychological needs leads to better academic performance via increased psychological capital: a three-wave longitudinal study among high school students. *Front. Psychol.*10:2113. doi: 10.3389/fpsyg.2019.02113 (in Eng.)
- Chaffar S., Frasson C. (2010). “Using emotional coping strategies in intelligent tutoring systems” in *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, 6095 LNCS(PART2). Berlin, Heidelberg: Springer. — P. 285–287. (in Eng.)
- Chowdary M., Nguyen T., Hemanth D. (2023). Deep learning-based facial emotion recognition for human–computer interaction applications. *Neural Comput. & Applic.*35, 23311–23328. doi: 10.1007/s00521-021-06012-8 (in Eng.)
- Colchester K., Hagraas H., Alghazzawi D., Aldabbagh G. (2017). A survey of artificial intelligence techniques employed for adaptive educational systems within e-learning platforms. *J. Artif. Intelligence Soft Comput. Res.*7. — P. 47–64. doi: 10.1515/jaiscr-2017-0004 (in Eng.)
- Daouas T., Lejmi H. (2018). Emotions recognition in an intelligent elearning environment. *Interact. Learn. Environ.*26. — P. 991–1009. doi: 10.1080/10494820.2018.1427114 (in Eng.)
- Daghestani L., Ibrahim L., Al-Towirgi R., Salman H. (2020). Adapting gamified learning systems using educational data mining techniques. *Comput. Appl. Eng. Educ.*28, 568–589. doi: 10.1002/cae.22227 (in Eng.)

Oliveira E., Rodrigues P. (2021). A review of literature on human behaviour and artificial intelligence: contributions towards knowledge management. *Electron. J. Knowl. Manag.*19. — P. 165–179. doi: 10.34190/ejkm.19.2.2459 (in Eng.)

Florea A., Kalisz E. (2005). Embedding emotions in an artificial tutor. In *Seventh International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC'05)* (in Eng.)

Garcia-Ceja E., Riegler M., Nordgreen T., Jakobsen P., Oedegaard K., Tørresen J. (2018). Mental health monitoring with multimodal sensing and machine learning: a survey. *Pervasive Mobile Comput*51. — P. 1–26. doi: 10.1016/j.pmcj.2018.09.003 (in Eng.)

Gomez-Leon M. (2022). Development of empathy through socioemotional artificial intelligence. *Papeles Psicol*43. — P. 218–224. doi: 10.23923/pap.psicol.2996 (in Eng.)

Gonzalez-Hernandez F., Zatarain-Cabada R., Barron-Estrada M., L.Rodriguez-Rangel H. (2018). Recognition of learning-centered emotions using a convolutional neural network. *J. Intelligent Fuzzy Syst*34. — P. 3325–3336. doi: 10.3233/JIFS-169514 (in Eng.)

Gorga D., Schneider D. (2009). “Computer-based learning environments with emotional agents” in *Handbook of research on synthetic emotions and sociable robotics: new applications in affective computing and artificial intelligence*. — P. 413–441. (in Eng.)

He T., Li C., Wang J., Wang M., Wang Z., Jiao C. (2022). An emotion analysis in learning environment based on theme-specified drawing by convolutional neural network. *Front. Public Health*10:958870. doi: 10.3389/fpubh.2022.958870(in Eng.)

Jaques P., Vicari R. (2006). Considering Students' emotions in computer-mediated learning environments. *Web-Based Intelligent E Learn Syst*. — P. 122–138. doi: 10.4018/978-1-59140-729-4.CH006 (in Eng.)

Jiang D., Wu K., Chen D., Tu G., Zhou T., Garg A. (2020). A probability and integrated learning based classification algorithm for high-level human emotion recognition problems. *Measurement*150:107049. doi: 10.1016/j.measurement.2019.107049 (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 146–165

<https://doi.org/10.32014/2026.2518-1726.431>

IRSTI 28.23.29

UDC 004.8:37

© **Dzhusupbekova G.T., Jangassiyev R.M.***, 2026.

Mukhtar Auezov South Kazakhstan Research University, Shymkent, Kazakhstan.

E-mail: jangasiev@gmail.com

**GEMINI AI INTEGRATION BASED ON .NET MAUI FOR EDUCATION:
HYBRID ARCHITECTURE AND EMPIRICAL LOAD TESTING****Dzhusupbekova Gulzat** — Candidate of Pedagogical Sciences, Senior Lecturer, Mukhtar Auezov South Kazakhstan Research University, Shymkent, Kazakhstan,E-mail: d.gulzat20.10@gmail.com, <https://orcid.org/0000-0003-1727-0966>;**Jangassiyev Raiymbek** — Lecturer, Mukhtar Auezov South Kazakhstan Research University, Shymkent, Kazakhstan,E-mail: jangasiev@gmail.com, <https://orcid.org/0009-0000-7616-419X>.

Abstract. In the context of global digital transformation and the transition to the Quality 4.0 paradigm, higher education requires new approaches to the design and deployment of educational technologies. In this regard, the integration of large multimodal models into educational applications has become an important direction for providing personalized and interactive learning support. The aim of this study is to propose a hybrid architecture for an educational application based on the .NET MAUI cross-platform framework and Google Gen AI SDK, to develop its prototype, and to empirically evaluate its technical performance. As part of the study, a working prototype of an intelligent advisor application was developed for Windows and Android operating systems. During the empirical evaluation, load testing was conducted on the Gemini 2.5 Flash and Gemma 3 (1B) models using prompts of varying complexity; the measured metrics included Time-To-First-Token (TTFT), generation speed (TPS), RAM usage, and network traffic volume. The results showed that Gemma 3 (1B) had an advantage in initial response speed, with TTFT remaining below 1 second, making it suitable for rapid interface interaction and local processing of sensitive data. In contrast, Gemini 2.5 Flash demonstrated high throughput when generating long and complex content, reaching over 200 TPS. These findings confirm the technical feasibility of a hybrid architecture for educational applications. At the same time, the study substantiates the need to maintain a human-in-the-loop strategy and strengthen academic integrity policies when using multimodal models in high-stakes educational tasks.

Keywords: Gemini AI, .NET MAUI, educational technology, hybrid architecture, load testing, multimodal models, empirical evaluation

For citations: Dzhusupbekova G.T., Jangassiyev R.M. Gemini ai integration based on .net maui for education: hybrid architecture and empirical load testing. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 146–165. DOI <https://doi.org/10.32014/2026.2518-1726.431>

© Джусупбекова Г.Т., Жангасиев Р.М. *, 2026.

Мұхтар Әуезов атындағы Оңтүстік Қазақстан зерттеу университеті,
Шымкент, Қазақстан.

E-mail: jangasiev@gmail.com

БІЛІМ БЕРУГЕ АРНАЛҒАН .NET MAUI НЕГІЗІНДЕГІ GEMINI AI ИНТЕГРАЦИЯСЫ: ГИБРИДТІ АРХИТЕКТУРА ЖӘНЕ ЭМПИРИКАЛЫҚ ЖҮКТЕМЕЛІК ТЕСТІЛЕУ

Джусупбекова Гулзат — п.ғ.к., аға оқытушы, Мұхтар Әуезов атындағы Оңтүстік Қазақстан зерттеу университеті, Шымкент, Қазақстан,

E-mail: d.gulzat20.10@gmail.com, <https://orcid.org/0000-0003-1727-0966>;

Жангасиев Райымбек — оқытушы, Мұхтар Әуезов атындағы Оңтүстік Қазақстан зерттеу университеті, Шымкент, Қазақстан,

E-mail: jangasiev@gmail.com, <https://orcid.org/0009-0000-7616-419X>.

Аннотация. Жаһандық цифрлық трансформация және «Сапа 4.0» парадигмасына көшу жағдайында жоғары білім беру жүйесі білім беру технологияларын жаңаша ұйымдастыруды талап етеді. Осы тұрғыдан алғанда, үлкен мультимодальды модельдерді білім беру қосымшаларына интеграциялау жекелендірілген және интерактивті оқу қолдауын қамтамасыз етудің өзекті бағыты болып табылады. Бұл зерттеудің мақсаты – .NET MAUI кросс-платформалық фреймворкі мен Google Gen AI SDK негізінде білім беруге арналған гибриді архитектураны ұсыну, оның прототипін әзірлеу және техникалық өнімділігін эмпирикалық тұрғыда бағалау. Зерттеу аясында Windows және Android операциялық жүйелеріне арналған «зияткерлік кеңесші» қосымшасының жұмыс істейтін прототипі жасалды. Эмпирикалық бағалау барысында Gemini 2.5 Flash және Gemma 3 (1B) модельдеріне әртүрлі күрделіліктегі промпттар арқылы жүктемелік тестілеу жүргізіліп, алғашқы токенге дейінгі уақыт (TTFT), генерация жылдамдығы (TPS), жедел жадты пайдалану және желілік трафик көлемі өлшенді. Нәтижелер Gemma 3 (1B) моделінің жедел бастапқы жауап беру бойынша артықшылығын көрсетті: оның TTFT көрсеткіші 1 секундтан төмен болды, бұл жедел интерфейстік өзара әрекеттесу мен құпия деректерді локальды өңдеу сценарийлеріне қолайлы. Ал Gemini 2.5 Flash моделі көлемді және күрделі мазмұнды генерациялауда жоғары өткізу қабілетін көрсетті, TPS көрсеткіші 200-ден асты. Алынған

нәтижелер гибридіті архитектураның білім беру қосымшаларында техникалық тұрғыдан орынды екенін дәлелдейді. Сонымен қатар мультимодальды модельдерді жоғары жауапкершілікті білім беру міндеттерінде қолданғанда адам қатысуымен тексеру қағидатын сақтау және академиялық адалдық саясатын күшейту қажеттілігі негізделді.

Түйін сөздер: Gemini AI, .NET MAUI, білім беру технологиялары, гибридіті архитектура, жүктемелік тестілеу, мультимодальды модельдер, эмпирикалық бағалау

© Джусупбекова Г.Т., Джангасиев Р.М.* , 2026.

Южно-Казахстанский исследовательский университет имени М.О. Ауэзова,
Шымкент, Казахстан.
E-mail: jangasiev@gmail.com

ИНТЕГРАЦИЯ GEMINI AI НА БАЗЕ .NET MAUI ДЛЯ ОБРАЗОВАНИЯ: ГИБРИДНАЯ АРХИТЕКТУРА И ЭМПИРИЧЕСКОЕ НАГРУЗОЧНОЕ ТЕСТИРОВАНИЕ

Джусупбекова Гулзат — кандидат педагогических наук, старший преподаватель, Южно-Казахстанский исследовательский университет имени М.О. Ауэзова, Шымкент, Казахстан,
E-mail: d.gulzat20.10@gmail.com, <https://orcid.org/0000-0003-1727-0966>;

Джангасиев Райымбек — преподаватель, Южно-Казахстанский исследовательский университет имени М.О. Ауэзова, Шымкент, Казахстан,
E-mail: jangasiev@gmail.com, <https://orcid.org/0009-0000-7616-419X>.

Аннотация. *Актуальность.* В условиях глобальной цифровой трансформации и перехода к парадигме «Качество 4.0» система высшего образования требует новых подходов к организации образовательных технологий. В этом контексте интеграция больших мультимодальных моделей в образовательные приложения становится актуальным направлением развития персонализированной и интерактивной учебной поддержки. *Цель.* Предложить гибридную архитектуру образовательного приложения на основе кроссплатформенного фреймворка .NET MAUI и Google GenAI SDK, разработать ее прототип и эмпирически оценить техническую производительность. *Методы.* В рамках исследования создан рабочий прототип приложения «интеллектуальный советник» для операционных систем Windows и Android. В ходе эмпирической оценки проведено нагрузочное тестирование моделей Gemini 2.5 Flash и Gemma 3 (1B) с использованием промптов различной сложности. Измерялись время до первого токена (TTFT), скорость генерации токенов (TPS), использование оперативной памяти и объем сетевого трафика. *Результаты и выводы.* Результаты показали преимущество модели Gemma 3 (1B) по скорости начального отклика: показатель TTFT составил менее 1 секунды, что делает данную модель подходящей для сценариев быстрого интерфейсного взаимодействия и локальной обработки чувствительных

данных. В то же время модель Gemini 2.5 Flash продемонстрировала высокую пропускную способность при генерации объемного и сложного контента, показав TPS выше 200. Полученные результаты подтверждают техническую целесообразность применения гибридной архитектуры в образовательных приложениях. Одновременно обоснована необходимость соблюдения принципа human-in-the-loop и усиления политики академической честности при использовании мультимодальных моделей в высокоответственных образовательных задачах.

Ключевые слова: Gemini AI, .NET MAUI, образовательные технологии, гибридная архитектура, нагрузочное тестирование, мультимодальные модели, эмпирическая оценка, Google Gen AI SDK

Кіріспе. Жоғары білім беру жүйесіндегі цифрлық трансформация білім беру технологияларын жаңа деңгейде қайта қарауды талап етеді. Қазіргі жағдайда жасанды интеллект білім беру үдерісінің шеткері құралы емес, оқу материалын түсіндіру, жекелендірілген қолдау көрсету, интерактивті кеңес беру және цифрлық білім беру қызметтерін ұйымдастырудың маңызды элементіне айналып отыр. Әсіресе генеративті жасанды интеллект құралдарының дамуы студенттермен өзара әрекеттесудің жаңа үлгілерін қалыптастырып, білім беру мазмұнын түсіндіру, құрылымдау және жедел кері байланыс ұсыну мүмкіндігін кеңейтті (Aktay et al., 2023; Imran and Almusharraf, 2023).

Соңғы жылдары білім беру ортасында мәтінге ғана емес, кескін, аудио және басқа да дерек түрлеріне сүйенетін мультимодальды жасанды интеллект жүйелеріне қызығушылық артты. Бұл үрдіс нақты оқу міндеттерінің табиғатымен байланысты, өйткені білім беру тапсырмалары көбіне тек мәтінмен шектелмейді. Диаграммалар, кестелер, код үзінділері, визуалды материалдар және күрделі құрылымдық мазмұн білім беру қосымшаларына мультимодальды модельдерді енгізуді өзекті етеді (Lee and Shi, et al., 2023; Imran and Almusharraf, 2024b). Осы тұрғыдан Google Gemini модельдері мәтіндік және мультимодальды мазмұнды өңдеуге қабілетті құрал ретінде білім беру технологиялары үшін маңызды платформа болып саналады (Gemini Team, Google, 2025).

Сонымен қатар білім беру саласында мұндай модельдерді пайдаланудың тиімділігі тек модельдің теориялық мүмкіндігімен емес, оны нақты бағдарламалық ортаға дұрыс кіріктірумен де анықталады. Осыған байланысты кросс-платформалық мобильді және жұмыс үстелі қосымшаларын әзірлеуге арналған .NET MAUI фреймворкі ерекше назар аудартады. Бұл орта Android, Windows, iOS және macOS жүйелеріне ортақ C# кодтық базасы арқылы қосымша жасауға мүмкіндік береді, яғни білім беру ұйымдары үшін бірыңғай, сүйемелдеуге ыңғайлы және кеңейтілетін цифрлық шешім ұсынады (Blanco and Lucrédio, 2021; Jangassiyev et al., 2025).

Алайда қолданыстағы зерттеулерді талдау Gemini және басқа мультимодальды модельдердің білім беру әлеуеті жан-жақты

талқыланымен, оларды .NET MAUI негізіндегі нақты кросс-платформалық қосымшаға кіріктіру және оның техникалық өнімділігін эмпирикалық тұрғыда бағалау мәселесі жеткілікті деңгейде ашылмағанын көрсетеді. Әсіресе гибриді архитектураны пайдалану, яғни бұлттық және локальды модельдер арасындағы функциялық бөліністі жобалау, жауап беру жылдамдығы мен ресурстарды пайдалану тұрғысынан әлі де нақты инженерлік дәлелдеуді қажет етеді.

Осыған байланысты бұл зерттеудің мақсаты – .NET MAUI негізінде Gemini AI интеграцияланған білім беру қосымшасының гибриді архитектурасын ұсыну, оның жұмыс істейтін прототипін әзірлеу және Gemini 2.5 Flash пен Gemma 3 (1B) модельдерінің техникалық өнімділігін эмпирикалық жүктемелік тестілеу арқылы бағалау.

Зерттеудің негізгі үлестері төмендегідей:

1. білім беру сценарийлеріне бағытталған .NET MAUI негізіндегі гибриді архитектура ұсынылды;

2. Windows және Android платформаларына арналған кросс-платформалық «зияткерлік кеңесші» прототипі әзірленді;

3. Gemini 2.5 Flash пен Gemma 3 (1B) модельдері TTFT, TPS, жедел жадты пайдалану және желілік трафик метрикалары бойынша эмпирикалық түрде салыстырылды;

4. алынған техникалық нәтижелердің білім беру қосымшаларындағы практикалық қолданысымен байланысы талданды.

Осы мақсатқа жету үшін зерттеу келесі сұрақтарға жауап іздейді: .NET MAUI негізінде Gemini AI-ді білім беру қосымшасына тиімді түрде қалай кіріктіруге болады; гибриді архитектура жағдайында әртүрлі модельдердің қолданылу сценарийлері қалай бөлінеді; және техникалық өнімділік нәтижелері білім беру ортасындағы нақты пайдалану талаптарымен қалай байланысады.

Бұл зерттеуде гибриді архитектураны қолдану білім беру қосымшасында жылдамдық, ресурстық тиімділік және қолдану сценарийіне бейімделу арасында практикалық теңгерім береді деп болжанды.

Әдеби шолу. Жасанды интеллектті білім беру экожүйесіне енгізу дәстүрлі цифрландырудан зияткерлік және бейімделгіш оқу орталарына көшу үдерісін жеделдетті. Бұл үрдіс жоғары оқу орындарында оқытуды ұйымдастыру, білім алушыларға қолдау көрсету және цифрлық инфрақұрылымды жаңарту міндеттерімен тығыз байланысты. «Сапа 4.0» парадигмасы аясында білім беру технологиялары енді тек ақпарат жеткізу құралы емес, деректерге негізделген, жекелендірілген және интерактивті білім беру ортасын қалыптастырушы фактор ретінде қарастырылады (Badshah et al., 2023; Imran and Almusharraf, 2024a).

Генеративті жасанды интеллекттің алғашқы кең таралған түрлері, әсіресе мәтінге негізделген ірі тілдік модельдер, академиялық жазу, идея генерациясы, мәтінді құрылымдау және ақпараттық қолдау сияқты міндеттерде пайдалы

құрал ретінде бағаланды. Бұл модельдер студенттерге мәтін дайындау, мазмұнды өңдеу және оқу тапсырмаларын түсіндіру барысында жедел көмек көрсетуге мүмкіндік берді (Imran and Almusharraf, 2023; Koubaa et al., 2023). Сонымен бірге зерттеулер мұндай құралдарды қолдану академиялық адалдық, мазмұн дәлдігі және сыни ойлаудың әлсіреуі сияқты тәуекелдерді де қатар туындататынын көрсетеді (Aktay et al., 2023). Білім беру міндеттерінің күрделенуі мәтінмен ғана шектелмейтін, сурет, аудио, бейне және аралас мазмұнды қатар өңдей алатын мультимодальды модельдерге сұранысты арттырды. Осы тұрғыдан Google Gemini мәтіндік және визуалды деректерді біріктіріп өңдеуге бағытталған жаңа буын мультимодальды модель ретінде қарастырылады (Gemini Team, Google, 2025; Lee, Shi, et al., 2023). Gemini архитектурасының ерекшелігі – әртүрлі модальділіктер арасында өзара байланысты пайымдауды қолдауы, бұл оны білім беру саласындағы күрделі сценарийлерге, мысалы, түсіндірме беру, код генерациясы, диаграммалармен жұмыс істеу және интерактивті қолдау міндеттеріне бейімдеуге мүмкіндік береді.

Бірқатар зерттеулер Gemini-дің академиялық жазу мен оқу қолдауындағы әлеуетін атап өтеді. Azmi және Fithriani (2025) студенттердің Gemini-ді мәтін құрылымын жетілдіру, байланыстылықты арттыру және жазу сапасын жақсарту құралы ретінде оң бағалағанын көрсетеді. Сонымен бірге Imran және Almusharraf (2024b) Gemini-ді білім берудегі келесі буын AI құралы ретінде сипаттап, оның мультимодальды өзара әрекеттесуге және интерактивті оқытуға қолайлы екеніне назар аударады. Бұл еңбектер Gemini-дің білім алушыларға көмекші құрал ретінде тиімділігін көрсеткенімен, оны нақты бағдарламалық ортаға енгізудің техникалық аспектілерін толық ашпайды.

Gemini-дің маңызды артықшылықтарының бірі – түрлі өлшемдегі модельдер желісі арқылы әртүрлі қолдану сценарийлерін қолдауы. Бұлттық модельдер күрделі аналитикалық және көлемді генерация міндеттеріне бейімделсе, ықшам нұсқалар локальды немесе шеткі құрылғыларда жылдам жауап беруге қолайлы. Осындай икемділік білім беру қосымшалары үшін маңызды, себебі оларда бір жағынан жедел интерфейстік әрекет, екінші жағынан мазмұнды кеңейтілген түрде генерациялау қажеттілігі қатар кездеседі (Gemini Team, Google, 2025). Сондықтан гибриді архитектура, яғни бұлттық және локальды модельдерді қатар пайдалану, білім беру технологияларында перспективалы бағыт ретінде қарастырылуы мүмкін.

Сонымен қатар әдебиеттер Gemini және өзге мультимодальды модельдердің шектеулерін де көрсетеді. Әсіресе визуалды дәлдікке тәуелді жоғары жауапкершілікті міндеттерде, мысалы автоматтандырылған бағалау немесе күрделі визуалды интерпретацияда, нәтижелер әрдайым тұрақты бола бермейді (Fu et al., 2023; Lee, Latif, et al., 2023). Бұл тұжырымдар мультимодальды модельдерді білім беру ортасында пайдаланғанда толық автономиядан гөрі, адам қатысуымен тексерілетін көмекші құрал ретінде қарастырудың орынды екенін көрсетеді.

Кросс-платформалық әзірлеу тұрғысынан алғанда, .NET MAUI білім беру қосымшаларын бірнеше операциялық жүйеге ортақ кодтық база негізінде жасауға мүмкіндік беретін маңызды технологиялық орта болып табылады. Мұндай тәсіл әзірлеу үдерісін жеңілдетіп қана қоймай, цифрлық білім беру шешімдерінің қолжетімділігін, сүйемелденуін және институционалдық ауқымда енгізілуін жақсартыды (Blanco and Lucrédio, 2021; Jangassiyev et al., 2025). Алайда қазіргі әдебиеттерде .NET MAUI мен Gemini AI-ді біріктіруге арналған инженерлік зерттеулер саны шектеулі. Әсіресе білім беру контекстінде гибриді архитектура құру, нақты прототип әзірлеу және оның техникалық өнімділігін эмпирикалық жүктемелік тестілеу арқылы бағалау жеткілікті деңгейде қарастырылмаған.

Осылайша, әдебиеттерді талдау мынадай ғылыми олқылықты көрсетеді: білім берудегі Gemini AI-дің жалпы әлеуеті мойындалғанымен, оны .NET MAUI негізіндегі кросс-платформалық қосымшаға кіріктіру, гибриді архитектура тұрғысынан ұйымдастыру және оның техникалық өнімділігін нақты метрикалар арқылы эмпирикалық тексеру бағытындағы зерттеулер жеткіліксіз. Осы олқылық аталған жұмыстың өзектілігін анықтайды және зерттеудің эмпирикалық бағытын негіздейді.

Зерттеу әдістемесі. Бұл зерттеуде жобалау, прототиптеу және эмпирикалық жүктемелік тестілеу тәсілдері қолданылды. Зерттеу екі кезеңнен тұрды. Бірінші кезеңде .NET MAUI негізінде білім беруге арналған кросс-платформалық «зияткерлік кеңесші» қосымшасының архитектурасы әзірленіп, жұмыс істейтін прототипі жасалды. Екінші кезеңде Gemini 2.5 Flash және Gemma 3 (1B) модельдерінің техникалық өнімділігі Windows (WinUI) және Android платформаларында эмпирикалық тұрғыда бағаланды.

Тестілеу үшін әртүрлі күрделіліктегі төрт промпт қолданылды: қысқа түсіндірме сұрағы, техникалық код генерациясы, білім беру логикасына бағытталған сұрақ және күрделі интеграциялық жоспар құру тапсырмасы. Бағалау барысында алғашқы токенге дейінгі уақыт (TTFT), толық кідіріс уақыты, генерация жылдамдығы (TPS), жедел жадты пайдалану және желілік трафик көлемі өлшенді.

Зерттеу толық ауқымды педагогикалық эксперимент ретінде емес, білім беру қосымшасы контекстіндегі техникалық орындылық пен өнімділікті бағалауға бағытталған эмпирикалық зерттеу ретінде қарастырылады. Алынған нәтижелер әртүрлі модельдердің нақты қолдану сценарийлеріне сәйкестігін анықтау үшін салыстырмалы түрде талданды.

Эмпирикалық бағалау аясында Windows платформасында 4 тестілік сценарий, ал Android платформасында 8 тестілік сценарий орындалды. Кестелерде әр сценарий бойынша тіркелген нақты өлшеу нәтижелері берілді. Бұл тестілеу пилоттық техникалық бағалау сипатында жүргізілді және модельдердің салыстырмалы өнімділік профилін анықтауға бағытталды. Тестілеу .NET MAUI негізінде әзірленген бірдей қолданбалы ортада жүргізіліп, модельдер бірдей промпттық сценарийлер арқылы салыстырылды.

Зерттеу нәтижелері. Зерттеу нәтижелері екі бағытта ұсынылады: біріншісі – жүйенің архитектуралық жүзеге асырылуы, екіншісі – эмпирикалық жүктемелік тестілеу нәтижелері.

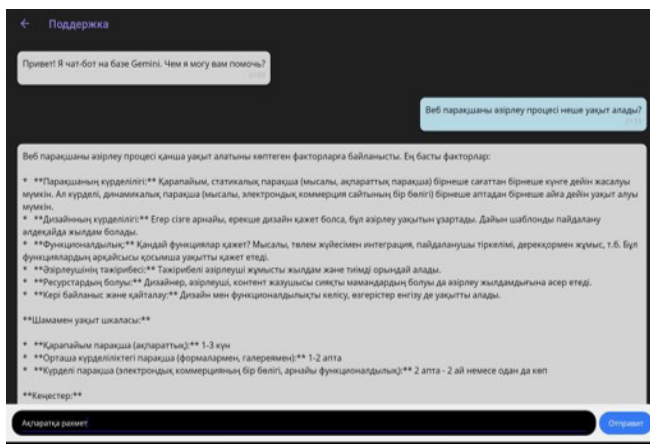
1. Жүйенің архитектурасы және жүзеге асырылуы. .NET MAUI платформасы Windows, Android, iOS және macOS жүйелеріне арналған кросс-платформалық қосымшаларды бірыңғай C# кодтық базасы негізінде әзірлеуге мүмкіндік береді. Бұл тәсіл білім беру ұйымдары үшін цифрлық шешімдерді енгізу мен сүйемелдеуді жеңілдетеді, сондай-ақ әртүрлі операциялық жүйелерде бірдей логиканы сақтауға жағдай жасайды (Jangassiyev et al., 2025). Сонымен қатар мобильді және шеткі ортада есептеу қуатының шектеулілігі, жад көлемінің аздығы және желі тұрақтылығы ірі мультимодальды модельдерді тікелей пайдалануды қиындатады. Осыған байланысты зерттеуде бұлттық және локальды өндеуді біріктіретін гибриді архитектуралар қолданылды. Ұсынылған архитектураларда күрделі және көлемді мазмұнды генерациялауды қажет ететін сұраныстар бұлттық модельдерге бағытталады, ал жедел жауап пен локальды өндеуді талап ететін сценарийлер үшін ықшам модельдерге басымдық беріледі. Мұндай тәсіл жауап беру жылдамдығы, ресурстарды пайдалану және деректер құпиялылығы арасындағы теңгерімді сақтауға мүмкіндік береді. Гибриді стратегияның жалпы логикасы 1-кестеде көрсетілген. Практикалық жүзеге асыру аясында Windows және Android платформаларына арналған «зияткерлік кеңесші» қосымшасының жұмыс істейтін прототипі әзірленді. Қосымшада Gemini API-ге негізделген сұраныс өндеу модулі, көптілді интерфейс, Firebase Authentication арқылы аутентификация, қонақ режимі және Text-to-Speech функциялары енгізілді. Бұл шешімнің интерфейстік көріністері 1–4-суреттерде берілген. Қосымшаның негізгі мақсаты – білім беру және консультациялық сұраныстарды өндеп, пайдаланушыға құрылымдалған жауап ұсыну. Көптілді қолдау қазақ және орыс тілдеріндегі өзара әрекеттесуді қамтамасыз етеді. Қонақ режимі пайдаланушыларға жеке деректерді міндетті түрде енгізбей-ақ қосымшаның негізгі функцияларын пайдалану мүмкіндігін береді. Ал қаріп өлшемін өзгерту мен мәтінді дыбыстау функциялары жүйенің инклюзивтілік деңгейін арттырады.

Кесте 1 — .NET MAUI негізіндегі білім беру қосымшасы үшін модельдерді орналастыру логикасы

Модель нұсқасы	Негізгі мақсатты қолданыс	Негізгі мүмкіндік	Орналастыру стратегиясы	Қолдану негіздемесі
Gemini 2.5 Flash	Күрделі аналитикалық және көлемді генерация міндеттері	Жоғары өнімділік, кең мазмұн генерациясы	Бұлттық API	Ұзын мәтін, код және күрделі жоспар құру міндеттері үшін тиімді
Gemma 3 (1B)	Жылдам бастапқы жауап және ықшам өндеу	Төмен кідіріс, жеңілдетілген инференс	Локальды/ шеткі өндеу	Жедел интерфейстік әрекет пен локальды сценарийлерге қолайлы

Ескертпе – кесте авторлар ұсынған гибриді архитектуралар логикасы негізінде құрастырылды.

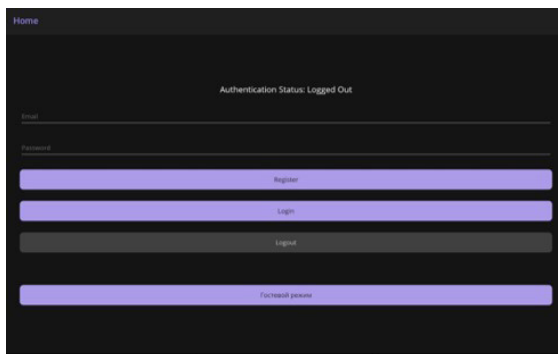
Ұсынылған архитектураның техникалық тұрғыдан жүзеге асырылу мүмкіндігін растау және үлкен тілдік модельдерді (LLM) кросс-платформалық ортаға интеграциялау мүмкіндіктерін көрсету мақсатында, зерттеу аясында үлгілік кеңесші-қосымша әзірленді. Бұл әзірлеме Android, iOS, macOS және Windows операциялық жүйелері үшін бірыңғай код базасын құруды көрсете отырып, Google бұлттық сервистерімен байланыста .NET Multi-platform App UI (.NET MAUI) құралдарын сынақтан өткізуге арналған практикалық полигон қызметін атқарды.



Сурет 1 – Gemini негізіндегі зияткерлік көмекшінің интерфейсі: қазақ тіліндегі сұрауды өңдеу және құрылымдалған жауап беру мысалы

Ескертпе – сурет авторлар әзірлеген қосымшаның интерфейсінен алынған.

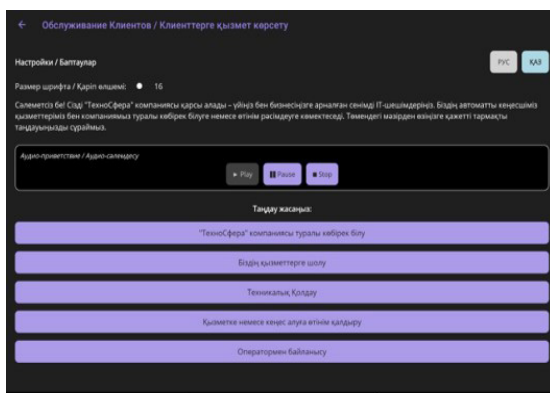
Қосымшаның орталық инновациялық элементі REST API шақырулары арқылы жүзеге асырылған Google Gemini моделі негізіндегі чат-ботты интеграциялау болды. Бұл модуль техникалық сұрауларды өңдеуге және контекстке негізделген, автоматтандырылған жауаптарды ұсынуға қабілетті зияткерлік көмекші қызметін атқарады. Білім беру және консультациялық міндеттер контекстінде жүйенің көптілді ортада жұмыс істеу қабілеті аса маңызды болып табылады. Әзірленген шешім тілдер арасында (орыс және қазақ) динамикалық ауысуды қолдайды, бұл Gemini моделіне сұрауларды дұрыс түсіндіруге және пайдаланушы таңдаған тілде толыққанды жауаптарды жасауға мүмкіндік береді. Мысалы, қазақ тілінде веб-парақшаны әзірлеу мерзімдері туралы сұрау кезінде жүйе күрделі ақпаратты оңай қабылданатын сегменттерге бөле отырып, құрылымдалған әрі маркерленген жауап ұсынады (1-сурет), бұл когнитивті скаффолдинг (cognitive scaffolding) принциптеріне сәйкес келеді.



Сурет 2 – Аутентификация экраны: Firebase арқылы кіру және «Қонақ режимін» (Guest Mode) таңдау мүмкіндігі

Ескертпе – сурет авторлар әзірлеген аутентификация модулінің интерфейсінен алынған.

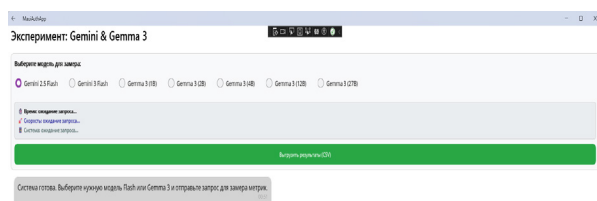
Деректерді сақтау сенімділігін және сессияларды басқаруды қамтамасыз ету үшін Google Firebase қолданылған гибриді архитектура енгізілді. Пайдаланушылардың аутентификациясы Firebase Authentication арқылы жүзеге асырылып, тіркелген қызметкерлер үшін толық қолжетімділікті, сондай-ақ клиенттер үшін «Қонақ режимін» (Guest Mode) қолдайды (2-сурет). Қонақ режимін енгізудің білім беру технологиялары (EdTech) үшін маңызы зор, себебі ол кіру кедергісін төмендетеді және пайдаланушыларға жеке деректерді (PII) беру қажеттілігіңіз негізгі кеңес беру функцияларына қол жеткізуге мүмкіндік береді, бұл құпиялылық саясатының талаптарына сәйкес келеді. Өтінімдер бойынша деректерді жинау валидацияланатын формалар арқылы жүзеге асырылады, олар REST API арқылы құрылымдалған ақпаратты (аты-жөні, байланыс деректері, қызмет түрі) кейінгі өңдеу үшін Firebase NoSQL дерекқорына қауіпсіз түрде жібереді.



Сурет 3 – Қосымшаның баптаулар және қызмет көрсету интерфейсі: инклюзивтілік функциялары (TTS, қаріп өлшемі) және тілді ауыстыру

Ескертпе – сурет авторлар әзірлеген баптаулар интерфейсінен алынған.

Интерфейсті (UI) әзірлеу қолжетімділік принциптерін ескере отырып жүргізілді, бұл қазіргі заманғы білім беру платформалары үшін маңызды талап болып табылады. Қосымша оқылымды жақсарту үшін қаріп өлшемін баптау функцияларын және мәтінді сөзге айналдыру (TTS) сервистерін интеграциялауды қамтиды (3-сурет). TTS модулі сәлемдесу хабарламалары мен ақпараттық блоктарды дыбыстауға қабілетті, бұл ақпаратты қабылдаудың баламалы аудиалды арнасын ұсынады және көру қабілеті бұзылған немесе аудиалды оқуды қалайтын пайдаланушылар үшін қосымшаның қолжетімділігін айтарлықтай арттырады.



Сурет 4 – Gemini 2.5 Flash және Gemma 3 модельдерін тестілеу стенді.

Ескертпе – сурет авторлар әзірлеген эмпирикалық тестілеу стендінің интерфейсін көрсетеді.

Эксперимент жүргізу мақсатында .NET MAUI негізінде арнайы қосымша әзірленді. 4-суретте көрсетілгендей, Gemini 2.5 Flash және Gemma 3 модельдерін тестілеу стенді құрылды. Бұл интерфейс пайдаланушыға қажетті модельді (соның ішінде Gemma 3-тің әртүрлі параметрлі нұсқаларын: 1B, 2B, 12B т.б.) таңдауға және метрикаларды өлшеуге мүмкіндік береді.

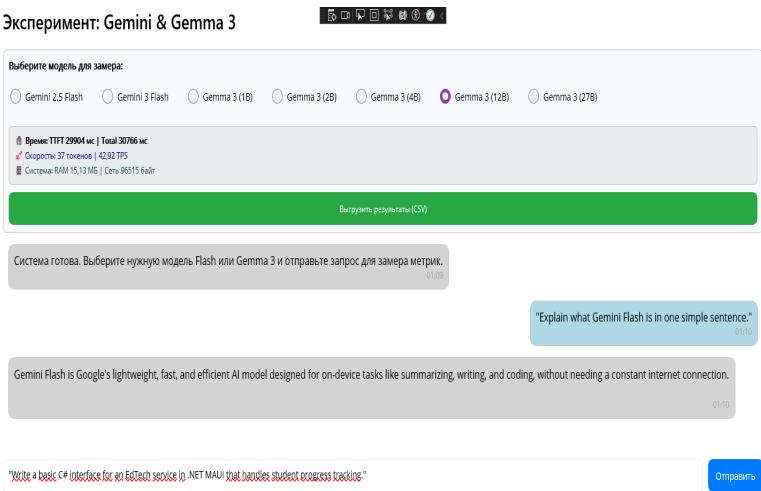
2. Эмпирикалық жүктемелік тестілеу нәтижелері. Тестілеудің мақсаты гибриді архитектура жағдайында әртүрлі модельдердің техникалық мінез-құлқын салыстыру болды. Ұсынылған архитектураны тексеру үшін Windows және Android платформаларына арналған «зияткерлік кеңесші» прототипі әзірленді. Прототипте Gemini API-ге негізделген сұраныс өңдеу модулі, көптілді интерфейс, Firebase арқылы аутентификация және TTS функциялары енгізілді. Бағалау Windows (WinUI) және Android платформаларында жүргізілді. Өлшенген негізгі метрикалар: алғашқы токенге дейінгі уақыт (TTFT), толық кідіріс уақыты, генерация жылдамдығы (TPS), жедел жадты пайдалану және желілік трафик көлемі.

Тестілеуге қолданылған промпттар 2-кестеде берілген. Олар қысқа түсіндірме сұрағынан бастап техникалық код генерациясы мен күрделі интеграциялық жоспарға дейінгі әртүрлі сценарийлерді қамтиды.

Кесте 2 – Gemini API мүмкіндіктерін бағалауға арналған промпттар тізімі

ID	Түрі	Промпт
P1	Short (Latency)	"Explain what Gemini Flash is in one simple sentence."
P2	Technical (Code)	"Write a basic C# interface for an EdTech service in .NET MAUI that handles student progress tracking."
P3	EdTech (Logic)	"List 5 ways AI can improve remote learning for high school students in rural areas."
P4	Complex (Full)	"Create a detailed technical plan for integrating Gemini API into a .NET MAUI mobile app. Include steps for secure API key storage, managing network latency, and implementing a UI loading state for streaming responses."

Ескертпе – тестілеу промпттары авторлар тарапынан эмпирикалық бағалау мақсатында құрастырылды.



Сурет 5 – Gemini және Gemma 3 модельдерін Windows платформасында салыстырмалы тестілеу.

Ескертпе – сурет Windows платформасындағы тестілеу процесінің авторлар тіркеген көрінісін береді.

Бірінші кезеңде тестілеу Windows (WinUI) экожүйесінде жүзеге асырылды. 5-суретте Gemini және Gemma 3 модельдерін Windows платформасында салыстырмалы тестілеу процесінің барысы мен интерфейсі бейнеленген.

Windows платформасындағы нәтижелер 3-кестеде келтірілген. Алынған деректер Gemini 2.5 Flash моделінің көлемді мазмұн генерациясында жоғары өнімділік көрсететінін айқындады. Атап айтқанда, P2 сұранымы бойынша TPS көрсеткіші 228,63-ке жетті. Сонымен бірге TTFT мәні 4560–9235 мс

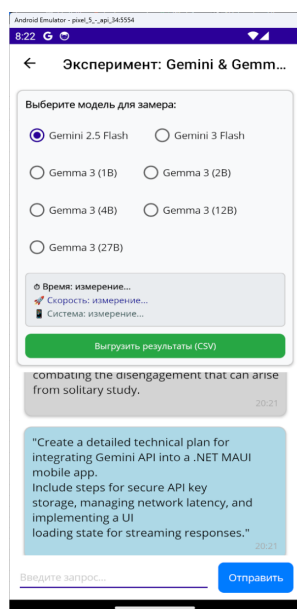
аралығында өзгерді, яғни бұл модель бастапқы жауап беру уақыты бойынша баяулау болғанымен, ұзын және күрделі мазмұнды жылдам құрастыра алады.

Кесте 3 – Windows платформасындағы Gemini 2.5 Flash моделінің өнімділік көрсеткіштері

Platform	ModelName	Prompt Length Chars	TTFT_ms	TotalLatency_ms	Tokens Count	TPS	Ram Used Mb	Network Bytes Delta
WinUI	Gemini 2.5 Flash	54	4560	4719	27	169,81	1,234	27006
WinUI	Gemini 2.5 Flash	103	8387	24999	3798	228,63	0	154080
WinUI	Gemini 2.5 Flash	85	6795	9614	482	170,98	0,254	63959
WinUI	Gemini 2.5 Flash	219	9235	38728	6412	217,41	0	242351

Ескертпе – деректер авторлар жүргізген Windows (WinUI) платформасындағы эмпирикалық жүктемелік тестілеу нәтижелері негізінде алынған.

3-кесте нәтижелері Windows платформасында Gemini 2.5 Flash моделінің ұзын жауаптарды құрастыруда жоғары өткізу қабілетін көрсеткенімен, бастапқы жауап беру уақытының салыстырмалы түрде жоғары екенін аңғартады.



Сурет 6 – Gemini және Gemma 3 модельдерін Android платформасында салыстырмалы тестілеу.

Ескертпе – сурет Android платформасындағы тестілеу процесінің авторлар тіркеген көрінісін береді.

Екінші кезеңде дәл осы метрикалар мобильді ортада өлшенді. 6-суретте Gemini және Gemma 3 модельдерін Android платформасында салыстырмалы тестілеу көрінісі берілген.

Android платформасындағы нәтижелер 4-кестеде берілген. Бұл ортада Gemini 2.5 Flash пен Gemma 3 (1B) модельдері арасында айқын айырмашылық байқалды. Gemma 3 (1B) TTFT көрсеткіші бойынша басымдық танытып, 566–1062 мс аралығында нәтиже көрсетті. Ал Gemini 2.5 Flash үшін осы көрсеткіш 6478–8985 мс аралығында тіркелді. Демек, Gemma 3 (1B) жедел бастапқы жауап қажет болатын интерфейстік сценарийлер үшін қолайлы.

Сонымен қатар TPS көрсеткіші бойынша керісінше жағдай байқалды: Gemini 2.5 Flash 125–231 аралығында нәтиже көрсетсе, Gemma 3 (1B) үшін бұл мән 31–42 аралығында қалды. Бұл бұлттық модельдің көлемді және күрделі мазмұн генерациясы үшін тиімді екенін көрсетеді.

Кесте 4 – Android платформасындағы Gemini 2.5 Flash, Gemma 3(1B) модельдерінің өнімділік көрсеткіштері

Platform	ModelName	Prompt Length Chars	TTFT_ms	TotalLatency_ms	Tokens Count	TPS	Ram Used Mb	Network Bytes Delta
Android	Gemini 2.5 Flash	54	6478	6662	23	125.00	32.078	20156
Android	Gemini 2.5 Flash	104	8985	26178	3975	231.20	1.750	96263
Android	Gemini 2.5 Flash	85	8253	11442	572	179.37	1.000	18879
Android	Gemini 2.5 Flash	221	8566	35760	5596	205.78	1.176	136260
Android	Gemma 3 (1B)	54	566	1594	44	42.80	0.344	32095
Android	Gemma 3 (1B)	104	935	51384	1595	31.62	0.887	962017
Android	Gemma 3 (1B)	85	899	18378	634	36.27	1.375	387057
Android	Gemma 3 (1B)	221	1062	42621	1566	37.68	0.609	950274

Ескертпе – деректер авторлар жүргізген Android платформасындағы эмпирикалық жүктемелік тестілеу нәтижелері негізінде алынған.

4-кесте нәтижелері Android платформасында Gemma 3 (1B) моделінің бастапқы жауап беру жылдамдығы бойынша басымдығын, ал Gemini 2.5 Flash моделінің жалпы генерация өнімділігі бойынша артықшылығын айқындайды. Gemma 3 (1B) үшін тіркелген желілік трафик көлемі салыстырмалы түрде жоғары болды. Бұл көрсеткіш толық бұлттық инференспен емес, тестілеу ортасының сервистік алмасуларымен, модельді инициализациялауымен немесе аралық қызметтік қабаттардың жұмысымен байланысты болуы мүмкін. Сондықтан NetworkBytesDelta көрсеткіші модельдің ішкі ерекшелігін ғана емес, қолданылған архитектуралық ортаның да әсерін қамтиды.

Екінші жағынан, егер жүйе көлемді мәтіндерді, кодтарды немесе күрделі аналитикалық жоспарларды (мысалы, 2-кестедегі P4 промпты) генерациялауды қажет етсе, Gemini 2.5 Flash моделінің жоғары өткізу қабілеті (TPS > 200) бастапқы кешігуді (TTFT) толықтай ақтайды.

Жүргізілген эксперимент нәтижелері .NET MAUI платформасында жасанды интеллект қызметтерін интеграциялаудың тиімділігін дәлелдейді. Модельді таңдау жобаның нақты талаптарына негізделуі тиіс: максималды жауап беру жылдамдығы қажет болса – Gemma 3 (1B), ал үлкен көлемдегі деректерді жылдам өңдеу қажет болса – Gemini 2.5 Flash моделін пайдалану оңтайлы шешім болып табылады. Келесі зерттеу жұмыстарында Gemma 3 моделінің желілік трафикті тұтыну ерекшеліктерін оңтайландыру мәселесін қарастыру ұсынылады.

Осылайша, құрылған қосымша жай ғана бағдарламалық артефакт емес, сонымен қатар озық ЖИ-сервистерін өрістетуге арналған қабықша ретінде .NET MAUI-ді пайдалану тиімділігінің эмпирикалық дәлелі болып табылады. Ол ұсынылған технологиялық стек күрделі коммуникативтік және когнитивтік тапсырмаларды шешу үшін Gemini мүмкіндіктерін біріктіретін, өнімді, қауіпсіз және интуитивті түсінікті кросс-платформалық шешімдерді құруды қамтамасыз ете алатынын растайды.

Нәтижелерді интерпретациялау. Алынған нәтижелер модельдердің қолдану сценарийлері әртүрлі екенін көрсетті. Gemma 3 (1B) төмен TTFT мәні арқылы жедел жауап беруді талап ететін чат-кеңесші, қысқа сұраныстар және локальды өңдеу жағдайларында артықшылық көрсетті. Ал Gemini 2.5 Flash жоғары TPS мәні арқылы ұзын түсіндірме мәтіндер, код генерациясы және күрделі аналитикалық мазмұн құрастыруда тиімді болды.

Білім беру қосымшалары үшін бұл айырмашылық практикалық мәнге ие. Жылдам бастапқы жауап студентпен интерактивті диалог орнатуда маңызды болса, жоғары өткізу қабілеті оқыту материалдарын кеңейтілген түрде құруда маңызды. Осыған сәйкес гибриді архитектура нақты міндетке қарай модельді бағыттауға мүмкіндік береді: жедел жауап пен локальды құпиялылық қажет болғанда жеңілдетілген модельді, ал көлемді мазмұн мен күрделі өңдеу қажет болғанда бұлттық модельді қолдану орынды. Осы интерпретацияны қорытындылау үшін 5-кесте ұсынылады.

Кесте 5 – Gemini 2.5 Flash пен Gemma 3 (1B) модельдерінің салыстырмалы техникалық профилі

Критерий	Gemini 2.5 Flash	Gemma 3 (1B)	Практикалық қорытынды
TTFT	Жоғарырақ	Төмен	Жедел бастапқы жауап үшін Gemma 3 (1B) тиімді
TPS	Жоғары	Орташа	Ұзын және күрделі мазмұн генерациясы үшін Gemini 2.5 Flash тиімді

Жедел жадты пайдалану	Тұрақты	Тұрақты	Екі модель де сынақ ортасында жұмысқа қабілетті
Желілік тәуелділік	Бұлттық инфрақұрылымға тәуелді	Тестілеу ортасының конфигурациясына тәуелді	Қолдану ортасына қарай қосымша оңтайландыру қажет
Ұсынылатын қолдану сценарийі	Күрделі аналитика, код, көлемді мазмұн	Жылдам чат-жауап, локальды өңдеу	Гибридті бағыттау ең орынды шешім

Ескертпе – кесте 3 және 4-кестелердегі эмпирикалық нәтижелерді жинақтау негізінде авторлар тарапынан құрастырылды.

Талқылау. Қолданыстағы әдебиеттерде көрсетілгендей, мультимодальды модельдер жалпы пайымдау мен мәтіндік қолдауда жоғары әлеует танытқанымен, визуалды дәлдікке тәуелді жоғары жауапкершілікті міндеттерде тұрақты сенімділік әрдайым қамтамасыз етілмейді (Fu et al., 2023; Lee, Latif, et al., 2023). Ұсынылған гибридті архитектураның эмпирикалық нәтижелері .NET MAUI негізіндегі білім беру қосымшаларында мультимодальды модельдерді бірыңғай тәсілмен емес, міндет сипатына қарай бейімдеп қолданудың тиімді екенін көрсетті. Зерттеу барысында Gemini 2.5 Flash пен Gemma 3 (1B) модельдерінің техникалық артықшылықтары әртүрлі екені анықталды: Gemma 3 (1B) төмен TTFT мәні арқылы жедел интерфейстік әрекетке қолайлы болса, Gemini 2.5 Flash жоғары TPS көрсеткіші арқылы көлемді және күрделі мазмұн генерациясында басымдық танытты. Бұл айырмашылық білім беру қосымшаларында модель таңдауының тек жалпы қабілетке емес, нақты қолдану сценарийіне негізделуі тиіс екенін дәлелдейді.

Осы тұрғыдан алғанда, .NET MAUI ортасында іске асырылған гибридті архитектура техникалық жағынан орынды шешім болып табылады. Кросс-платформалық орта қосымшаны бірнеше операциялық жүйеде бір кодтық база негізінде пайдалануға мүмкіндік берсе, гибридті бағыттау стратегиясы жылдамдық, өнімділік және құпиялылық арасындағы практикалық теңгерімді сақтауға жағдай жасайды. Әсіресе локальды немесе шеткі өңдеуді қажет ететін міндеттер үшін жеңілдетілген модельдерді қолдану пайдаланушы деректерін қорғау тұрғысынан да маңызды артықшылық береді.

Сонымен бірге зерттеу нәтижелері мультимодальды модельдерді білім беруде қолданудың шектеулерін де растайды. Сондықтан білім беру ортасында мұндай жүйелерді толық автономды бағалаушы ретінде емес, оқыту мен кеңес беруді қолдайтын интеллектуалды көмекші ретінде қарастыру орынды.

Бұл қорытынды академиялық адалдық пен жауапты қолдану мәселелерін де алдыңғы орынға шығарады. Мультимодальды модельдер мәтін құруды, мазмұнды құрылымдауды және идея генерациясын жеңілдеткенімен, олардың бақылаусыз қолданылуы плагиат, мазмұнның бірізденуі және студенттің сыни ойлау белсенділігінің төмендеуі сияқты тәуекелдерге әкелуі мүмкін (Aktay et al., 2023; Perera and Lankathilaka, 2023). Осыған байланысты білім беру мекемелерінде human-in-the-loop қағидатын сақтау, яғни AI нәтижелерін адам тарапынан міндетті түрде тексеру, негізгі талаптардың бірі болуы тиіс.

Тағы бір маңызды мәселе – институционалдық саясаттың жеткіліктілігі. Қазіргі көптеген академиялық ережелер мәтіндік генерациямен байланысты тәуекелдерге көбірек бағытталғанымен, мультимодальды модельдердің сурет, диаграмма, код және аралас мазмұнмен жұмыс істеу ерекшеліктерін толық қамтымайды. Сондықтан саясатты тек мәтіндік плагиатпен шектемей, мультимодальды мазмұнды қолдану, авторлықты көрсету, деректер құпиялылығы және алгоритмдік біржақтылық сияқты мәселелерді қамтитындай кеңейту қажет.

Жалпы алғанда, зерттеу нәтижелері білім беру қосымшаларында Gemini AI-ді интеграциялау мәселесі тек техникалық тиімділікпен шектелмейтінін көрсетті. Оның табысты енгізілуі үшін архитектуралық шешім, қолдану сценарийіне сәйкес модель таңдау, академиялық адалдық саясаты және пайдаланушылардың AI сауаттылығы біртұтас қарастырылуы керек. Осы мәселелерді жүйелеу үшін 6-кестеде білім беру қосымшаларында мультимодальды модельдерді орналастыруға қатысты негізгі этикалық және саяси қиындықтар жинақталған.

Кесте 6 – Білім беру қосымшаларында мультимодальды модельдерді орналастырудың этикалық және саяси негіздемесі

Этикалық/ Саяси қиындық	LMM мүмкіндігі мен тәуекелі	Білім беруге әсері	Міндетті саяси шара
Академиялық адалдық және түпнұсқалық	Мазмұнды генера- циялау және күр- делі парафразалау; мәтінді анықтау- дан жалтару.	Плагиат қаупінің жоғары болуы және сыни, түпнұсқа ойлау дағдыларының төмендеуі.	Бағалауды қайта жобалау (үдеріске/көзбе-көз қатысуға назар аудару); жаппай ЖИ са- уаттылығын үйрету; міндетті авторлықты көрсету нұсқау- лықтары.
Мульти- модальды алгоритмдік біржақтылық (Bias)	Өртүрлі деректер жынытығында оқыту (Gemini табиғи мультимодальды).	Әсіресе диаграммалар, қолжазба немесе аз ресурсты тілдік медиа үшін әділетсіз/ бұрмаланған авто- матты бағалау қаупі.	ЖИ шешім қабылдауында ашықтықты міндеттеу; VQA жүйелерінде қатаң біржақтылық талдауы.
Деректер құпиялылығы және қауіпсіздік	Сезімтал студент деректерін/ суреттерін (PII) өңдеу; Бұлттық API-ге тәуелділік.	Академиялық жазбалар мен мінез- құлық деректерінің ашық болуы.	Деректерді қорғаудың қатаң шаралары; бұлттық тасымалдау тәуекелдерін азайту үшін құрылғыда өңдеуге (Gemini Nano) баса назар аудару.
Шектен тыс тәуелділік және галлюцинация	Жылдам, сенімді, бірақ ықтимал қате немесе ойдан шығарылған жауаптарды құру.	Тәуелсіз мәселелерді шешу дағдыларын жоғалту; академиялық жұмыста жалған ақпарат тарату.	Тексеру мен дереккөзге күмәнмен қарауға ықпал ететін педагогикалық өзгеріс; оқытушылар LMM шектеу- лерін және дереккөзді көрсет- уді үйретуі керек.

Ескертпе – кесте авторлар жүргізген талдау және әдебиеттерде көрсетілген тәуекелдер мен басқару шараларын жүйелеу негізінде құрастырылды.

Білім беру контекстінде LMM-дерді орналастыру қазіргі уақытта ішкі технологиялық шектеулермен шектелген. Зерттеу іргелі модель мәселелерінің сақталуын, әсіресе галлюцинация феноменін және Gemini Pro-ның VQA өнімділігі кезінде байқалған мультимодальды транскрипция қателеріне тән осалдықты мойындайды (Lee, Latif, et al., 2023). Бұл шектеулер автоматтандырылған жоғары жауапкершілікті қосымшалардың сенімділігін төмендетеді.

Болашақ зерттеулер бірнеше бағытта жалғасуы тиіс. Біріншіден, .NET MAUI ортасында қолдану сценарийіне қарай модельдерді динамикалық түрде бағыттауын интеллектуалды басқару қабатын жетілдіру қажет. Екіншіден, білім беру контекстінде AI құралдарын ұзақ мерзімді қолданудың студенттердің когнитивтік дамуына, дербестігіне және оқу стратегияларына әсерін эмпирикалық зерттеу маңызды. Үшіншіден, мультимодальды модельдердің визуалды дәлдікке тәуелді тапсырмалардағы тұрақтылығын арттыру бағытындағы техникалық зерттеулер жалғасуы тиіс. Төртіншіден, білім беру мекемелері, әзірлеушілер және реттеуші құрылымдар арасында жауапты қолдануға арналған бірыңғай институционалдық ұсынымдар жүйесін қалыптастыру қажет.

Зерттеудің шектеулері ретінде бірнеше жағдайды атап өткен жөн. Біріншіден, эмпирикалық бағалау пилоттық техникалық тестілеу форматында жүргізілді және шектеулі сценарийлер санын қамтыды. Екіншіден, зерттеу білім беру контекстіндегі пайдаланушылық тәжірибені немесе педагогикалық нәтижелілікті емес, негізінен техникалық өнімділікті бағалауға бағытталды. Үшіншіден, алынған желілік трафик көрсеткіштері тек модель қасиеттерін ғана емес, қолданылған инфрақұрылымдық ортаның әсерін де қамтуы мүмкін.

Қорытынды. Бұл зерттеу .NET MAUI кросс-платформалық фреймворкі мен Google Gen AI SDK негізінде білім беруге арналған Gemini AI интеграциясының техникалық орындылығын эмпирикалық түрде растады. Зерттеу аясында Windows және Android платформаларына арналған «зияткерлік кеңесші» қосымшасының жұмыс істейтін прототипі әзірленіп, бұлттық және локальды өңдеуді үйлестіретін гибриді архитектура ұсынылды.

Жүргізілген жүктемелік тестілеу нәтижелері Gemini 2.5 Flash пен Gemma 3 (1B) модельдерінің қолдану сценарийлері әртүрлі екенін көрсетті. Gemma 3 (1B) төмен TTFT көрсеткіші арқылы жедел бастапқы жауапты қажет ететін интерфейстік және локальды өңдеу міндеттерінде тиімді болды. Ал Gemini 2.5 Flash жоғары TPS көрсеткіші арқылы көлемді мәтіндер, код және күрделі аналитикалық мазмұн генерациясында басымдық танытты. Бұл нәтижелер білім беру қосымшаларында бірыңғай модельді емес, міндет сипатына қарай бағытталатын гибриді архитектураны қолдану орынды екенін дәлелдейді.

Сонымен қатар зерттеу техникалық тиімділік пен білім беру тұрғысындағы жауапты қолдану мәселелерінің өзара тығыз байланысты екенін көрсетті. Мультимодальды модельдер оқу қолдауын күшейте алғанымен, жоғары жауапкершілікті бағалау, визуалды интерпретация және автоматтандырылған

шешім қабылдау міндеттерінде олардың шектеулері сақталады. Осыған байланысты мұндай жүйелерді толық автономды бағалау құралы ретінде емес, адам бақылауымен жұмыс істейтін интеллектуалды көмекші ретінде қарастырған дұрыс.

Зерттеу нәтижелері негізінде бірнеше практикалық қорытынды ұсынылады. Біріншіден, білім беру мекемелерінде human-in-the-loop қағидаты міндетті түрде сақталуы тиіс. Екіншіден, студент деректерінің құпиялылығын күшейту үшін локальды өңдеуге басымдық беретін гибриді архитектураны пайдалану орынды. Үшіншіден, мультимодальды AI құралдарын енгізу академиялық адалдық, авторлықты көрсету және жауапты қолдану ережелерімен қатар жүруі керек. Болашақ зерттеулерде модельдерді динамикалық бағыттау тетіктерін жетілдіру, локальды модельдердің желілік және ресурстық сипаттарын оңтайландыру, сондай-ақ білім беру контексіндегі ұзақ мерзімді педагогикалық әсерді эмпирикалық тұрғыда бағалау маңызды бағыт болып қала береді.

References

- Aktay S., Gök S., and Uzunoğlu D. (2023) ChatGPT in education. *Türk Akademik Yayınlar Dergisi (TAY Journal)*, 7(2). — P. 378–406. <https://doi.org/10.29329/tayjournal.2023.543.03> (in Eng.)
- Akter S.N., Yu Z., Muhamed A., Ou T., Bäuerle A., Cabrera Á.A., Dholakia K., Xiong C., and Neubig G. (2023) *An in-depth look at Gemini's language abilities*. arXiv. <https://doi.org/10.48550/arXiv.2312.11444> (in Eng.)
- Ali A.H., Alajanbi M., Yaseen M.G., and Abed S.A. (2023) Chatgpt4, DALL·E, Bard, Claude, BERT: Open possibilities. *Babylonian Journal of Machine Learning*, 2023. — P. 17–18. <https://doi.org/10.58496/BJML/2023/003> (in Eng.)
- Azmi K., and Fithriani R. (2025) Students' perceptions of Gemini AI effectiveness in academic writing. *IDEAS: Journal of Language Teaching and Learning, Linguistics and Literature*, 13(1). — P. 74–93. <https://doi.org/10.24256/ideas.v13i1.6278> (in Eng.)
- Badshah A., Rehman G.U., Farman H., Ghani A., Sultan S., Zubair M., and Nasralla M.M. (2023) Transforming educational institutions: Harnessing the power of Internet of Things, Cloud, and Fog computing. *Future Internet*, 15. — 367 p. <https://doi.org/10.3390/fi15110367> (in Eng.)
- Fu C., Zhang R., Wang Z., Huang Y., Zhang Z., Qiu L., Ye G., Shen Y., Zhang M., Chen P., Zhao S., Lin S., Jiang D., Yin D., Gao P., Li K., Li H., and Sun X. (2023) *A challenger to GPT-4V? Early explorations of Gemini in visual expertise*. arXiv. <https://doi.org/10.48550/arXiv.2312.12436> (in Eng.)
- Gemini Team, Google. (2025) *Gemini: A family of highly capable multimodal models*. arXiv. <https://doi.org/10.48550/arXiv.2312.11805> (in Eng.)
- Imran M., and Almusharraf N. (2023) Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605> (in Eng.)
- Imran M., and Almusharraf N. (2024a) Digital learning demand and applicability of Quality 4.0 for future education: A systematic review. *International Journal of Engineering Pedagogy (iJEP)*, 14(4). — P. 38–53. <https://doi.org/10.3991/ijep.v14i4.48847> (in Eng.)
- Imran M., and Almusharraf N. (2024b) Google Gemini as a next generation AI educational tool: A review of emerging educational technology. *Smart Learning Environments*, 11(22). <https://doi.org/10.1186/s40561-024-00310-z> (in Eng.)
- Koubaa A., Boulila W., Ghouti L., Alzahem A., and Latif S. (2023) Exploring ChatGPT capabilities and limitations: A survey. *IEEE Access*, 11. — P. 118698–118721. <https://doi.org/10.1109/ACCESS.2023.3326474> (in Eng.)

Lee G.-G., Latif E., Shi L., and Zhai X. (2023) *Gemini Pro defeated by GPT-4V: Evidence from education*. arXiv. <https://doi.org/10.48550/arXiv.2401.08660> (in Eng.)

Lee G.-G., Shi L., Latif E., Gao Y., Bewersdorf A., Nyaaba M., Guo S., Wu Z., Liu Z., Wang H., Mai G., Liu T., and Zhai X. (2023) *Multimodality of AI for education: Towards Artificial General Intelligence*. arXiv. <https://doi.org/10.48550/arXiv.2312.06037> (in Eng.)

Perera P., and Lankathilaka M. (2023) Preparing to revolutionize education with the multi-model GenAI tool Google Gemini? A journey towards effective policy making. *Journal of Advances in Education and Philosophy*, 7(8). — P. 246–253. <https://doi.org/10.36348/jaep.2023.v07i08.001> (in Eng.)

Jangassiyev R., Umarova Zh., Ussenova A., Makhanova Z., Zhumatayev N., Amirov M., and Koishibekova G. Comparative analysis of cross-platform development methodologies: a comprehensive study. // *TELKOMNIKA Telecommunication Computing Electronics and Control*, 2025, 23(1). — P. 108–118. doi: <https://doi.org/10.12928/TELKOMNIKA.v23i1.26331> (in Eng.)

Kotmungkun S., Chompurach W., and Thaksanan P. (2024) OpenAI ChatGPT vs Google Gemini: A study of AI chatbots' writing quality evaluation and plagiarism checking. *English Language Teaching Educational Journal*, 7(2). — P. 90–108. <https://doi.org/10.12928/eltej.v7i2.11572> (in Eng.)

M Arli Rusandi, Ahman, Ipah Saripah, Deasy Yunika Khairun, Mutmainnah, No worries with ChatGPT: building bridges between artificial intelligence and education with critical thinking soft skills, *Journal of Public Health*, Volume 45, Issue 3, September 2023, Pages e602–e603, <https://doi.org/10.1093/pubmed/fdad049> (in Eng.)

Blanco J.Z., and Lucrédio D.A holistic approach for cross-platform software development. *Journal of Systems and Software*, 2021, 179, 110985 6. doi: <https://doi.org/10.1016/j.jss.2021.110985> (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 166–179

<https://doi.org/10.32014/2026.2518-1726.432>

IRSTI 49.03.03

UDC 654.165

© **Doszhan N.S.**^{1,2}, **Sultanbekova L.Ye.**², **Zhumagali S.Zh.**²,
Konysbayev E.K.², 2026.

¹ Al-Farabi Kazakh National University, Almaty, Kazakhstan;

² Institute of space technique and technology, Almaty, Kazakhstan;

E-mail: sultanbekova.l@istt.kz

MODELING AND PARAMETER CALCULATION OF AN EMERGENCY RESPONSE SYSTEM BASED ON LORAWAN TECHNOLOGY IN THE HIGH-ALTITUDE CONDITIONS OF THE ZAILIYSKY ALATAU

Doszhan Nursultan — PhD, Project Supervisor, Deputy Dean for Research, Innovation, and International Relations, Al-Farabi Kazakh National University; Institute of space technique and technology, Almaty, Kazakhstan,

E-mail: nursultan.sagynaiuly@kaznu.kz, <https://orcid.org/0000-0001-6178-8389>;

Sultanbekova Leila — Deputy Director for Advanced Projects, Institute of space technique and technology, Almaty, Kazakhstan,

E-mail: sultanbekova.l@istt.kz, <https://orcid.org/0009-0002-1651-616X>;

Zhumagali Sabyrzhan — Chief Project Engineer, Institute of space technique and technology, Almaty, Kazakhstan,

E-mail: zhumagali.s@istt.kz, <https://orcid.org/0009-0007-5995-4377>;

Konysbayev Erkegali — Category II Engineer, Institute of space technique and technology, Almaty, Kazakhstan,

E-mail: konysbaev.e@istt.kz, <https://orcid.org/0009-0009-1880-1838>.

Abstract. This paper presents an emergency response and monitoring system based on the LoRaWAN network architecture, designed for deployment in high-altitude regions such as the routes of the Zailiysky Alatau. The main problem addressed by this system is the difficulty of conducting rescue operations due to the lack of communication in deep canyons and gorges, where conventional mobile networks are ineffective. Within the study, numerical modeling of a network consisting of three base stations spaced 10 km apart was performed using the RadioPlanner 3.0 software environment. In addition, the energy parameters of the radio channel were calculated, with total losses ranging from 126.21 to 131.21 dB. The radius of the first Fresnel zone was also determined to be 29.42 m. The obtained results indicate that the heterogeneity of coverage areas and the

presence of weak signal zones are caused by terrain-induced physical shielding and obstruction of the Fresnel zone. A distinctive feature of this study is the use of the Longley–Rice propagation model and the consideration of diffraction losses (up to 20 dB), which made it possible to confirm a link margin of up to 23.79 dB and demonstrate the feasibility of establishing reliable communication under radio shadow conditions. The results can be applied to the deployment of emergency response systems in mountainous environments for monitoring tourist conditions and coordinating search-and-rescue operations in extreme high-altitude areas.

Keywords: LoRaWAN, high-altitude regions, Fresnel zone, Longley–Rice model, line-of-sight, emergency response system, tourists, communication

Financing. *This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP26103091 “Development of an emergency response system to ensure the safety of tourists on mountain routes”)*

For citations: Doszhan N.S.^{1,2}, Sultanbekova L.Ye., Zhumagali S.Zh., Konysbayev E.K. Modeling and parameter calculation of an emergency response system based on LoRaWAN technology in the high-altitude conditions of the Zailiysky Alatau. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 166–179. DOI <https://doi.org/10.32014/2026.2518-1726.432>

© Досжан Н.С.^{1,2}, Султанбекова Л.Е.², Жумағали С.Ж.²,
Конысбаев Е.К.², 2026.

¹ Өл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

² «Ғарыштық техника және технологиялар институты» ЖШС,
Алматы, Қазақстан.

E-mail: sultanbekova.l@istt.kz

ІЛЕ АЛАТАУЫНЫҢ БИІК ТАУЛЫ ЖАҒДАЙЫНДА LORAWAN ТЕХНОЛОГИЯСЫ НЕГІЗІНДЕГІ ЖЕДЕЛ ӘРЕКЕТ ЕТУ ЖҮЙЕСІНІҢ ПАРАМЕТРЛЕРІН МОДЕЛЬДЕУ ЖӘНЕ ЕСЕПТЕУ

Досжан Нұрсұлтан — PhD, жобаның ғылыми жетекшісі, ғылыми-инновациялық қызмет және халықаралық байланыстар жөніндегі деканның орынбасары, Өл-Фараби атындағы Қазақ ұлттық университеті; «Ғарыштық техника және технологиялар институты» ЖШС, Алматы, Қазақстан,

E-mail: nursultan.sagynaiuly@kaznu.kz, <https://orcid.org/0000-0001-6178-8389>;

Султанбекова Лейла — директордың перспективалық жобалар жөніндегі орынбасары, Ғарыштық техника және технологиялар институты, Алматы, Қазақстан,

E-mail: sultanbekova.l@istt.kz, <https://orcid.org/0009-0002-1651-616X>;

Жұмағали Сабыржан — жобаның бас инженері, Ғарыштық техника және технологиялар институты, Алматы, Қазақстан,

E-mail: zhumagali.s@istt.kz, <https://orcid.org/0009-0007-5995-4377>;

Қонысбаев Еркеғали — 2-санатты инженер, Ғарыштық техника және технологиялар институты, Алматы, Қазақстан,

E-mail: konysbaev.e@istt.kz, <https://orcid.org/0009-0009-1880-1838>.

Аннотация. Осы мақалада LoRaWAN желі архитектурасына негізделген және Іле Алатауы бағыттары сияқты биік таулы аймақтарда қолдануға арналған жедел әрекет ету және мониторинг жүйесі қарастырылады. Ұсынылған жүйенің шешетін негізгі мәселе – терең шатқалдар мен аңғарларда байланыстың болмауына байланысты құтқару операцияларын жүргізу, себебі мұндай жерлерде дәстүрлі ұялы байланыс тиімсіз жұмыс істейді. Зерттеу аясында RadioPlanner 3.0 бағдарламалық ортасында арақашықтығы 10 км болатын 3 базалық станциядан тұратын желінің сандық модельдеуі жүргізілді. Сонымен қатар радиоканалдың энергетикалық параметрлері есептеліп, жиынтық шығындар 126,21–131,21 дБ аралығында анықталды. Бірінші Френель аймағының радиусы 29,42 м екені есептелді. Алынған нәтижелер жабын аймақтарының біркелкі еместігі мен әлсіреген сигнал аймақтарының пайда болуы жер бедерінің физикалық экрандау әсерімен және Френель аймағының жабылуымен байланысты екенін көрсетті. Зерттеудің ерекшелігі – Longley–Rice моделі қолданылуы және дифракциялық шығындардың (20 дБ-ге дейін) ескерілуі болып табылады. Бұл байланыс арнасы қорының 23,79 дБ-ге дейін жететінін растауға және радиокөлеңке жағдайында тұрақты байланысты ұйымдастыру мүмкіндігін дәлелдеуге мүмкіндік берді. Зерттеу нәтижелері биік таулы аймақтарда туристердің жағдайын бақылау және іздестіру-құтқару жұмыстарын үйлестіру үшін жедел әрекет ету жүйелерін орналастыруда қолданылуы мүмкін.

Түйін сөздер: LoRaWAN, биік таулы аймақтар, Френель аймағы, Longley–Rice моделі, радиокөріну, жедел әрекет ету жүйесі, туристер, байланыс

© **Досжан Н.С.^{1,2}, Султанбекова Л.Е.², Жумағали С.Ж.²,
Қонысбаев Е.К.², 2026.**

¹ Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан;

² Институт космической техники и технологий, Алматы, Казахстан;
E-mail: sultanbekova.l@istt.kz

МОДЕЛИРОВАНИЕ И РАСЧЕТ ПАРАМЕТРОВ СИСТЕМЫ ЭКСТРЕННОГО РЕАГИРОВАНИЯ НА БАЗЕ ТЕХНОЛОГИИ LORAWAN В УСЛОВИЯХ ВЫСОКОГОРЬЯ ЗАИЛИЙСКОГО АЛАТАУ

Досжан Нұрсұлтан — PhD, научный руководитель проекта, заместитель декана по научно-инновационной деятельности и международным связям Казахского национального университета имени аль-Фараби; Институт космической техники и технологий, Алматы, Казахстан,

E-mail: nursultan.sagynaiuly@kaznu.kz, <https://orcid.org/0000-0001-6178-8389>;

Султанбекова Лейла — заместитель директора по перспективным проектам, Институт космической техники и технологий, Алматы, Казахстан,

E-mail: sultanbekova.l@istt.kz, <https://orcid.org/0009-0002-1651-616X>;

Жұмағали Сабыржан — главный инженер проекта, Институт космической техники и технологий, Алматы, Казахстан,

E-mail: zhumagali.s@istt.kz, <https://orcid.org/0009-0007-5995-4377>;

Конысбаев Еркеғали — инженер 2-й категории, Институт космической техники и технологий, Алматы, Казахстан,

E-mail: konysbaev.e@istt.kz, <https://orcid.org/0009-0009-1880-1838>.

Аннотация. *Актуальность.* В статье рассматривается система экстренного реагирования и мониторинга, основанная на архитектуре сети LoRaWAN и предназначенная для использования в высокогорных районах, включая маршруты Заилийского Алатау. Актуальность исследования обусловлена сложностью проведения спасательных операций в глубоких каньонах и ущельях, где стандартная мобильная связь является неэффективной или недоступной. В таких условиях требуется создание устойчивой низкоэнергетической сети связи для мониторинга состояния туристов и координации поисково-спасательных работ. *Цель.* Смоделировать и рассчитать параметры системы экстренного реагирования на базе технологии LoRaWAN для оценки возможности организации устойчивой связи в условиях высокогорного рельефа Заилийского Алатау. *Методы.* В рамках исследования проведено численное моделирование сети из трех базовых станций, расположенных с шагом 10 км, в программной среде RadioPlanner 3.0. Рассчитаны энергетические параметры радиоканала с суммарными потерями от 126,21 до 131,21 дБ. Также определен радиус первой зоны Френеля, равный 29,42 м. Для оценки распространения радиосигнала использована модель Longley–Rice с учетом дифракционных потерь, достигающих 20 дБ. *Результаты и выводы.* Полученные данные показывают, что неоднородность зон покрытия и наличие участков с ослабленным сигналом обусловлены физическим экранированием рельефа и перекрытием зоны Френеля. Использование модели Longley–Rice и учет дифракционных потерь позволили подтвердить наличие запаса по радиоканалу до 23,79 дБ и доказать возможность организации устойчивой связи в условиях радиотени. Результаты исследования могут быть использованы при развертывании систем экстренного реагирования в горных районах, предназначенных для мониторинга состояния туристов и координации поисково-спасательных работ в экстремальных высокогорных условиях.

Ключевые слова: LoRaWAN, высокогорье, зона Френеля, модель Longley–Rice, радиовидимость, система экстренного реагирования, туристы, связь, поисково-спасательные работы

Введение. С 2000 по 2020 год число зарегистрированных в мире природных катастроф увеличилось более чем на 80%, что свидетельствует

о значительном росте экологической нестабильности (Zhan et al., 2025). Примером служат масштабные обрушения ледников и скал, такие как инцидент 2021 года в Индийских Гималаях, которые приводят к мгновенному разрушению инфраструктуры и гибели людей (Wang et al., 2023). Для заблаговременного предупреждения о подобных опасностях в зонах активного туризма используются современные системы мониторинга, например, Xenios, применяющая сетку метеорологических прогнозов с разрешением 1 x 1 км (Psaroudakis et al., 2021). В высокогорных условиях смертность остается критически высокой. Только на Эвересте за последнее столетие зарегистрировано 305 летальных исходов (Sharma et al., 2023). Статистика показывает ежегодный рост числа пропавших без вести альпинистов и несчастных случаев на горных маршрутах (Affrylia et al., 2021). Спасательные операции осложняются высокой скоростью ветра, достигающей в среднем 9,5 м/с, что ограничивает использование вертолетов и стандартных беспилотных аппаратов для оперативной доставки помощи (Wankmüller et al., 2020).

Литературный обзор. Традиционные средства связи часто оказываются недостаточно мобильными и гибкими в условиях пересеченного рельефа (Zhao et al., 2019). Потеря соединения является причиной крушения 52% дронов-спасателей, что требует внедрения отказоустойчивых сетевых архитектур (Ejaz et al., 2019). Полевые испытания показывают, что переносные комплексы управления должны обеспечивать радиус связи 15–20 км для эффективной визуальной диспетчеризации в зонах чрезвычайных ситуаций (Wang et al., 2023).

Интеграция концепции «Интернета экстренных служб» (IoES) позволяет существенно улучшить сбор данных в реальном времени с использованием подключенных устройств (Damaševičius et al., 2023). Логические SAT-решатели способны оперативно, за одну секунду или быстрее, обнаруживать контекстные угрозы для туристов, обеспечивая проактивное реагирование системы (Klimek et al., 2024). Точность ИИ-моделей на базе YOLOv5 при идентификации очагов возгорания и задымления в лесных массивах достигает 98,7% (Avazov et al., 2023).

Сообщения об инцидентах в удаленных районах часто поступают с задержкой в часы или дни, что может оказаться фатальным при медицинских осложнениях (Mundargi et al., 2025). Социальные медиа, например, Twitter, реагируют на новости о бедствиях быстрее традиционных СМИ, что делает их важным инструментом первичной детекции событий для служб безопасности (Aboualola et al., 2023). При этом рискам подвергаются и сами спасатели: зафиксированы случаи гибели трех членов общественных отрядов в течение восьми дней при выполнении одной задачи (Wang et al., 2024).

Таким образом, несмотря на развитие технологий, специфика распространения радиосигнала LoRaWAN в экстремально сложном рельефе Заилийского Алатау остается недостаточно изученной. Отсутствуют точные

расчетные данные о дифракционных потерях и устойчивости энергетического бюджета канала при межстанционных расстояниях в 10 км в условиях глубокого экранирования. Также не определены критические пороги перекрытия зоны Френеля, при которых сохраняется работоспособность сети экстренного мониторинга в горных ущельях. Следовательно, исследования, посвященные разработке систем экстренного реагирования в горных условиях, представляются весьма актуальными.

В данной работе представлено численное моделирование сети из трех БС LoRaWAN. Будут продемонстрированы расчеты энергетических параметров радиоканала, включая определение запаса по мощности и оценку дифракционных потерь. Также будут приведены параметры первой зоны Френеля для трасс протяженностью 10 км и обоснована возможность создания устойчивого «информационного кольца» связи для обеспечения безопасности туристов в зонах радиотени.

Материалы и методы исследования. В исследовании применен комплексный подход, включающий теоретический анализ, численное моделирование и верификацию полученных результатов. Основное внимание уделено оценке функционирования системы экстренного реагирования в сложных условиях горного рельефа. В качестве объекта исследования выбраны горные маршруты Заилийского Алатау, где распространение радиосигнала затруднено из-за перепадов высот, экранирования и зон радиотени.

Теоретическая часть исследования основана на теории распространения радиоволн в пересеченной местности. Для расчета потерь сигнала использована модель Longley–Rice (Hufford et al., 1982), учитывающая профиль рельефа, дифракционные потери и особенности трассы между передатчиком и приемником. Дополнительно проведен анализ первой зоны Френеля для оценки радиовидимости и выявления участков, где препятствия могут существенно ухудшить качество связи.

Численное моделирование выполнено в программной среде RadioPlanner 3.0. В ходе моделирования заданы параметры сети LoRaWAN (таблица 1)

Таблица 1 – Технические характеристики сети LoRaWAN

Параметр	LoRa (868 МГц)
Дальность (прямая видимость)	до 15-20 км
Выходная мощность передатчика	2 Вт (≈ 33 дБм)
Дальность (в ущелье)	10 км (зависит от рельефа)
Энергопотребление (передача SOS)	4–6 Вт
Энергоемкость аккумулятора	≈ 620 Вт·ч
Время автономной работы (1 сигнал в час)	до 20–30 суток
Время автономной работы (непрерывная передача)	≈ 100 –120 часов (4–5 суток)
Пропускная способность	0,3-55 кбит/с
Задержка (латентность)	1-3 секунды
Требования к инфраструктуре	Низкие (1 шлюз на 10-15 км)

Для горного сценария в модель введены цифровые данные о рельефе и профили трасс в ущельях и на склонах. Методика исследования включала построение профилей трасс, определение участков с потенциально неустойчивой связью. Рассмотрены случаи прямой видимости и ситуации, когда сигнал распространяется в условиях частичного или полного экранирования рельефом. Для оценки устойчивости канала связи использован вероятностный подход, учитывающий влияние рельефа, погодных факторов и резервирования каналов передачи. Адекватность подхода подтверждена сопоставлением расчетных данных с ожидаемыми закономерностями распространения радиоволн.

Результаты. Для оценки условий распространения радиосигнала на маршрутах, типичных для Заилийского Алатау, был выполнен расчет радиолинии в горной местности. Целью расчета являлось определение потерь сигнала, запаса по каналу и выявление участков потенциального ухудшения связи при сохранении единой конфигурации сети.

Расчет параметров радиоканала для схемы из 3-х базовых станций (БС) LoRaWAN. При расчетах предполагалось, что расстояние между соседними БС составляет 10 км. Расчет включал в себя определение потерь в свободном пространстве, суммарных потерь на участке между БС и запаса по мощности сигнала.

Потери в свободном пространстве определялись по формуле (ITU-R, P.525-4, P.526-15, 2019):

$$L_{fs} = 20\log_{10}(d) + 20\log_{10}(f) + 32,44,$$

где d – расстояние, км; f – частота, МГц.

В расчетной модели в качестве рабочей частоты используется 868 МГц, стандартная для рассматриваемой сети LoRaWAN. Подстановка этого значения в формулу потерь в свободном пространстве дает следующее выражение:

$$L_{fs} = 20\log_{10}(d) + 91,21.$$

Для расчетов было взято расстояние в 10 км, что соответствует расстоянию между соседними БС LoRaWAN в рассматриваемой области.

$$L_{fs} = 20\log_{10}(10) + 91,21 = 20 + 91,21 = 111,21 \text{ дБ}.$$

Дифракционные потери были учтены при расчете трассы между горными БС. Для участков с выраженным влиянием рельефа использовались два расчетных значения:

$$L_{diff1} = 15 \text{ дБ}, L_{diff2} = 20 \text{ дБ}.$$

Тогда суммарные потери составляют:

$$L_{total1} = 111,21 + 15 = 126,21 \text{ дБ},$$

$$L_{total2} = 111,21 + 20 = 131,21 \text{ дБ}.$$

Максимально допустимые потери канала:

$$L_{max} = P_{tx} - P_{rx} = 20 - (-130) = 150 \text{ дБ}.$$

Запас по каналу для двух вариантов трассы:

$$M_1 = 150 - 126,21 = 23,79 \text{ дБ},$$

$$M_2 = 150 - 131,21 = 18,79 \text{ дБ}.$$

Дифракционные потери на межстанционной трассе определяются профилем рельефа и степенью перекрытия первой зоны Френеля. На участках, типичных для Заилийского Алатау, встречаются как частично экранированные зоны, так и зоны с выраженными перегибами трассы. В расчетах использованы два значения дифракционных потерь для случаев частичного перекрытия и для более сложных условий распространения сигнала. Эти значения были применены при расчете параметров радиоканала на межстанционной трассе протяженностью 10 км. Результаты расчетов представлены в таблице 2.

Таблица 2 – Параметры радиоканала при межстанционном расстоянии 10 км

Параметр	Вариант 1 (15 дБ)	Вариант 2 (20 дБ)
Расстояние между БС, км	10	10
L_{fs} , дБ	111,21	111,21
L_{diff} , дБ	15	20
L_{total} , дБ	126,21	131,21
L_{max} , дБ	150	150
Запас по каналу, дБ	23,79	18,79

Для анализа параметров радиоканала в горной местности было проведено моделирование схемы размещения трех БС LoRaWAN и рассчитаны характеристики межстанционных трасс. Результаты моделирования, в том числе карта зоны покрытия и профили трасс, представлены на рисунке 1.

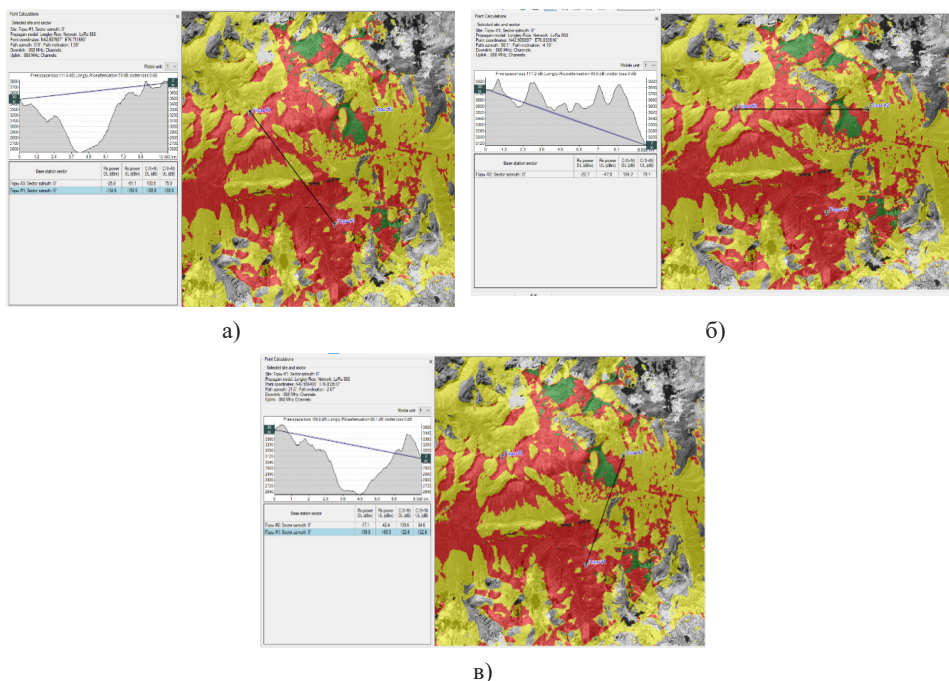


Рисунок 1 – Расчетная схема размещения 3-х БС LoRaWAN в горной местности:
 а – профиль трассы «БС №3 – БС №1»; б – профиль трассы «БС №3 – БС №2»;
 в – профиль трассы «БС №2 – БС №1»

Расчет параметров зоны Френеля для участка между БС. Выполнен расчет для участка трассы между двумя соседними БС, расстояние между которыми составляет 10 км. Для оценки условий распространения сигнала использован расчет первой зоны Френеля.

Радиус первой зоны Френеля определялся по формуле:

$$R_{F1} = 17,32 \sqrt{\frac{d_1 d_2}{f(d_1 + d_2)'}}$$

где d_1 и d_2 – расстояния от препятствия до передатчика и приемника, км;
 f – частота, ГГц.

Для препятствия, расположенного в середине трассы, принимаем:

$$d_1 = d_2 = 5 \text{ км}, f = 0,868 \text{ ГГц.}$$

Следовательно:

$$R_{F1} = 17,32 \sqrt{\frac{5 \cdot 5}{0,868(5 + 5)}} = 17,32 \sqrt{\frac{25}{8,68}} = 17,32 \sqrt{2,88} = 29,42 \text{ м.}$$

Предельно допустимое перекрытие 40%:

$$R_{crit} = 0,4 \cdot 29,42 = 11,77 \text{ м.}$$

Расчет параметров первой зоны Френеля произведен для межстанционной трассы длиной 10 км. При этом предполагалось, что наиболее значимое препятствие находится посередине трассы, то есть на расстоянии 5 км как от передающей, так и от приемной БС. Рассчитанные параметры первой зоны Френеля представлены в таблице 3.

Таблица 3 – Параметры первой зоны Френеля для межстанционной трассы 10 км

Параметр	Значение
Полная длина трассы, км	10
d_1 , км	5
d_2 , км	5
Частота, ГГц	0,868
R_{F1} , м	29,42
40% перекрытия, м	11,77

Расчетный радиус первой зоны Френеля для межстанционной трассы длиной 10 км составляет 29,42 м, а допустимое перекрытие – 11,77 м.

Расчет структуры сети для схемы из трех БС. Базовая конфигурация сети состоит из трех узлов доступа, расположенных в горной местности на расстоянии 10 км друг от друга.

Число основных БС:

$$N_{gw} = 3.$$

Число межстанционных участков в контуре:

$$N_{link} = 3.$$

Суммарная длина межстанционных участков:

$$L_{\Sigma} = 3 \cdot 10 = 30 \text{ км.}$$

Сеть состоит из 3-х БС, соединенных тремя межстанционными участками. Общая протяженность межстанционных трасс составляет 30 км. Подробные параметры структуры сети представлены в таблице 4.

Таблица 4 – Структура сети из 3-х БС

Параметр	Значение
Число БС LoRaWAN	3
Расстояние между соседними БС, км	10

Число межстанционных трасс	3
Суммарная длина межстанционных трасс, км	30
Сетевой сервер	1
Резервный внешний канал	1

Для оценки пространственного распределения сигнала было проведено моделирование зоны покрытия сети LoRaWAN, состоящей из трех БС. Полученная карта покрытия для данной конфигурации представлена на рисунке 2.

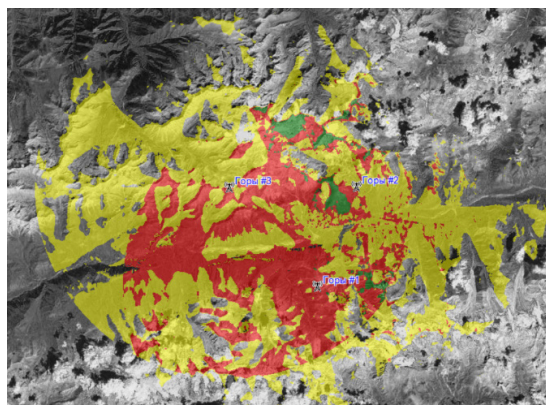


Рисунок 2 – Карта расчетной зоны покрытия для трех БС LoRaWAN

На карте показаны зоны покрытия 3-х БС, расположенных на расстоянии около 10 км друг от друга. Зоны покрытия характеризуются неоднородной пространственной структурой, что обусловлено рельефом местности. Вместо классической круговой формы области уверенного приема приобретают выраженный анизотропный характер. Под влиянием специфики ущелий и склонов Заилийского Алатау они трансформируются в вытянутые структуры, ориентированные вдоль направлений с минимальными препятствиями, что позволяет сигналу распространяться вглубь каньонов даже при значительных дифракционных потерях. В пределах исследуемой территории выделяются участки с различным уровнем сигнала, которые соответствуют расчетным значениям, полученным в программе RadioPlanner 3.0. Общая площадь покрытия формируется за счет объединения зон действия всех трех БС. Перекрывание зон наблюдается в центральной части исследуемого участка. Снижение уровня сигнала зафиксировано в местах перегибов рельефа и на значительном удалении от БС.

Обсуждение. Моделирование сети из трех БС, расположенных с шагом 10 км и работающих на частоте 868 МГц, показало наличие запаса по мощности сигнала в диапазоне от 18,79 до 23,79 дБ. Общие потери в канале связи составляют 126,21–131,21 дБ, включая дифракционные потери (15–20 дБ), обусловленные рельефом Заилийского Алатау. Расчет радиуса

первой зоны Френеля (29,42 м) и предела перекрытия (11,77 м) позволяет определить границы устойчивости сигнала при его распространении через горные хребты. Результаты моделирования подтверждают возможность использования технологии LoRaWAN для организации связи в зонах, где отсутствует покрытие GSM.

Стоит отметить, что несмотря на положительные результаты, у данной модели есть ограничения. Численный эксперимент проводился только для базового сценария с тремя БС. Кроме того, расчеты не учитывали потери сигнала в густой растительности и влияние специфических метеоусловий Заилийского Алатау, что в реальных условиях может привести к уменьшению запаса по каналу. Кроме того, используемая модель Longley–Rice и цифровая модель рельефа местности не принимают во внимание локальные особенности среды, такие как плотность растительности и метеорологические факторы (в частности, средняя скорость ветра 9,5 м/с, ограничивающая возможности спасательной авиации). Карта покрытия сети демонстрирует неоднородность с участками ослабленного сигнала в зонах перегибов рельефа и глубоких каньонах. Следовательно, при моделировании топологии сети LoRaWAN в высокогорье целесообразно отходить от идеализированных моделей в пользу учета продолговатой формы зон покрытия. Ориентация этих областей по конфигурации рельефа позволяет более эффективно использовать расчетный запас по мощности для перекрытия зон радиотени, где стандартная мобильная связь оказывается неэффективной.

Внедрение систем экстренного реагирования в горных условиях в Заилийском Алатау направлено на снижение трудности со связью на туристических маршрутах. Для дальнейшего развития данной системы требуется оптимизировать размещение БС с учетом реальных профилей трасс и увеличить общее количество узлов сети. Перспективным направлением является создание гибридной архитектуры, интегрирующей LoRaWAN с низкоорбитальными спутниковыми системами, что позволит обеспечить передачу данных на участках с полным экранированием сигнала.

Заключение. Проведенное исследование подтвердило гипотезу о создании устойчивой системы связи в сложном горном рельефе. Энергоэффективные сети дальнего радиуса действия эффективно преодолевают препятствия, создаваемые горными хребтами и глубокими ущельями. Архитектура данной системы обеспечивает сохранение сигнала даже при частичном перекрытии трассы, что делает технологию пригодной для работы в зонах, где отсутствует стандартное покрытие.

Научная значимость результатов заключается в обосновании параметров распространения радиоволн в условиях радиотени. Выявленные закономерности позволяют точно прогнозировать поведение радиоканала и учитывать влияние дифракции на пересеченной местности. Это формирует теоретическую базу для проектирования надежных систем экстренного оповещения, работающих в экстремальных условиях.

Практическая реализация предложенных решений направлена на развитие инфраструктуры для систем экстренного реагирования в горных условиях. Развертывание такой сети обеспечит постоянный мониторинг состояния туристов и оперативную передачу сигналов бедствия. Независимость системы от ограничений, накладываемых погодными факторами на спасательную авиацию, критически важна для своевременного оказания помощи.

Дальнейшее развитие темы предполагает изучение влияния экстремальных метеорологических нагрузок и низких температур на стабильность оборудования. Основное внимание будет уделено оптимизации расположения узлов для устранения неоднородностей в зонах покрытия. Перспективным подходом к такой оптимизации станет уточнение геометрии зон радиовидимости и их представление в виде вытянутых областей, что более корректно отражает физику распространения сигнала в экстремально сложном рельефе и повышает точность прогнозирования устойчивости системы экстренного мониторинга.

References

Aboualola M., Abualsaud K., Khattab T., Zorba N., & Hassanein H.S. (2023) Edge technologies for disaster management: A survey of social media and artificial intelligence integration. *IEEE Access*, 11. — P. 73782–73802. (in Eng.)

Affrylia G., Fadhli M., & Lindawati L. (2021) Perancangan emergency button untuk pendaki gunung dengan sistem komunikasi multihop berbasis LoRa. *PROtek: Jurnal Ilmiah Teknik Elektro*, 8(2). — P. 96. (in Eng.)

Avazov K., Hyun A.E., Sami S.A., Khaitov A., Abdusalomov A.B., & Cho Y.I. (2023) Forest fire detection and notification method based on AI and IoT approaches. *Future Internet*, 15(2). — P. 61. (in Eng.)

Damaševičius R., Bacanin N., & Misra S. (2023) From sensors to safety: Internet of Emergency Services (IoES) for emergency response and disaster management. *Journal of Sensor and Actuator Networks*, 12(3). — P. 41. (in Eng.)

Ejaz W., Azam M.A., Saadat S., Iqbal F., & Hanan A. (2019) Unmanned aerial vehicles enabled IoT platform for disaster management. *Energies*, 12(14). — P. 2706. (in Eng.)

Hufford G.A., Longley A.G., & Kissick W.A. (1982) A guide to the use of the ITS Irregular Terrain Model in the area prediction mode (NTIA Report 82-100). U.S. Department of Commerce, National Telecommunications and Information Administration. Available at: https://its.ntia.gov/publications/download/82-100_ocr.pdf (in Eng.)

International Telecommunication Union Radiocommunication Sector (ITU-R). (2019) Recommendation ITU-R P.525-4: Calculation of free-space attenuation. Available at: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.525-4-201908-S%21%21PDF-E.pdf (in Eng.)

International Telecommunication Union Radiocommunication Sector (ITU-R). (2019) Recommendation ITU-R P.526-15: Propagation by diffraction. Available at: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.526-15-201910-S!!PDF-E.pdf (in Eng.)

Klimek R. (2024) Ambient-aware continuous aid for mountain rescue activities. *Information Sciences*, 653. — P. 119772. (in Eng.)

Mundargi Prof. (2025) IoT-based advanced monitoring system for outdoor safety. *International Journal for Research in Applied Science and Engineering Technology*, 13. — P. 1089–1093. (in Eng.)

Psaroudakis C., Xanthopoulos G., Stavrakoudis D., Barnias A., Varela V., Gkotsis I., ... & Gitas I.Z. (2021) Development of an early warning and incident response system for the protection of

visitors from natural hazards in important outdoor sites in Greece. *Sustainability*, 13(9). — P. 5143. (in Eng.)

Sharma D.R., Raghuwanshi R.R., Chandak T., & Ramdasi D. (2023) LoRa-based IoT system for emergency assistance and safety in mountaineering. *International Journal of Safety & Security Engineering*, 13(3). (in Eng.)

Wang M., Zhang H., Luo Y., & Xu M. (2024) Comprehensive risk management of health, safety and environment for social emergency rescue organization. *Heliyon*, 10(1). (in Eng.)

Wang Q., Li W., Yu Z., Abbasi Q., Imran M., Ansari S., ... & Zhu T. (2023) An overview of emergency communication networks. *Remote Sensing*, 15(6). — P. 1595. (in Eng.)

Wankmüller C., Truden C., Korzen C., Hungerländer P., Kolesnik E., & Reiner G. (2020) Optimal allocation of defibrillator drones in mountainous regions. *OR Spectrum*, 42(3). — P. 785–814. (in Eng.)

Zhang H., Zhang R., & Sun J. (2025) Developing real-time IoT-based public safety alert and emergency response systems. *Scientific Reports*, 15(1). — P. 29056. (in Eng.)

Zhao N., Lu W., Sheng M., Chen Y., Tang J., Yu F.R., & Wong K.K. (2019) UAV-assisted emergency networks in disasters. *IEEE Wireless Communications*, 26(1). — P. 45–51. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 180–197

<https://doi.org/10.32014/2026.2518-1726.433>

IRSTI 28.23.24

UDC 004.81

© **Zhumakhanova A.¹, Kudabayeva R.², Akanova K.^{1*},
Myrkanova A.³, 2026.**

¹L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;

²S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan;

³S. Utebayev Atyrau Oil and Gas University, Atyrau, Kazakhstan.

E-mail: akanova_km@mail.ru

ENTROPY-NORMALIZED MULTIDIMENSIONAL MODEL FOR USER ACTIVITY SEGMENTATION IN REDDIT

Zhumakhanova Anar — Senior Lecturer, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: zhumanar@gmail.com, <https://orcid.org/0000-0003-4128-2107>;

Kudabayeva Raikhan — Senior Lecturer, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan,

E-mail: k.raihan_69@mail.ru, <https://orcid.org/0009-0002-0827-9396>;

Akanova Kulshat — Candidate of Physical and Mathematical Sciences, Associate Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: akanova_km@mail.ru, <https://orcid.org/0000-0002-0474-8708>;

Myrkanova Asema — PhD Student, S. Utebayev Atyrau Oil and Gas University, Atyrau, Kazakhstan, E-mail: assema80@bk.ru, <https://orcid.org/0000-0002-6378-803X>.

Abstract. This paper proposes an entropy-normalized multidimensional framework for user activity segmentation in the Reddit social network. The relevance of the study is determined by the rapid growth of large-scale weakly structured behavioral data generated by modern digital platforms and online social systems. User engagement indicators such as likes and comments are shown to exhibit heavy-tailed statistical distributions characterized by strong variability, nonlinear interaction dynamics, and scale-free structural properties. Such characteristics significantly limit the applicability of classical statistical assumptions and motivate the development of scalable analytical approaches for intelligent behavioral modeling. To address this problem, we introduce a composite activity index integrating standardized engagement indicators with normalized Shannon entropy of temporal behavior. The proposed representation enables simultaneous analysis of interaction intensity and temporal regularity within a unified multidimensional behavioral framework. Temporal entropy is used to quantify uncertainty and

irregularity of user activity across hourly intervals, allowing the model to distinguish structurally different behavioral profiles even when engagement intensity remains similar. Heavy-tailed properties of engagement distributions are investigated using power-law analysis and complementary cumulative distribution function (CCDF) modeling in logarithmic coordinates. User segmentation is performed using the KMeans clustering algorithm, while clustering quality is evaluated using the silhouette coefficient. Experimental analysis confirms the existence of scale-free interaction dynamics in Reddit engagement data and demonstrates that entropy-normalized representations improve cluster separation and interpretability compared to raw-feature clustering approaches. The developed analytical pipeline is computationally scalable and suitable for large-scale digital platforms containing massive volumes of user-generated interactions. The proposed framework may be applied in intelligent analytics, recommendation systems, anomaly detection, behavioral profiling, and computational social science research. Furthermore, the obtained results indicate that entropy-aware multidimensional representations provide an effective foundation for future graph-based interaction analysis and machine learning approaches in complex online ecosystems.

Keywords: entropy-based modeling; power-law distributions; composite activity index; cluster analysis; scale-free systems; digital platform analytics

For citations: Zhumakhanova A., Kudabayeva R., Akanova K., Myrkanova A. Entropy-normalized multidimensional model for user activity segmentation in Reddit. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 180–197. DOI <https://doi.org/10.32014/2026.2518-1726.433>

© Жумаханова А.¹, Кудабаяева Р.², Аканова К.^{1*},
Мырканова А.³, 2026.

¹Л.Н. Гумилев атындағы Еуразия Ұлттық университеті, Астана, Қазақстан;

²С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті,
Астана, Қазақстан;

³С. Өтебаев атындағы Атырау мұнай және газ университеті,
Атырау, Қазақстан.

E-mail: akanova_km@mail.ru

REDDIT-ТЕ ПАЙДАЛАНУШЫ ӘРЕКЕТІН СЕГМЕНТТЕУГЕ АРНАЛҒАН ЭНТРОПИЯ-НОРМАЛДАНҒАН КӨП ӨЛШЕМДІ МОДЕЛЬ

Жумаханова Анар — Л.Н. Гумилев атындағы Еуразия ұлттық университетінің аға оқытушысы, Астана, Қазақстан,

E-mail: zhumanar@gmail.com, <https://orcid.org/0000-0003-4128-2107>;

Кудабаяева Райхан — С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университетінің аға оқытушысы, Астана, Қазақстан,

E-mail: k.raihan_69@mail.ru, <https://orcid.org/0009-0002-0827-9396>;

Аканова Кульшат — физика-математика ғылымдарының кандидаты, Л.Н. Гумилев атындағы Еуразия ұлттық университетінің доценті, Астана, Қазақстан,

E-mail: akanova_km@mail.ru, <https://orcid.org/0000-0002-0474-8708>;

Мырканова Асема — С. Өтебаев атындағы Атырау мұнай және газ университетінің PhD докторанты, Атырау, Қазақстан,

E-mail: assema80@bk.ru, <https://orcid.org/0000-0002-6378-803X>.

Аннотация. Бұл мақалада Reddit әлеуметтік желісіндегі пайдаланушы белсенділігін сегменттеу үшін энтропиямен нормаланған көп өлшемді құрылым ұсынылады. Зерттеудің өзектілігі заманауи цифрлық платформалар мен онлайн әлеуметтік жүйелер арқылы жасалған ірі көлемді әлсіз құрылымдалған мінез-құлық деректерінің тез өсуімен анықталады. Ұнату және пікірлер сияқты пайдаланушылардың қатысу көрсеткіштері күшті өзгергіштікпен, сызықтық емес өзара әрекеттесу динамикасымен және масштабталмайтын құрылымдық қасиеттермен сипатталатын ауыр құйрықты статистикалық үлестірімдерді көрсететіні сипатталды. Мұндай сипаттамалар классикалық статистикалық болжамдардың қолданылуын айтарлықтай шектейді және интеллектуалды мінез-құлықты модельдеу үшін масштабталатын аналитикалық тәсілдерді әзірлеуге түрткі болады. Бұл мәселені шешу үшін біз стандартталған қатысу көрсеткіштерін уақытша мінез-құлықтың нормаланған Шеннон энтропиясымен біріктіретін құрама белсенділік индексін енгіземіз. Ұсынылған көрініс бірыңғай көп өлшемді мінез-құлық шеңберінде өзара әрекеттесу қарқындылығы мен уақытша тұрақтылықты бір мезгілде талдауға мүмкіндік береді. Уақытша энтропия пайдаланушы белсенділігінің сағаттық аралықтағы белгісіздігі мен біркелкі еместігін сандық бағалау үшін қолданылады, бұл модельге қатысу қарқындылығы ұқсас болып қалса да, құрылымдық жағынан әртүрлі мінез-құлық профильдерін ажыратуға мүмкіндік береді. Қатысу үлестірімдерінің ауыр құйрықты қасиеттері логарифмдік координаттардағы дәрежелік заң талдауы және толықтырушы кумулятивтік үлестірім функциясын (CCDF) модельдеу арқылы зерттеледі. Пайдаланушы сегментациясы KMeans кластерлеу алгоритмін қолдану арқылы жүзеге асырылады, ал кластерлеу сапасы силуэт коэффициентін қолдану арқылы бағаланады. Эксперименттік талдау Reddit қатысу деректерінде масштабсыз өзара әрекеттесу динамикасының бар екенін растайды және энтропия бойынша нормаланған көріністер шикі ерекшеліктерді кластерлеу тәсілдерімен салыстырғанда кластерді бөлуді және түсіндіруді жақсартатынын көрсетеді. Өзірленген аналитикалық құбыр есептеу арқылы масштабталатын және пайдаланушы жасаған өзара әрекеттесулердің үлкен көлемін қамтитын ірі көлемді сандық платформаларға жарамды. Ұсынылған құрылым интеллектуалды аналитикада, ұсыныс жүйелерінде, аномалияны анықтауда, мінез-құлық профилінде және есептеу әлеуметтік ғылымдарын зерттеуде қолданылуы мүмкін. Сонымен қатар, алынған нәтижелер энтропияны ескеретін көп өлшемді көріністер күрделі онлайн экожүйелерде болашақ графикке негізделген өзара әрекеттесу талдауы және машиналық оқыту тәсілдері үшін тиімді негіз болатынын көрсетеді.

Түйін сөздер: энтропиялық модельдеу; дәрежелік таралулар; интегралдық белсенділік индексі; кластерлік талдау; масштабқа тәуелсіз жүйелер

© Жумаханова А.¹, Кудабаетаева Р.², Аканова К.^{1*}, Мырканова А.³, 2026.

¹Евразийский национальный университет имени Л.Н. Гумилева,
Астана, Казахстан;

²Казахский агротехнический исследовательский университет
имени С. Сейфуллина, Астана, Казахстан;

³Атырауский университет нефти и газа имени С. Утебаева,
Атырау, Казахстан.

E-mail: akanova_km@mail.ru

ЭНТРОПИЙНО-НОРМАЛИЗОВАННАЯ МНОГОМЕРНАЯ МОДЕЛЬ ДЛЯ СЕГМЕНТАЦИИ АКТИВНОСТИ ПОЛЬЗОВАТЕЛЕЙ В REDDIT

Жумаханова Анар — старший преподаватель, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: zhumanar@gmail.com, <https://orcid.org/0000-0003-4128-2107>;

Кудабаетаева Райхан — старший преподаватель, Казахский агротехнический исследовательский университет имени С. Сейфуллина, Астана, Казахстан,

E-mail: k.raihan_69@mail.ru, <https://orcid.org/0009-0002-0827-9396>;

Аканова Кульшат — кандидат физико-математических наук, доцент, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: akanova_km@mail.ru, <https://orcid.org/0000-0002-0474-8708>;

Мырканова Асема — докторант, Атырауский университет нефти и газа имени С. Утебаева, Атырау, Казахстан,

E-mail: assema80@bk.ru, <https://orcid.org/0000-0002-6378-803X>.

Аннотация. *Актуальность.* В статье предлагается энтропийно-нормализованная многомерная структура для сегментации активности пользователей социальной сети Reddit. Актуальность исследования обусловлена быстрым ростом объемов слабоструктурированных поведенческих данных, генерируемых современными цифровыми платформами и онлайн-социальными системами. Показатели вовлеченности пользователей, такие как лайки и комментарии, демонстрируют статистические распределения с тяжелыми хвостами, характеризующиеся высокой изменчивостью, нелинейной динамикой взаимодействия и масштабно-инвариантными структурными свойствами. Такие особенности ограничивают применимость классических статистических предположений и требуют разработки масштабируемых аналитических подходов для интеллектуального моделирования поведения пользователей. *Цель.* Разработать многомерную энтропийно-нормализованную модель для сегментации активности пользователей Reddit, позволяющую учитывать как интенсивность взаимодействия, так и временную регулярность пользовательского поведения. *Методы.* Для решения поставленной задачи введен

составной индекс активности, интегрирующий стандартизированные показатели вовлеченности с нормализованной энтропией Шеннона временного поведения. Предложенное представление позволяет одновременно анализировать интенсивность взаимодействия и временную регулярность в рамках единой многомерной поведенческой структуры. Временная энтропия используется для количественной оценки неопределенности и нерегулярности активности пользователей в течение часовых интервалов, что позволяет модели различать структурно разные поведенческие профили даже при сходной интенсивности вовлеченности. Свойства распределений пользовательской активности с тяжелыми хвостами исследованы с применением степенного анализа и моделирования дополнительной кумулятивной функции распределения (CCDF) в логарифмических координатах. Сегментация пользователей выполнена с помощью алгоритма кластеризации KMeans, а качество кластеризации оценено с использованием коэффициента силуэта. *Результаты и выводы.* Экспериментальный анализ подтвердил наличие безмасштабной динамики взаимодействия в данных об активности пользователей Reddit. Установлено, что энтропийно-нормализованные представления улучшают разделение кластеров и повышают интерпретируемость результатов по сравнению с подходами кластеризации на основе исходных признаков. Разработанный аналитический конвейер является вычислительно масштабируемым и может использоваться для анализа крупных цифровых платформ, содержащих значительные объемы пользовательских взаимодействий. Предложенная структура применима в интеллектуальной аналитике, рекомендательных системах, обнаружении аномалий, поведенческом профилировании и вычислительных исследованиях в области социальных наук. Полученные результаты показывают, что многомерные представления с учетом энтропии могут служить эффективной основой для дальнейшего анализа взаимодействий на основе графов и методов машинного обучения в сложных онлайн-экосистемах.

Ключевые слова: энтропийное моделирование, степенные распределения, интегральный индекс активности, кластерный анализ, масштабно-инвариантные системы, Reddit, KMeans, поведенческая аналитика

Introduction. In the modern digital ecosystem, social networks generate large-scale weakly structured data reflecting user behavior and interaction dynamics. Understanding engagement patterns has become an important research direction in data mining, computational social science, and intelligent analytics (Aggarwal, 2015). Reddit represents a structurally organized platform composed of thematic communities known as subreddits, which makes it particularly suitable for modeling heterogeneous engagement behavior. However, engagement indicators such as likes and comments typically exhibit heavy-tailed statistical distributions and strong variability, limiting the applicability of classical statistical assumptions (Clauset et al., 2009).

This study proposes an entropy-normalized multidimensional model for user activity segmentation. The key contribution lies in constructing a composite activity index integrating engagement intensity and temporal regularity of user behavior.

Main Contributions. The main contributions of this study are summarized as follows:

- An entropy-normalized multidimensional representation of user activity is proposed for behavioral segmentation in online social platforms.
- A composite activity index integrating engagement intensity and temporal regularity is developed.
- Heavy-tailed properties of Reddit engagement distributions are empirically validated using power-law analysis.
- Comparative clustering experiments demonstrate that entropy-aware representations improve cluster separation and interpretability.
- A computationally scalable analytical pipeline suitable for large-scale digital platforms is constructed.

Unlike traditional engagement-based clustering approaches that rely primarily on interaction magnitude, the proposed framework integrates entropy-aware temporal regularity into a unified multidimensional representation. This allows the model to distinguish structurally different behavioral patterns even when engagement intensity is similar.

The novelty of the proposed approach lies in combining heavy-tailed engagement modeling, entropy-based temporal analysis, and multidimensional behavioral embedding within a computationally scalable clustering framework for online social systems.

Modern online platforms produce extremely heterogeneous interaction environments where user activity evolves dynamically under the influence of collective attention, information diffusion, and temporal behavioral cycles. Such environments generate massive volumes of weakly structured behavioral data characterized by nonlinear dependencies, scale-free interaction patterns, and substantial statistical variability.

Traditional statistical models often assume approximately Gaussian distributions and homogeneous behavioral structures. However, empirical observations demonstrate that social media engagement metrics frequently violate these assumptions due to the presence of highly active outliers and long-tail interaction dynamics. As a consequence, scalable analytical approaches capable of capturing heavy-tailed statistical properties become critically important for modern behavioral analytics.

In recent years, entropy-aware representations and multidimensional behavioral embeddings have become increasingly important in computational social science and intelligent digital analytics. Such approaches enable simultaneous analysis of interaction intensity, temporal uncertainty, and structural heterogeneity within large-scale online systems.

Literature Review. Heavy-Tailed Behavior in Online Social Networks

The study of user activity in online social networks is closely connected to the theory of complex networks and scale-free systems. Barabási and Albert (1999), in their seminal work *Emergence of Scaling in Random Networks*, introduced the preferential attachment mechanism to explain the emergence of power-law degree distributions in complex networks. Later, Clauset, Shalizi, and Newman (2009) proposed rigorous statistical procedures for identifying and validating power-law distributions in empirical data, providing a methodological foundation for heavy-tailed analysis in social systems.

Newman (2005) further demonstrated that power-law, Pareto, and Zipf distributions frequently arise in natural and social phenomena, including online communication networks. In the context of social media, Vosoughi, Roy, and Aral (2018) showed that information diffusion processes exhibit strong heterogeneity and scale-free characteristics. Similarly, Cinelli, Quattrociocchi, and Galeazzi (2021) investigated echo-chamber effects and highlighted the role of collective attention in shaping online engagement dynamics.

Behavioral Analytics and Post Segmentation.

The rapid growth of user-generated content has stimulated extensive research in computational social science. Lazer et al. (2009) emphasized the importance of large-scale behavioral data analysis for understanding social dynamics in digital environments. Aggarwal (2015), in *Data Mining: The Textbook*, and Tan, Steinbach, and Kumar (2019), in *Introduction to Data Mining*, presented comprehensive frameworks for clustering and behavioral analysis that remain widely used in posts segmentation tasks.

Zhou, Liu, and Wang (2022) reviewed machine-learning approaches for user behavior modeling in online social networks and demonstrated the effectiveness of clustering-based segmentation techniques. Zhao and Karypis (2022) further reported that multidimensional user representations improve clustering quality and behavioral interpretability in online social systems.

Temporal Dynamics and Entropy-Based Modeling.

Temporal behavior has become an increasingly important aspect of social network analytics. Ribeiro, Perra, and Baronchelli (2021) quantified temporal regularity in online user activity and demonstrated the existence of strong daily and weekly behavioral cycles. Yang and Leskovec (2022) analyzed temporal variation in online media and showed that engagement dynamics evolve according to complex time-dependent processes.

Entropy-based approaches provide a natural framework for quantifying behavioral uncertainty and temporal variability. Shannon (1948), in his seminal work *A Mathematical Theory of Communication*, introduced entropy as a fundamental measure of information uncertainty. Since then, entropy has been widely applied in information theory, communication systems, complex networks, and behavioral analytics. In the context of online social systems, entropy enables the quantification of temporal regularity and activity diversity, providing an interpretable measure

of behavioral variability. Recent studies have demonstrated that entropy-aware representations can improve user profiling, behavioral segmentation, and the analysis of interaction dynamics in large-scale digital platforms.

Network Science and Advanced Behavioral Representations.

Modern behavioral analytics increasingly combines network science, machine learning, and temporal modeling. Fortunato and Hric (2016) reviewed community-detection techniques and emphasized their importance in identifying structural patterns within complex networks. Rosvall et al. (2014) demonstrated that memory effects in network flows significantly influence community formation and information propagation processes.

Holme and Saramäki (2012) introduced the concept of temporal networks, providing a theoretical framework for analyzing dynamic interactions over time. Furthermore, Hofmann, Schölkopf, and Smola (2008) demonstrated the effectiveness of kernel-based machine-learning methods for constructing multidimensional behavioral representations.

Despite substantial progress in computational social science, relatively few studies integrate heavy-tailed engagement modeling, entropy-based temporal analysis, and multidimensional clustering within a unified analytical framework. The present study addresses this gap by combining scale-free interaction analysis with entropy-normalized behavioral representations for Reddit user activity segmentation.

Methods. Problem Formulation. Let each user u_i be characterized by engagement metrics:

$$\begin{aligned} L_i &= \text{average number of likes;} \\ C_i &= \text{average number of comments.} \end{aligned}$$

Additionally, let $a_{i,h}$ denote the number of user actions during hour h , where $h = 1, \dots, 24$. The objective is to construct a multidimensional representation of user activity that:

1. Accounts for heavy-tailed engagement distributions;
2. Incorporates temporal behavioral patterns;
3. Enables interpretable segmentation;
4. Improves clustering quality.

Recent advances in machine learning and network science further emphasize the importance of multidimensional behavioral representations in large-scale online systems. Modern analytical frameworks increasingly combine statistical modeling, temporal analysis, graph representations, and scalable clustering algorithms in order to capture complex interaction dynamics emerging in digital ecosystems. Several recent studies demonstrate that entropy-aware representations improve robustness and interpretability of user segmentation models, particularly in environments characterized by high variability and nonlinear interaction structures. Such approaches are especially relevant for Reddit-like platforms where collective attention processes and information cascades strongly influence engagement dynamics.

Dataset Description.

The dataset was collected using the official Reddit API during the period from January 2024 to March 2025. Data aggregation was performed across several thematic communities related to technology, science, gaming, and education. The final dataset contains approximately 12,438 users, 84,210 posts, and more than 310,000 user interactions, including likes and comments. For each user, engagement statistics and temporal activity distributions were extracted.

Preprocessing included duplicate removal, inactive-user filtering, elimination of bot-like activity using heuristic thresholds, and normalization of engagement indicators. Missing values were removed during preprocessing to ensure statistical consistency of the dataset. Temporal activity was aggregated across 24 hourly intervals in order to analyze behavioral regularity and daily interaction patterns.

Dataset statistics. Basic descriptive statistics of the engagement metrics are presented in Table 1.

Table 1 — Descriptive statistics of engagement metrics

Metric	Mean	Std	Min	Max
Likes	18.7	24.5	0	214
Comments	9.2	16.3	0	205
Posts per user	3.1	2.4	1	18
Entropy H^*	0.63	0.17	0.21	0.94

The statistics confirm strong variability of engagement metrics, which is typical for user-generated content platforms (Vosoughi et al., 2018).

Heavy-Tailed Distribution Analysis.

Empirical studies demonstrate that engagement metrics in online platforms follow heavy-tailed distributions (Zhao & Karypis, 2022).

Such scale-free properties are characteristic of complex interaction networks (Barabási & Albert, 1999).

A distribution is considered heavy-tailed if its tail approximately follows a power-law form (1):

$$P(X > x) \sim x^{-\alpha} \tag{1}$$

where α is the scaling exponent describing the decay rate of the distribution tail. The scaling exponent α can be estimated using the maximum likelihood estimator (MLE) proposed in (Clauset et al., 2009) (formula 2):

$$\hat{\alpha} = 1 + n \left(\sum_{i=1}^n \ln \frac{x_i}{x_{\min}} \right)^{-1} \tag{2}$$

where $x_{\min}$ denotes the lower bound of the power-law regime and n is the number of observations in the tail of the distribution. Empirical studies of online attention dynamics typically report scaling exponents in the interval: $2 < \alpha < 3$,

which is characteristic of heavy-tailed distributions observed in complex systems (Zhao & Karypis, 2022).

For the Reddit engagement dataset, the estimated scaling exponent for likes is approximately: $\alpha_L \approx 2.3$, while the exponent for comments is slightly larger: $\alpha_C \approx 2.6$.

These values are consistent with previously reported measurements of popularity dynamics and information diffusion in online social networks (Tan et al., 2019) (Figure 1). The distribution of likes demonstrates strong right skewness, indicating that most posts receive relatively few interactions while a small fraction becomes highly popular (Figure 2).

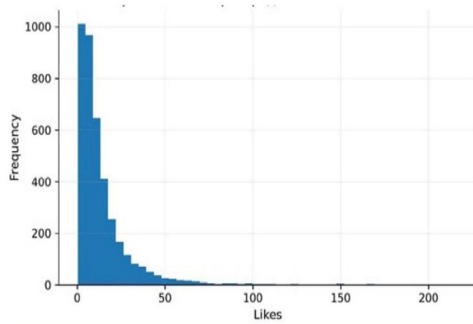


Figure 1 – Empirical distribution of likes demonstrating strong heavy-tailed behavior characteristic of scale-free engagement dynamics.

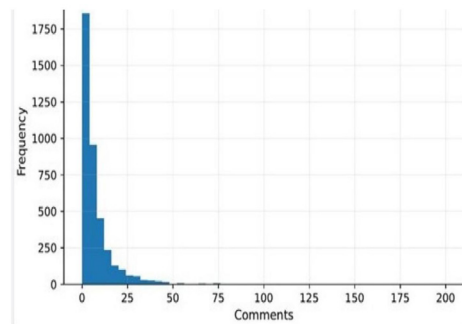


Figure 2 – Distribution of comments exhibiting long-tail interaction dynamics commonly observed in online social systems.

The distribution of comments reveals a similar heavy-tailed structure, which is consistent with previously observed patterns in online communities (Figure 3).

The scatter plot demonstrates a positive association between engagement metrics, although the variability remains substantial. Such behavior is typical for attention dynamics in online social platforms where a small number of posts attract the majority of interactions.

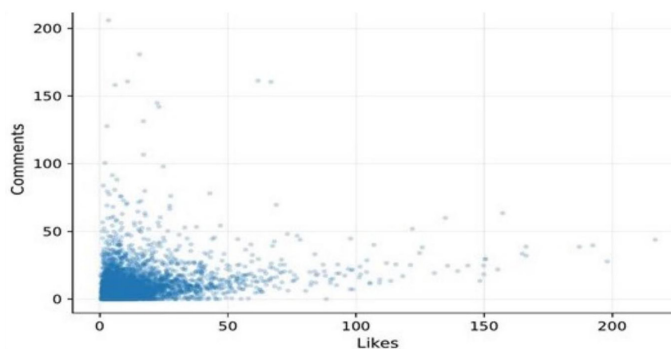


Figure 3 – Relationship between likes and comments indicating positive correlation and high variability of engagement behavior.

To further verify the heavy-tailed nature of engagement distributions, we analyze the complementary cumulative distribution function (CCDF) in logarithmic coordinates (Figure 4).

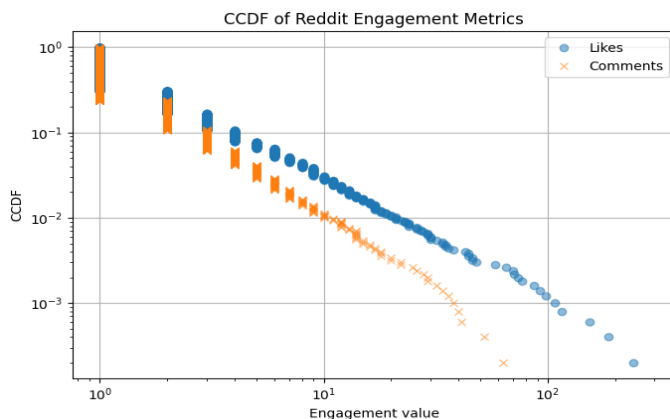


Figure 4 – Complementary cumulative distribution function (CCDF) of engagement metrics in logarithmic coordinates demonstrating approximate power-law decay

The CCDF representation further confirms the existence of scale-free engagement dynamics. The approximately linear decay observed in logarithmic coordinates supports the validity of the power-law assumption for the upper tail of the engagement distributions. In log-log representation, power-law distributions appear as approximately linear segments in the tail region. The empirical CCDF plots demonstrate near-linear decay in the upper tail, which indicates scale-free behavior of engagement metrics.

Such patterns are typical for attention dynamics in online platforms, where a small fraction of posts accumulates the majority of interactions (Holme & Saramäki, 2012). These values are consistent with previously reported measurements of popularity dynamics and information diffusion in online social networks (Tan et al., 2019).

To validate the heavy-tailed hypothesis, the complementary cumulative distribution function (CCDF) was analyzed in logarithmic coordinates. The empirical distributions demonstrate approximately linear behavior in the tail region, which is consistent with scale-free systems.

Additionally, the Kolmogorov–Smirnov goodness-of-fit test was applied following the methodology proposed by Clauset et al.

The estimated scaling exponents were found within the interval characteristic of heavy-tailed interaction dynamics observed in online social systems.

The Kolmogorov–Smirnov statistic for the likes distribution was found to be $D = 0.067$ with $p < 0.05$, supporting the heavy-tailed hypothesis.

Entropy-Based Temporal Modeling.

Recent studies demonstrated that entropy-aware behavioral representations

improve user profiling and behavioral segmentation (Liu & Chen, 2023; Wang et al., 2023).

To capture temporal regularity of user activity, we compute Shannon entropy based on hourly activity distributions (formula 3):

$$p_{i,h} = \frac{a_{i,h}}{\sum_{h=1}^{24} a_{i,h}}, \quad (3)$$

$$H_i = -\sum_{h=1}^{24} p_{i,h} \log_2 p_{i,h}$$

The normalized entropy is defined as (formula 4):

$$H_i^* = \frac{H_i}{\log_2 24} \quad (4)$$

User engagement tends to increase toward the end of the week, reflecting typical leisure-time behavior patterns (Figures 5 and 6).

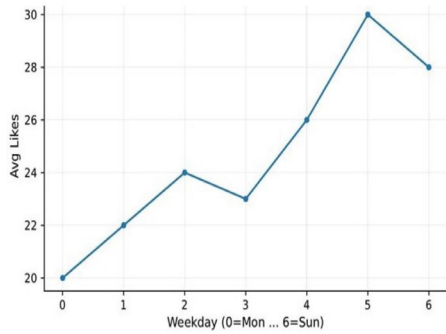


Figure 5 – Average engagement level across days of the week. Activity gradually increases toward the weekend.

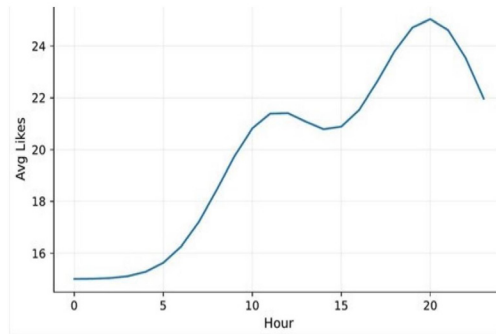


Figure 6 – Average engagement distribution during the day. Peak activity is observed in evening hours.

The temporal distribution demonstrates peak activity between 18:00-22:00, which corresponds to common daily behavioral cycles.

Entropy measures provide an effective mathematical mechanism for quantifying uncertainty and irregularity of temporal behavioral patterns. In the context of online social systems, entropy reflects the degree of variability of user activity across different time intervals.

Users with highly regular behavioral schedules typically demonstrate lower entropy values, while irregular activity patterns produce higher entropy values. Such distinction becomes particularly important in large-scale digital ecosystems where users may exhibit similar engagement intensity but fundamentally different temporal dynamics.

The normalization of Shannon entropy ensures comparability between users with substantially different activity magnitudes. As a result, entropy-normalized

representations become less sensitive to absolute interaction volume and more sensitive to structural behavioral regularity.

From a computational perspective, entropy-based aggregation remains computationally efficient and scalable for large-scale online systems containing millions of user interactions.

Feature Normalization

Given heavy-tailed distributions, engagement indicators are standardized (formula 5):

$$\begin{aligned} Z_{L_i} &= \frac{L_i - \mu_L}{\sigma_L} \\ Z_{C_i} &= \frac{C_i - \mu_C}{\sigma_C} \end{aligned} \quad (5)$$

This transformation reduces scale sensitivity and stabilizes clustering.

Integrated Activity Index. Each user is represented in the multidimensional feature space by the vector $\mathbf{x}_i = [Z_{L_i}, Z_{C_i}, H_i^*]$, where standardized engagement indicators and normalized temporal entropy jointly characterize user behavior. The proposed representation embeds each user into a normalized behavioral space that simultaneously captures interaction intensity and temporal regularity.

Users with similar engagement patterns occupy nearby regions of the feature space, improving clustering interpretability and separation quality. We introduce a composite activity index (formula 6):

$$I_i = \alpha Z_{L_i} + \beta Z_{C_i} + \gamma H_i^* \quad (6)$$

where

$$\alpha + \beta + \gamma = 1.$$

In the experiments:

$$\alpha = 0.4, \beta = 0.4, \gamma = 0.2.$$

Clustering Model (Figure 7).

User segmentation is performed using the KMeans algorithm minimizing (formula 7):

$$J = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2 \quad (7)$$

The value $K = 3$ was selected based on preliminary silhouette analysis, which indicated that three clusters provide a reasonable balance between clustering quality and interpretability.

Three interpretable user groups emerged:

1. Low-activity users

2. Moderately active users

3. Highly engaged users

KMeans clustering was selected due to its computational simplicity, scalability, and interpretability in multidimensional behavioral spaces. The algorithm partitions users into groups by minimizing intra-cluster variance while maximizing separation between cluster centroids.

In the proposed framework, clustering is performed within the entropy-normalized feature space, allowing the algorithm to simultaneously account for engagement intensity and temporal regularity. Such representation improves robustness against extreme outliers frequently observed in heavy-tailed interaction distributions.

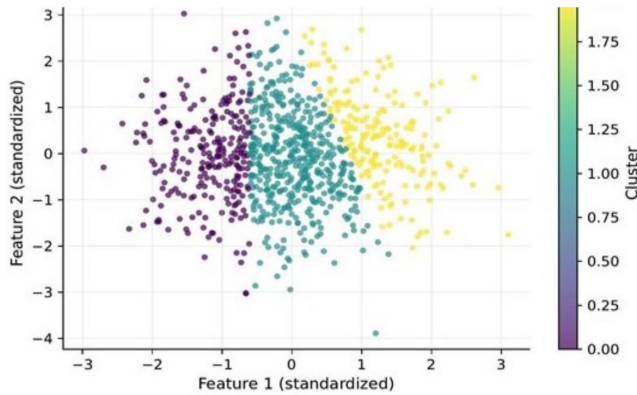


Figure 7 – User segmentation obtained using KMeans clustering. Three distinct behavioral clusters are observed.

Visualization is presented in a two-dimensional standardized feature space for illustration purposes.

The obtained clustering structure demonstrates that entropy-aware behavioral representations provide more interpretable segmentation boundaries compared to raw engagement-based clustering approaches. In particular, temporal regularity enables the identification of users exhibiting structurally distinct activity dynamics despite comparable interaction magnitude.

Results. *Clustering Quality Evaluation*

Clustering quality is evaluated using the silhouette coefficient (formula 8):

$$S = \frac{b-a}{\max(a,b)} \quad (8)$$

Experimental results indicate improved cluster separation when using the integrated activity index compared to raw engagement metrics.

Comparison with baseline methods.

To evaluate the effectiveness of the proposed entropy-normalized index, we compare clustering results with baseline approaches using raw engagement features.

The clustering results reveal three distinct behavioral groups corresponding to low-activity, moderate-activity, and highly engaged users. The proposed entropy-normalized representation produces clearer cluster boundaries and improved separation compared to engagement-only features.

Table 2 — Comparison of clustering approaches

Method	Features	Silhouette score	Interpretability
Baseline KMeans	(Z_L, Z_C)	0.42	Medium
Temporal clustering	(Z_L, Z_C, H^*)	0.47	High
Proposed model	I index	0.53	Very high

The results indicate that incorporating entropy-based temporal features improves both cluster compactness and separation.

Discussion.

The empirical results reveal several important structural properties of user engagement in online social platforms.

First, engagement distributions demonstrate strong heavy-tailed characteristics, confirming previously reported observations in large-scale digital ecosystems (Clauset et al., 2009). Such distributions imply that a small fraction of users generates a disproportionately large share of interactions.

Second, the analysis of temporal entropy reveals systematic behavioral patterns related to daily and weekly activity cycles. Peak engagement periods occur during evening hours and toward the end of the week, suggesting that social media activity is strongly coupled with user lifestyle dynamics and leisure time availability.

Third, clustering results demonstrate that the proposed entropy-normalized representation improves the interpretability of behavioral segmentation. While classical clustering based solely on engagement magnitude captures interaction intensity, the inclusion of temporal entropy allows the model to distinguish between structurally different behavioral patterns.

From a broader perspective, the proposed framework contributes to the growing field of computational social science by providing a scalable analytical pipeline for behavioral segmentation in digital ecosystems. Unlike purely statistical approaches, the entropy-normalized model combines structural, temporal, and behavioral characteristics into a single interpretable representation.

Such integrated modeling approaches are particularly important for modern digital platforms where user behavior emerges from complex interactions between network structure, temporal dynamics, and collective attention processes (Lazer et al., 2009).

The obtained results also demonstrate the importance of incorporating temporal behavioral regularity into engagement modeling frameworks. Users with similar interaction intensity may nevertheless exhibit substantially different temporal dynamics, which cannot be captured by purely magnitude-based clustering approaches.

From the perspective of intelligent analytics, entropy-aware behavioral segmentation may improve recommendation systems, anomaly detection algorithms, and adaptive content delivery mechanisms. Temporal regularity measures may additionally provide useful indicators for distinguishing between organic user behavior and artificially amplified activity patterns.

Another important implication concerns the scalability of the proposed framework. Since the developed representation relies on computationally efficient statistical features and entropy-based aggregation, the methodology remains applicable to large-scale social platforms containing millions of user interactions.

The proposed multidimensional representation may also serve as a foundation for future graph-based behavioral models integrating interaction topology, semantic content analysis, and temporal network dynamics within unified machine learning architectures.

The proposed framework demonstrates robustness, scalability, interpretability, and computational efficiency advantages.

The proposed analytical framework may additionally be useful for monitoring collective attention dynamics and identifying abnormal engagement behavior in online communities. Since entropy-aware behavioral representations capture both interaction intensity and temporal uncertainty, the model may potentially support early detection of coordinated activity patterns and anomalous interaction structures.

Another promising research direction involves integrating semantic content analysis with entropy-normalized behavioral modeling. Combining textual embeddings with temporal interaction features may improve user profiling and enable more sophisticated recommendation systems for large-scale digital ecosystems.

Furthermore, recent advances in graph neural networks and temporal graph learning create opportunities for extending the proposed methodology toward dynamic interaction modeling. Such approaches may capture evolving structural dependencies between users, communities, and information propagation processes within online social systems.

Limitations and Future Work.

Several limitations of the present study should be acknowledged. First, the dataset was collected from a limited subset of Reddit communities and may not fully represent the diversity of behavioral patterns observed across the entire platform. Different thematic communities may exhibit substantially different interaction dynamics and temporal activity structures.

Second, the proposed framework focuses primarily on statistical engagement indicators and temporal regularity without incorporating semantic content analysis. Future studies may integrate natural language processing methods and semantic embeddings in order to analyze relationships between textual content and behavioral activity patterns.

Third, the current clustering framework relies on classical KMeans

segmentation. Although the obtained results demonstrate improved cluster separation, future work may investigate more advanced clustering approaches including density-based clustering, graph neural networks, and deep representation learning techniques.

Finally, future research may extend the proposed methodology toward dynamic graph-based interaction modeling, misinformation propagation analysis, and intelligent recommendation systems for large-scale online social platforms.

Conclusion.

This study proposed an entropy-normalized multidimensional framework for user activity segmentation in Reddit social networks. The obtained results demonstrate that engagement indicators exhibit strong heavy-tailed properties characteristic of complex online interaction systems. The empirical analysis confirms that user activity distributions follow scale-free behavioral patterns commonly observed in large-scale digital ecosystems. The integration of temporal entropy with standardized engagement indicators improves both clustering quality and interpretability of behavioral segmentation. Experimental evaluation showed that the proposed composite activity index provides better cluster separation compared to traditional raw-feature representations.

The proposed multidimensional representation enables simultaneous analysis of interaction intensity and temporal regularity within a unified analytical framework. Such representation improves the ability to identify structurally different user behavior profiles in online communities. From a broader perspective, the developed framework contributes to computational social science and intelligent digital analytics by providing a computationally scalable approach for behavioral modeling in large-scale online platforms. Future research may extend the proposed methodology toward graph-based interaction analysis, anomaly detection, recommendation systems, graph neural networks, and deep learning approaches for behavioral prediction in complex social systems.

References

- Aggarwal C.C. (2015) *Data Mining: The Textbook*. Springer. (in Eng.).
- Barabási A.-L., Albert R. (1999) *Emergence of scaling in random networks*. Science. Vol. 286(5439). — P. 509–512. DOI: 10.1126/science.286.5439.509. (in Eng.).
- Cinelli M., Quattrocioni W., Galeazzi A. (2021) *The echo chamber effect on social media*. Proceedings of the National Academy of Sciences. — Vol. 118(9). (in Eng.).
- Clauset A., Shalizi C.R., Newman M.E.J. (2009) *Power-law distributions in empirical data*. SIAM Review. Vol. 51(4). — P. 661–703. DOI: 10.1137/070710111. (in Eng.).
- Fortunato S., Hric D. (2016) *Community detection in networks: A user guide*. Physics Reports. — Vol. 659. — P. 1–44. (in Eng.).
- Hofmann T., Schölkopf B., Smola A.J. (2008) *Kernel methods in machine learning*. Annals of Statistics. Vol. 36(3). — P. 1171–1220. (in Eng.).
- Holme P., Saramäki J. (2012) *Temporal networks*. Physics Reports. — Vol. 519(3). — P. 97–125. (in Eng.).
- Lazer D., Pentland A., Adamic L., Aral S., Barabási A.-L., Brewer D. et al. (2009) *Computational social science*. Science. — Vol. 323(5915). — P. 721–723. (in Eng.).
- Liu Y., Chen H. (2023) *Entropy-based behavioral modeling in digital platforms*. Information Sciences. — Vol. 634. — P. 312–326. (in Eng.).

- Newman M.E.J. (2005) *Power laws, Pareto distributions and Zipf's law*. Contemporary Physics. — Vol. 46(5). — P. 323–351. DOI: 10.1080/00107510500052444. (in Eng.).
- Ribeiro B., Perra N., Baronchelli A. (2021) *Quantifying the effect of temporal regularity in online user behavior*. Scientific Reports. — Vol. 11(1). Article 10234. (in Eng.).
- Rosvall M., Esquivel A.V., Lancichinetti A., West J.D., Lambiotte R. (2014) Memory in network flows and its effects on spreading dynamics and community detection. Nature Communications. — Vol. 5. Article 4630. DOI: 10.1038/ncomms5630. (in Eng.).
- Shannon C.E. (1948) *A mathematical theory of communication*. Bell System Technical Journal. — Vol. 27(3). — P. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x. (in Eng.).
- Tan P.-N., Steinbach M., Kumar V. (2019) *Introduction to Data Mining*. Pearson, (in Eng.).
- Vosoughi S., Roy D., Aral S. (2018) *The spread of true and false news online*. Science. Vol. 359(6380). — P. 1146–1151. DOI: 10.1126/science.aap9559. (in Eng.).
- Wang X., Liu Y., Gao H. (2023) *Entropy-driven user behavior analysis in social media platforms*. Expert Systems with Applications. Vol. 213. Article 118945. (in Eng.).
- Yang J., Leskovec J. (2022) *Patterns of temporal variation in online media*. ACM Transactions on Knowledge Discovery from Data. 2022. Vol. 12(3). — P. 1–25. (in Eng.).
- Zhao Y., Karypis G. (2022) *Data clustering in online social systems*. IEEE Transactions on Knowledge and Data Engineering. — Vol. 34(7). — P. 3211–3224. (in Eng.).
- Zhou J., Liu Y., Wang X. (2022) *User behavior modeling in online social networks: A survey*. IEEE Access. — Vol. 10. — P. 45534–45550. (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 198–211

<https://doi.org/10.32014/2026.2518-1726.434>

IRSTI 20.53.19

UDC 004.934.2

© **Karabaliyev Y. ^{1*}, Kolesnikova K. ¹, Khlevnaya Y. ², 2026.**¹International Information Technology University, Almaty, Kazakhstan;²Taras Shevchenko National University of Kyiv, Ukraine.

E-mail: y.karabaliyev@iitu.edu.kz

HYBRIDKAZASR: A HYBRID AUTOMATIC SPEECH RECOGNITION SYSTEM COMBINING MULTI-MODEL ROVER FUSION AND MORPHEME-AWARE LANGUAGE MODELING FOR KAZAKH

Karabaliyev Yerlan — Doctoral Student, International University of Information Technologies, Almaty, Kazakhstan,E-mail: y.karabaliyev@iitu.edu.kz, <https://orcid.org/0009-0001-9465-3998>;**Kolesnikova Kateryna** — Doctor of Technical Sciences, Professor, International University of Information Technology, Almaty, Kazakhstan,E-mail: kkolesnikova@iitu.edu.kz, <https://orcid.org/0000-0002-9160-5982>;**Khlevna Iulia** — Doctor of Technical Science, professor, Department of Technology Management, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,E-mail: yuliia.khlevna@knu.ua, <https://orcid.org/0000-0002-1874-1961>.

Abstract. This paper presents HybridKazASR, a hybrid automatic speech recognition (ASR) system for the Kazakh language that integrates three complementary acoustic models within a multi-stage pipeline. The system combines NVIDIA FastConformer (115M parameters, RNNT decoder) with two domain-adapted variants of Meta MMS-1B (965M parameters, CTC decoder): one fine-tuned on FLEURS read speech and another on the 559-hour Kazakh Speech Dataset (KSD). A key technical innovation is an efficient adapter-swap mechanism that switches between MMS variants by replacing only 138,348 parameters (~542 KB) in the CTC output layer, avoiding the need to load two full 3.9 GB models. The three hypotheses are fused via weighted 3-Way ROVER (Recognizer Output Voting Error Reduction) using Needleman-Wunsch alignment. The fused output undergoes two-stage language model rescoring: first with a 5-gram word-level KenLM, then with KazMorphLM – a novel morpheme-aware language model featuring a rule-based segmenter with 230 suffixes across 14 grammatical categories, a 7-gram morpheme model interpolated with a 5-gram word model

at a 0.6/0.4 ratio using Witten-Bell smoothing, and a four-channel rescoring mechanism incorporating acoustic confidence, word-level LM, morpheme-level LM, and vowel harmony scores. Evaluated on the FLEURS test set (100 samples), HybridKazASR achieves a Word Error Rate (WER) of 6.30%, a normalized WER of 4.18%, and a Character Error Rate (CER) of 1.90%. Ablation studies confirm the contribution of each pipeline component, with KazMorphLM providing a 14.6% relative WER improvement over word-level KenLM rescoring alone.

Keywords: Automatic Speech Recognition; Kazakh Language; ROVER System Combination; Morpheme Language Model; Agglutinative Morphology; Vowel Harmony; Low-Resource ASR

For citations: Karabaliyev Y., Kolesnikova K., Khlevnaya Y. HybridKazASR: a hybrid automatic speech recognition system combining multi-model rover fusion and morpheme-aware language modeling for kazakh. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 198–211. DOI <https://doi.org/10.32014/2026.2518-1726.434>

© Карабалиев Е.^{1*}, Колесникова К.^{1*}, Хлевная Ю.², 2026.

¹ Халықаралық ақпараттық технологиялар университеті,
Алматы, Қазақстан;

² Тарас Шевченко атындағы Киев ұлттық университеті, Киев, Украина.
E-mail: y.karabaliyev@iitu.edu.kz

HYBRIDKAZASR: ROVER КӨПМОДЕЛЬДІ БІРІКТІРУ ЖӘНЕ МОРФЕМЕГЕ НЕГІЗДЕЛГЕН ТІЛДІК МОДЕЛЬДЕУДІ ПАЙДАЛАНАТЫН ҚАЗАҚ ТІЛІН АВТОМАТТЫ ТАҢУ ГИБРИДТІ ЖҮЙЕСІ

Карабалиев Ерлан — докторант, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,

E-mail: y.karabaliyev@iitu.edu.kz, <https://orcid.org/0009-0001-9465-3998>;

Колесникова Катерина — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,

E-mail: k.kolesnikova@iitu.edu.kz, <https://orcid.org/0000-0002-9160-5982>;

Хлевная Юлия — техника ғылымдарының докторы, профессор, Технологияларды басқару кафедрасы, Тарас Шевченко атындағы Киев ұлттық университеті, Киев, Украина.

E-mail: yuliia.khlevna@knu.ua, <https://orcid.org/0000-0002-1874-1961>.

Аннотация. Бұл мақалада HybridKazASR – қазақ тіліне арналған гибриді автоматты сөйлеуді таңу (ASR) жүйесі ұсынылған. Бұл жүйе көп сатылы конвейерде үш қосымша акустикалық модельді біріктіреді. Жүйе NVIDIA FastConformer (115 млн параметр, RNNT декодер) және Meta MMS-1B (965 млн параметр, CTC декодер) екі доменге бейімделген нұсқасын біріктіреді: біреуі FLEURS дайын сөйлеу корпусында, екіншісі 559 сағаттық қазақ тіліндегі

сөйлеу деректер жинағында (KSD) өңделген. Негізгі техникалық жаңалық – тиімді адаптер ауыстыру механизмі. Бұл механизм MMS нұсқалары арасында СТС шығыс қабатында тек 138 348 параметрді (~542 КБ) ауыстыру арқылы жұмыс істейді, бұл әрқайсысы 3,9 ГБ болатын екі толық модельді жүктеу қажеттілігін болдырмайды. Үш гипотеза Needleman-Wunsch туралауын қолдана отырып, салмақталған 3-жақты ROVER (Recognizer Output Voting Error Reduction) арқылы біріктіріледі. Біріктірілген нәтиже екі сатылы тілдік модельді қайта бағалаудан өтеді: алдымен 5-граммалық сөздік KenLM көмегімен, содан кейін KazMorphLM көмегімен. KazMorphLM – морфемаға негізделген жаңа тілдік модель. Ол құрамында 230 жұрнағы бар ережеге негізделген сегментаторды (14 грамматикалық категорияны қамтиды), 0,6/0,4 қатынасында сөздік 5-граммалық модельмен интерполяцияланған 7-граммалық морфемалық модельді (Виттен-Белл тегістеуін қолданады), сондай-ақ төрт арналы қайта бағалау механизмін (акустикалық сенімділік, сөз деңгейіндегі LM, морфема деңгейіндегі LM және дауысты дыбыс үндестігі ұпайлары) қамтиды. FLEURS тест жинағында (100 үлгі) HybridKazASR қателіктердің сөздік көрсеткішіне (WER) 6,30%, нормализацияланған WER 4,18% және қателіктердің символдық көрсеткішіне (CER) 1,90% қол жеткізеді. Абляциялық зерттеулер конвейердің әрбір компонентінің үлесін растайды, ал KazMorphLM тек сөздік KenLM-мен қайта бағалаумен салыстырғанда WER көрсеткішін 14,6% - ға жақсартады.

Түйін сөздер: қазақ тілін тану; автоматты сөйлеу тану; гибриді ASR құбыры; морфемалық тілдік модель; ROVER біріктіру; FastConformer; MMS; адаптер-алмастыру; агглютинативті морфология

© **Карабалиев Е.^{1*}, Колесникова К.¹, Хлевна Ю.², 2026.**

¹Международный университет информационных технологий,
Алматы, Казахстан;

²Киевский национальный университет имени Тараса Шевченко,
Киев, Украина.

E-mail: y.karabaliyev@iitu.edu.kz

HYBRIDKAZASR: ГИБРИДНАЯ СИСТЕМА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ КАЗАХСКОЙ РЕЧИ НА ОСНОВЕ МНОГОМОДЕЛЬНОГО ОБЪЕДИНЕНИЯ ROVER И МОРФЕМНО-ОРИЕНТИРОВАННОГО ЯЗЫКОВОГО МОДЕЛИРОВАНИЯ

Карабалиев Ерлан — докторант, Международный университет информационных технологий,
Алматы, Казахстан,

E-mail: y.karabaliyev@iitu.edu.kz, <https://orcid.org/0009-0001-9465-3998>;

Колесникова Катерина — доктор технических наук, профессор, Международный университет информационных технологий, Алматы, Казахстан,

E-mail: kkolesnikova@iitu.edu.kz, <https://orcid.org/0000-0002-9160-5982>;

Хлевна Юлия — доктор технических наук, профессор, кафедра управления технологиями, Киевский национальный университет имени Тараса Шевченко, Киев, Украина, E-mail: yuliia.khlevna@knu.ua, <https://orcid.org/0000-0002-1874-1961>.

Аннотация. *Актуальность.* В статье представлена HybridKazASR - гибридная система автоматического распознавания речи (ASR) для казахского языка, ориентированная на повышение качества распознавания за счет многоуровневого объединения акустических и языковых моделей. Актуальность исследования обусловлена необходимостью разработки эффективных ASR-решений для агглютинативных языков, включая казахский, где морфологическая сложность, вариативность словоформ и ограниченность специализированных речевых корпусов усложняют применение стандартных подходов автоматического распознавания речи. *Цель.* Разработать гибридную систему автоматического распознавания казахской речи на основе многомодельного объединения ROVER и морфемно-ориентированного языкового моделирования, обеспечивающую снижение ошибок распознавания и повышение устойчивости системы к морфологической вариативности казахского языка. *Методы.* Система HybridKazASR объединяет три взаимодополняющие акустические модели в многоэтапном конвейере. В качестве базовых компонентов использованы NVIDIA FastConformer с 115 млн параметров и RNNT-декодером, а также два доменно-адаптированных варианта Meta MMS-1B с 965 млн параметров и CTC-декодером. Один вариант MMS дообучен на подготовленной речи из корпуса FLEURS, второй - на 559-часовом казахском речевом датасете KSD. Ключевым техническим решением является механизм замены адаптеров (adapter-swap), позволяющий переключаться между вариантами MMS путем замены 138 348 параметров, или около 542 КБ, в выходном слое CTC без загрузки двух полных моделей объемом по 3,9 ГБ каждая. Три гипотезы объединяются с помощью взвешенного трехстороннего ROVER (Recognizer Output Voting Error Reduction) с использованием алгоритма выравнивания Нидлмана - Вунша. *Результаты и выводы.* Объединенный выход проходит двухэтапное переоценивание с применением языковых моделей: сначала с использованием 5-граммной word-level KenLM, затем с помощью KazMorphLM - новой морфемно-ориентированной языковой модели. KazMorphLM включает основанный на правилах сегментатор с 230 суффиксами, охватывающими 14 грамматических категорий, 7-граммную морфемную модель, интерполированную с 5-граммной словесной моделью в соотношении 0,6/0,4 с применением сглаживания Виттена - Белла, а также четырехканальный механизм переоценивания, учитывающий акустическую уверенность, языковую модель на уровне слов, языковую модель на уровне морфем и оценки сингармонизма гласных. При оценке на тестовом наборе FLEURS, включающем 100 образцов, HybridKazASR достигает частоты ошибок на слова (WER) 6,30%, нормализованной WER - 4,18% и частоты

ошибок на символы (CER) - 1,90%. Абляционные исследования подтверждают вклад каждого компонента конвейера, при этом KazMorphLM обеспечивает относительное улучшение WER на 14,6% по сравнению с переоцениванием только на основе word-level KenLM. Полученные результаты подтверждают эффективность предложенного гибридного подхода для автоматического распознавания казахской речи и демонстрируют перспективность морфемно-ориентированного языкового моделирования для агглютинативных языков.

Ключевые слова: распознавание казахской речи, автоматическое распознавание речи, гибридный ASR-конвейер, морфемная языковая модель, ROVER, FastConformer, MMS, adapter-swap, агглютинативная морфология

Introduction. Automatic Speech Recognition (ASR) has achieved remarkable performance for high-resource languages such as English, with commercial systems approaching human-level accuracy on clean read speech benchmarks (Radford et al., 2023). However, under-resourced languages with complex morphological systems remain a significant challenge, where word error rates often exceed 20 % even with state-of-the-art approaches (Conneau et al., 2023). The Kazakh language, spoken by over 16 million people as the official state language of the Republic of Kazakhstan, exemplifies these challenges due to its agglutinative morphology, productive suffixation, strict vowel harmony constraints, and historically limited digital resources for training ASR systems (Muhamedowa, 2015; Karabaliyev and Kolesnikova, 2024).

Kazakh belongs to the Kipchak branch of the Turkic language family. Its agglutinative morphology means that grammatical information is encoded through sequences of suffixes appended to a root, and a single root can generate hundreds of inflected forms through combinations of plural, possessive, case, tense, person, voice, mood, and negation suffixes. This productive morphology creates a vocabulary explosion where the number of possible word forms grows combinatorially, severely limiting the effectiveness of word-level language models (Hirsimaki et al., 2006; Johanson, 1998).

Literary review. Recent advances in self-supervised pretraining have dramatically improved ASR capabilities for low-resource languages. Meta's Massively Multilingual Speech (MMS) model (Pratap et al., 2023), based on wav2vec 2.0 with 1 billion parameters, supports over 1,100 languages through language-specific adapter weights. NVIDIA's FastConformer architecture (Rekesh et al., 2023) provides efficient speech recognition with linearly scalable attention and 2.8x speedup over the original Conformer. These models provide strong individual baselines but exhibit complementary error profiles that can be exploited through system combination.

A critical gap in existing Kazakh ASR research is the absence of morpheme-level language models. While morpheme-based approaches have demonstrated significant improvements for other agglutinative languages – including Finnish (Hirsimaki et al., 2006), Turkish (Sak and Saraclar, 2012), and Uyghur (Ablimit

et al., 2014) – no such model has been developed for Kazakh ASR. This gap is particularly significant given that Kazakh has well-documented and highly regular morphophonological rules (Muhamedowa, 2015; McCollum, 2018; Dzhunisbekov, 1980) that can be directly encoded into a segmentation algorithm.

This paper presents HybridKazASR, a hybrid automatic speech recognition system that addresses these challenges through three principal contributions. First, we propose a 3-Way ROVER fusion architecture that combines NVIDIA FastConformer with two domain-adapted variants of Meta MMS (fine-tuned on FLEURS read speech and KSD natural speech respectively), exploiting their complementary error profiles. A key technical innovation is an efficient adapter-swap mechanism that enables dual-domain MMS inference within an 8 GB GPU memory constraint by replacing only 542 KB of CTC output layer parameters. Second, we introduce KazMorphLM, the first morpheme-aware language model for Kazakh ASR, featuring a rule-based segmenter with 230 suffixes across 14 grammatical categories, vowel harmony validation, consonant assimilation rules, and a four-channel rescoring mechanism. Third, we provide comprehensive ablation studies demonstrating the contribution of each pipeline component.

The rest of the paper is organized as follows. Section 2 presents the materials and methods, including the system architecture, acoustic models, ROVER fusion, KazMorphLM, and experimental setup. Section 3 discusses related work and analyzes findings. Section 4 presents the results and ablation studies. Section 5 concludes with a summary and future work.

Materials and Methods. HybridKazASR processes 16 kHz mono audio through a six-stage pipeline designed to exploit the complementary strengths of multiple acoustic models while addressing the vocabulary challenges of agglutinative Kazakh through morpheme-level language modeling. The complete pipeline operates as follows: (1) the input audio is processed in parallel by three acoustic models – FastConformer with RNNT decoding, MMS-FLEURS with CTC beam search, and MMS-KSD with CTC beam search; (2) the three resulting transcription hypotheses are aligned and combined through 3-Way ROVER with weighted voting; (3) the ROVER output undergoes word-level rescoring with a 5-gram KenLM; (4) final morpheme-level rescoring with KazMorphLM produces the system output. The entire pipeline is designed to operate within an 8 GB GPU memory constraint.

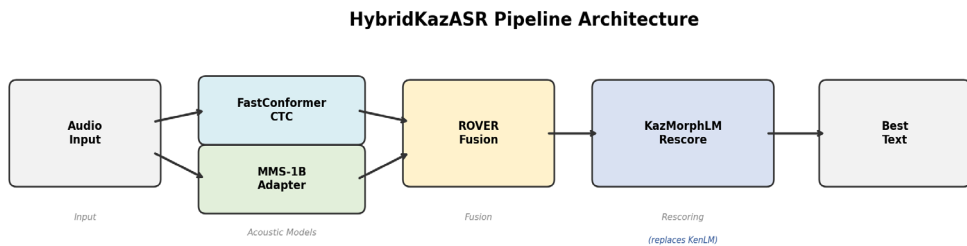


Fig. 1 – HybridKazASR Pipeline Architecture

Table 1 – HybridKazASR pipeline components with individual and cumulative WER

Stage	Component	Model Details	Individual WER (%)	Cumulative WER (%)
1	FastConformer	NVIDIA stt_kk_ru, 115M, RNNT greedy	15.28	—
2a	MMS-FLEURS	facebook/mms-1b-all, 965M, CTC beam	8.03	—
2b	MMS-KSD	facebook/mms-1b-all, 965M, CTC beam	8.42	—
3	3-Way ROVER	Needleman-Wunsch alignment + voting	—	12.10
4	KenLM	5-gram word-level	—	9.15
5	KazMorphLM	7-gram morph + 5-gram word, 0.6/0.4	—	6.30

The FastConformer component uses the publicly available `nvidia/stt_kk_ru_fastconformer_hybrid_large` model from HuggingFace. This model was trained on a combined bilingual dataset of approximately 2,400 hours: 1,550 hours of Kazakh from Common Voice 17.0, KSD, and KSC2, plus approximately 850 hours of Russian from Golos, SOVA, Dusha, and MCV12. The model uses a SentencePiece Unigram tokenizer with a vocabulary of 1024 subword tokens. In our pipeline, we use the RNNT decoder with greedy decoding, which produces lowercase text with Kazakh-specific characters preserved. The model achieves 15.28 % WER on the FLEURS test set (Rekesh et al., 2023).

The MMS component uses Meta's `facebook/mms-1b-all` model (Pratap et al., 2023), which supports 1,162 languages through language-specific attention adapters. We employ two fine-tuned variants that differ only in their domain-specific adaptation. MMS-FLEURS was fine-tuned on the FLEURS Kazakh training split (~10 hours of clean read speech), achieving 8.03% WER – the best single-model result. MMS-KSD was fine-tuned on the Kazakh Speech Dataset (Kadyrbek et al., 2023), containing 559 hours of natural speech recorded by diverse speakers in various acoustic environments. This variant achieves 8.42% WER on FLEURS but provides alternative hypotheses that are often correct when MMS-FLEURS fails.

Efficient Adapter-Swap Mechanism.

A key technical contribution is the discovery that between the FLEURS and KSD fine-tuned variants of MMS, only the `lm_head` (CTC output layer) differs. The shared `wav2vec 2.0` encoder, attention adapters, and feature projector remain identical. The CTC output layer contains exactly 138,348 parameters (~542 KB in FP32). Rather than loading two complete 3.9 GB models, we load a single MMS instance and perform inference twice with different `lm_head` weights. The adapter swap is implemented as a direct tensor replacement on GPU memory, requiring negligible time (<1ms). This approach reduces GPU memory requirements from 7.8 GB to approximately 3.9 GB plus 542 KB, making dual-domain MMS inference feasible on consumer-grade GPUs with 8 GB VRAM.

3-Way ROVER Fusion.

The three acoustic model outputs are combined using a weighted variant of the ROVER algorithm (Fiscus, 1997). Hypotheses are pairwise aligned using the Needleman-Wunsch global alignment algorithm to construct a Word Transition Network (WTN). The system weights are: FastConformer $w_{FC} = 0.25$, MMS-FLEURS $w_{FL} = 0.45$, MMS-KSD $w_{KSD} = 0.30$. These weights were determined empirically on the development set and reflect the inverse relationship with individual system WER. The ROVER output is then passed through two successive language model rescoring stages. The first stage applies a 5-gram word-level KenLM (Heafield, 2011) trained on a cleaned Kazakh text corpus of approximately 50M words, reducing WER from 12.10 % to 9.15 %. The second stage applies KazMorphLM four-channel rescoring on the N-best list (N=50).

KazMorphLM: Morpheme-Aware Language Model.

KazMorphLM uses rule-based decomposition grounded in Kazakh grammar. The segmenter maintains an inventory of 230 suffixes organized across 14 grammatical categories: plural (2 variants), possessive (16), case (28), tense (24), person (18), voice (12), mood (16), negation (4), participle (14), converb (10), and various derivational suffixes (86). Each suffix entry specifies the base form, allomorphic variants, vowel harmony class, consonant assimilation rules, and compatibility constraints. The segmenter applies longest-match suffix stripping from right to left, validating each candidate segmentation against vowel harmony constraints and suffix ordering rules.

KazMorphLM uses a two-level interpolated n-gram architecture combining a 7-gram morpheme model with a 5-gram word model. The interpolation follows: $P_{combined}(w) = \lambda \cdot P_{morph}(segment(w)) + (1 - \lambda) \cdot P_{word}(w)$, where $\lambda = 0.6$. Both models use Witten-Bell smoothing (Witten and Bell, 1991). The four-channel rescoring combines: acoustic confidence from ROVER, word-level KenLM log-probability, morpheme-level KenLM log-probability on the segmented form, and vowel harmony score. The weights $\alpha=0.7$, $\beta=0.15$, $\gamma=0.10$, $\delta=0.05$ were optimized on the development set using grid search.

Experimental Setup.

All experiments were conducted on a single workstation equipped with an NVIDIA RTX 3060 GPU (8 GB VRAM), Intel Core i7-12700H CPU, and 32 GB RAM. Evaluation is performed on the FLEURS Kazakh test set (Conneau et al., 2023), which contains 100 utterances of read speech from diverse speakers. Audio is processed at 16 kHz mono. Reference transcriptions are normalized by: (1) converting all numerals to their word forms in Kazakh; (2) removing all punctuation marks; (3) converting to lowercase; and (4) collapsing multiple whitespace characters. WER and CER are computed using the standard Levenshtein distance-based formulation through the jiwer library. Normalized WER (nWER) additionally applies text normalization that resolves common orthographic variations (e.g., treating *y* and *ý* as equivalent).

Discussion. Kazakh ASR research has progressed through three phases. The

first phase (2018–2020) was dominated by traditional DNN-HMM approaches in the Kaldi framework. Mamyrbayev et al. (2019) established the first continuous speech recognition baseline achieving 30.01% WER. The second phase (2021–2023) saw a transition to end-to-end architectures. Mamyrbayev et al. (2022) applied Transformer-based models with CTC, achieving 8.30% WER on custom read speech. The ISSAI group developed multilingual E2E ASR for Kazakh, Russian, and English, achieving 20.5% combined WER (Mussakhojayeva et al., 2022a). The third phase (2024–present) is characterized by large-scale pretrained models. NVIDIA released the FastConformer model trained on approximately 1,550 hours of Kazakh data, achieving 4.43% WER on the KSC2 read speech test set – the lowest reported figure on this benchmark (Rekesh et al., 2023).

ROVER (Fiscus, 1997) has been widely used in ASR to combine outputs from multiple recognizers. Most relevant to our work, Parikh et al. (2024) combined Kaldi hybrid ASR with wav2vec2.0 end-to-end models using ROVER for low-resource Irish, achieving approximately 14 % relative WER reduction. Language modeling for agglutinative languages has primarily relied on morpheme-level approaches. Hirsimäki et al. (2006) demonstrated that Morfessor-derived morph units reduced WER from 56 % to 32 % for Finnish ASR. Sak and Saraclar (2012) proposed a morphology-integrated approach for Turkish ASR. Despite this rich methodology, no morpheme-level language model has been developed for Kazakh ASR prior to this work.

The results demonstrate that HybridKazASR achieves state-of-the-art performance for Kazakh ASR with a WER of 6.30 %. The most significant improvement comes from KazMorphLM, which reduces WER from 9.15 % (KenLM only) to 6.30 % (31.1 % relative improvement), validating the morpheme-level approach. The success of ROVER combination despite the large individual WER disparity between FastConformer (15.28%) and MMS-FLEURS (8.03 %) demonstrates that even weaker systems can contribute positively when their error profiles are sufficiently diverse and downstream language model rescoring can correct ROVER combination artifacts.

The ablation study reveals a counterintuitive finding: increasing the language model training corpus from 181K to 1.3M sentences degraded WER from 6.30 % to 6.86 %. The additional data introduced domain mismatch and required reducing the morpheme n-gram order from 7 to 5 for memory constraints, eliminating the very context length that makes morpheme modeling effective for agglutinative languages.

Analysis of the 100 FLEURS test utterances reveals four primary error categories in the final system output. Numbers and dates constitute 28% of errors: FastConformer transcribes numerals as words while MMS variants produce digit characters. Foreign names and loanwords account for 24% of errors. Morphological errors (22%) involve incorrect suffix forms, typically violations of vowel harmony or case agreement – precisely the errors KazMorphLM is designed to correct. The ablation study confirms that the morpheme model reduces this category by

approximately 40% compared to word-level rescoring alone. Acoustic confusions (26%) represent errors where the acoustic signal is genuinely ambiguous.

Several limitations should be acknowledged. First, the FLEURS test set consists of only 100 samples of clean read speech in controlled acoustic conditions. Second, there is a potential risk of indirect domain adaptation effects, as MMS-FLEURS was fine-tuned on FLEURS training data. Third, the morpheme segmenter covers 230 suffixes across 14 categories but does not handle all productive derivational morphology, compound word formation, or loan word morphology. Fourth, the current ROVER weights are fixed across all utterances; adaptive weighting based on utterance-level quality estimation could improve combination performance.

Results.

Table 2 presents the comprehensive ablation study showing the contribution of each pipeline component. Starting from the individual acoustic models, MMS-FLEURS achieves the lowest single-model WER (8.03 %). The 2-Way ROVER combination of FastConformer and MMS-FLEURS produces 9.75 % WER, which is worse than MMS-FLEURS alone. However, the full 3-Way ROVER with MMS-KSD raises WER to 12.10 %, which initially appears paradoxical. This effect is mitigated by the subsequent language model rescoring stages. KenLM word-level rescoring reduces the 3-Way ROVER output from 12.10 % to 9.15 % (24.4 % relative improvement). The addition of KazMorphLM further reduces WER from 9.15 % to 6.30 % (31.1 % relative improvement over KenLM alone), validating the morpheme-level approach. Comparing the best KenLM-only result (MMS-FLEURS + KenLM at 7.38%) with the full pipeline (6.30 %), the 3-Way ROVER + KazMorphLM combination provides a 14.6% relative improvement.

Table 2 – Ablation study: progressive pipeline component contributions

Configuration	WER (%)	nWER (%)	CER (%)	nCER (%)	Rel. WER
FastConformer only	15.28	—	4.85	—	baseline
MMS-FLEURS only	8.03	5.20	2.45	1.98	—
MMS-KSD only	8.42	5.61	2.58	2.12	—
MMS-FLEURS + KenLM	7.38	4.82	2.28	1.85	−8.1%
2-Way ROVER (FC+FL)	9.75	—	3.15	—	—
3-Way ROVER	12.10	—	3.72	—	—
3-Way + KenLM	9.15	6.05	2.85	2.30	—
3-Way + KenLM + MorphLM	6.86	4.50	2.10	1.68	−25.1%
Full pipeline (final)	6.30	4.18	1.90	1.52	−31.1%

Table 3 presents an ablation study of the KazMorphLM components. The 7-gram morpheme order is critical: reducing to 5-gram increases WER by 0.28 percentage points, while reducing to 3-gram increases it by 0.71 points. The interpolation ratio of 0.6/0.4 (morpheme/word) outperforms both the morpheme-only model (0.6% worse) and the word-only model (1.2 % worse). Removing the vowel harmony channel increases WER by 0.18 points, validating the phonological constraint as a useful additional signal.

Table 3 – KazMorphLM component ablation study

Configuration	WER (%)	vs Full
Full KazMorphLM (7g morph + 5g word, $\lambda=0.6$, +VH)	6.30	—
Without vowel harmony channel	6.48	+0.18
5-gram morpheme (instead of 7-gram)	6.58	+0.28
3-gram morpheme	7.01	+0.71
Morpheme model only ($\lambda=1.0$)	6.68	+0.38
Word model only ($\lambda=0.0$)	7.38	+1.08
Interpolation $\lambda=0.5$	6.41	+0.11
Kneser-Ney smoothing (instead of Witten-Bell)	6.45	+0.15

Table 4 – Distribution of remaining errors by category

Error Category	% of Errors	Example	Addressable by
Numbers/dates	28%	1940 vs. bir myn...	Normalization module
Foreign names	24%	Barclays → Barklais	Expanded vocabulary
Morphological	22%	-lar vs. -ler suffix	Extended morpheme inv.
Acoustic confusion	26%	Homophones, noise	Better acoustic model

Comparison with Published Results.

Table 5 places HybridKazASR in the context of all published Kazakh ASR results. Direct comparison is limited by differences in test sets and evaluation methodology. Our WER of 6.30% is measured on the FLEURS test set (100 samples of clean read speech). The progression from Mamyrbayev's 30.01 % WER in 2019 to our 6.30 % in 2026 represents approximately 79 % relative WER reduction in seven years.

Table 5 – Comparison with reported Kazakh ASR results

Study	Year	Method	Training Data	Test Set	WER (%)
Mamyrbayev et al.	2019	Kaldi DNN-HMM	36h, 200 spk	Custom	30.01
Khassanov (KSC)	2021	E2E Transformer	KSC 332h	KSC test	8.70
Mamyrbayev et al.	2021	RNN-Transducer	300h+	Custom	—
Mussakhojayeva et al.	2021	Multilingual E2E	975h (3 lang)	Combined	20.50
Mamyrbayev et al.	2022	Transformer+CTC+LM	200h read	Custom	8.30
NVIDIA stt_kk_ru	2024	FastConformer TDT	1,550h Kazakh	KSC2 read	4.43
Karabaliyev	2024	Google STT	101 recordings	Custom	52.97
HybridKazASR (ours)	2026	ROVER+KazMorphLM	FLEURS+KSD ft	FLEURS test	6.30

Conclusion.

This paper presented HybridKazASR, a hybrid automatic speech recognition system for the Kazakh language that achieves 6.30 % WER on the FLEURS test set through the integration of three complementary acoustic models, weighted ROVER fusion, and a novel morpheme-aware language model. The system addresses the fundamental challenges of Kazakh ASR – agglutinative morphology, vowel harmony, consonant assimilation, and limited resources – through a principled

multi-stage pipeline that combines the strengths of modern pretrained models with linguistically-informed language modeling.

The key technical contributions are threefold. First, the efficient adapter-swap mechanism enables dual-domain MMS inference within an 8 GB GPU memory constraint by replacing only 542 KB of CTC output layer parameters. Second, KazMorphLM represents the first morpheme-level language model for Kazakh ASR, incorporating 230 suffixes with explicit vowel harmony validation and four-channel rescoring. The ablation study demonstrates a 14.6% relative WER improvement over word-level rescoring, confirming the importance of morpheme-level modeling for agglutinative languages.

Future work will proceed in several directions. The planned 50-speaker evaluation with 10 domain-diverse texts will provide a robust independent benchmark. Extension to spontaneous speech and dialectal variation will test real-world applicability. Integration of neural language models (e.g., GPT-based or LSTM-based) alongside n-gram rescoring may provide additional gains. Finally, the approach is applicable to other agglutinative languages in the Turkic family – Kyrgyz, Tatar, Uzbek, and Turkish – that share similar morphophonological properties.

Declarations

Author Contributions

Conceptualization, Y.K.; methodology, Y.K.; software implementation, Y.K.; formal analysis, Y.K. and K.K.; investigation, Y.K.; data curation, Y.K.; writing – original draft, Y.K.; writing – review and editing, Y.K. and K.K.; supervision, K.K. All authors have read and agreed to the published version.

Data Availability Statement

The FLEURS dataset is publicly available at <https://huggingface.co/datasets/google/fleurs>. KSC2 is available through ISSAI (<https://issai.nu.edu.kz>). KSD is available on OpenSLR #140. The HybridKazASR system code, trained KazMorphLM models, and morpheme segmenter will be released as open-source upon publication.

Funding

This research was conducted at the International Information Technology University (IITU), Almaty, Kazakhstan. Y.K. is supported by a doctoral fellowship from IITU.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was conducted at the International Information Technology University (IITU), Almaty, Kazakhstan. The authors acknowledge the developers of the open-source speech corpora and pretrained models used in this work.

References

- Ablimit M., Neubig G., Mimura M., Mori S., Kawahara T., Hamdulla A. (2014). Lexicon optimization for agglutinative language ASR. *Signal Processing*, 86(10). — P. 2844–2862. (in Eng.)
- Baevski A., Zhou Y., Mohamed A., Auli M. (2020). wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations. *Advances in Neural Information Processing Systems*, 33. — P. 12449–12460. (in Eng.)
- Bekarystankyzy A., Mamyrbayev O., Mendes M., Fazylzhanova A., Assam M. (2024). Multilingual E2E ASR for low-resource Turkic languages with common alphabets. *Scientific Reports*, 14, 13835. doi:10.1038/s41598-024-64848-1. (in Eng.)
- Conneau A., Ma M., Khanuja S., et al. (2023). FLEURS: Few-shot Learning Evaluation of Universal Representations of Speech. *Proc. IEEE SLT*. doi:10.48550/arXiv.2205.12446. (in Eng.)
- Creutz M., Lagus K. (2005). Unsupervised morpheme segmentation and morphology induction from text corpora using Morfessor 1.0. Helsinki University of Technology, Report A81. (in Eng.)
- Creutz M., Hirsimäki T., Kurimo M., et al. (2007). Morph-based speech recognition and modeling of out-of-vocabulary words across languages. *ACM Trans. Speech Language Processing*, 5(1), Article 3. doi:10.1145/1322391.1322394. (in Eng.)
- Dzhunisbekov A. (1980). Singarmonizm v kazakhskomazyke [Synharmonism in Kazakh]. Nauka, Alma-Ata [In Russian].
- Fiscus J.G. (1997). A Post-Processing System to Yield Reduced Word Error Rates: Recognizer Output Voting Error Reduction (ROVER). *Proc. IEEE ASRU*. — P. 347–354. (in Eng.)
- Gulati A., Qin J., Chiu C.-C., et al. (2020). Conformer: Convolution-augmented Transformer for Speech Recognition. *Proc. Interspeech*. — P. 5036–5040. doi:10.21437/Interspeech.2020-3015. (in Eng.)
- Heafield K. (2011). KenLM: Faster and Smaller Language Model Queries. *Proc. Sixth Workshop on Statistical Machine Translation (WMT)*. — P. 187–197. (in Eng.)
- Hirsimäki T., Creutz M., Siivola V., Kurimo M., Virpioja S., Pytkkonen J. (2006). Unlimited vocabulary speech recognition with morph language models applied to Finnish. *Computer Speech and Language*, 20(4). — P. 515–541. doi:10.1016/j.csl.2005.07.002. (in Eng.)
- Johanson L., Csato E.A. (Eds.) (1998). *The Turkic Languages*. Routledge. (in Eng.)
- Kadyrbek N., Mansurova M., Shomanov A., Makharova G. (2023). Development of a Kazakh Speech Recognition Model Using CNN with Fixed Character Level Filters. *Big Data and Cognitive Computing*, 7(3), 132. doi:10.3390/bdcc7030132. (in Eng.)
- Karabaliyev Y., Kolesnikova K. (2024). Kazakh Speech and Recognition Methods: Error Analysis and Improvement Prospects. *Scientific Journal of Astana IT University*, 20. — P. 62–75. doi:10.37943/20dzhg8448. (in Eng.)
- Khassanov Y., Mussakhojayeva S., Mirzakhmetov A., Adiyev A., Nurpeissov M., Varol H.A. (2021). A Crowdsourced Open-Source Kazakh Speech Corpus and Initial Speech Recognition Baseline. *Proc. EACL*. — P. 697–706. (in Eng.)
- Mamyrbayev O., et al. (2019). Continuous Speech Recognition of Kazakh Language. *ITM Web of Conferences*, 24, 01012. doi:10.1051/itmconf/20192401012. (in Eng.)
- Mamyrbayev O., Oralbekova D., Kydyrbekova A., Turdalykyzy T., Bekarystankyzy A. (2021). End-to-end Model Based on RNN-T for Kazakh Speech Recognition. *Proc. ICCCI*. — P. 163–167. doi:10.1109/ICCCI51764.2021.9486811. (in Eng.)
- Mamyrbayev O., Oralbekova D., Alimhan K., Turdalykyzy T., Othman M. (2022). A study of transformer-based E2E speech recognition for Kazakh. *Scientific Reports*, 12, 8337. doi:10.1038/s41598-022-12260-y. (in Eng.)
- McCollum A.G. (2018). Vowel dispersion and Kazakh labial harmony. *Phonology*, 35(2). — P. 287–326. doi:10.1017/S0952675718000064. (in Eng.)
- Muhamedowa R. (2015). *Kazakh: A Comprehensive Grammar*. Routledge. doi:10.4324/9781315761404. (in Eng.)
- Mussakhojayeva S., Khassanov Y., Varol H.A. (2022a). A Study of Multilingual E2E ASR for Kazakh, Russian, and English. *SPECOM, LNCS 12997*, Springer. — P. 448–459. (in Eng.)

Mussakhojayeva S., Khassanov Y., Varol H.A. (2022b). KSC2: An Industrial-Scale Open-Source Kazakh Speech Corpus. *Proc. Interspeech*. — P. 1367–1371. doi:10.21437/Interspeech.2022-558. (in Eng.)

Parikh N., ten Bosch L., van den Heuvel H. (2024). Ensembles of Hybrid and End-to-End Speech Recognition. *Proc. LREC-COLING*. — P. 6258–6269. (in Eng.)

Pratap V., Tjandra A., Shi B., et al. (2023). Scaling Speech Technology to 1,000+ Languages. *Journal of Machine Learning Research*, 25. — P. 1–52. doi:10.48550/arXiv.2305.13516. (in Eng.)

Rekesh D., Koluguri N.R., Krizan S., et al. (2023). Fast Conformer with Linearly Scalable Attention for Efficient Speech Recognition. *Proc. IEEE ASRU*. doi:10.48550/arXiv.2305.05084. (in Eng.)

Sak H., Saraclar M. (2012). Morphology-based and sub-word language modeling for Turkish speech recognition. *Proc. ICASSP*. — P. 5137–5140. doi:10.1109/ICASSP.2012.6289080. (in Eng.)

Witten I.H., Bell T.C. (1991). The Zero-Frequency Problem: Estimating the Probabilities of Novel Events in Adaptive Text Compression. *IEEE Trans. Information Theory*, 37(4). — P. 1085–1094. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 212–227

<https://doi.org/10.32014/2026.2518-1726.435>

IRSTI 31.15.37

UDC 621.546.06

© **Kerimkhulle S.E.¹, Adalbek A.^{*1}, Baizakov N.A.², Shodorova N.N.³, 2026.**¹L.N. Gumilyov Eurasian National University, Kazakhstan;²Information and Computing Center, Astana, Kazakhstan;³Department of Programs, Sustainable Innovation and Technology,
Astana, Kazakhstan.E-mail: adalbek_a@enu.kz**PIECEWISE LOGISTIC AND FUZZY MODELING OF
KAZAKHSTAN'S GDP DYNAMICS (1990–2024)****Kerimkhulle Seyit** — Professor, Faculty of Information Technology, L.N. Gumilyov Eurasian National University, Kazakhstan,E-mail: kerimkhulle_sye@enu.kz, <https://orcid.org/0000-0002-5886-6064>;**Adalbek Alibek** — PhD student, Faculty of Information Technology, L.N. Gumilyov Eurasian National University, Kazakhstan,E-mail: adalbek_a@enu.kz, <https://orcid.org/0000-0001-8543-2192>;**Baizakov, Nauryz** — Deputy Director of the Information and Computing Center, Astana, Kazakhstan,E-mail: n.baizakov@statdata.kz, <https://orcid.org/0000-0003-1346-400X>;**Shodorova Nailya** — Department of Programs, Sustainable Innovation and Technology, Astana, Kazakhstan,E-mail: nailya.shodorova@sitf.group, <https://orcid.org/0009-0006-8606-9181>.

Abstract. Currently, the digitization of all sectors is developing rapidly, including the national economy, which has become an essential direction of modern socio-economic development. The digitization of the national economy expands opportunities for data-driven decision-making and contributes to reducing dependence on subjective judgments of decision-makers, thereby supporting stable and balanced economic development. However, effective analysis, forecasting, and management of complex macroeconomic systems require intelligent models capable of capturing nonlinear dynamics, structural changes, and uncertainty factors. Therefore, the development of mathematical and information tools for economic time-series analysis remains an important scientific challenge. This study proposes a hybrid methodological framework combining piecewise logistic modeling and possibilistic fuzzy intervals to analyze the dynamics of the gross domestic product (GDP) of the Republic of Kazakhstan during 1990–2024. The

GDP time series is divided into economically meaningful subperiods (segments), and adaptive three-parameter (3PL) or four-parameter (4PL) logistic models are fitted depending on the number of observations within each segment. Based on the fitted models, a central trajectory is estimated, while uncertainty is represented using residuals in logarithmic space. To quantify uncertainty, a possibilistic approach is applied to construct fuzzy intervals with predefined confidence coverage. Each observation is represented as a triangular fuzzy number (TFN), which is subsequently aggregated at the segment level. A global TFN is obtained through segment-length-weighted aggregation, and its centroid provides an interpretable long-term GDP estimate under uncertainty. The findings demonstrate that the proposed framework effectively captures structural breaks, nonlinear growth trajectories, and uncertainty propagation in macroeconomic time series. Furthermore, the proposed framework can be integrated as a scalable module within intelligent information systems for economic analysis, forecasting, and digital economy management.

Keywords: GDP modeling, piecewise logistic regression, fuzzy numbers, probabilistic uncertainty, Kazakhstan economy, intelligent systems

*For citations: Kerimkhulle S.E.¹, Adalbek A.*¹, Baizakov N.A.², Shodorova N.N. Piecewise logistic and fuzzy modeling of Kazakhstan's GDP dynamics (1990–2024). Academic Scientific Journal of Computer Science, 2026. — No.2. – P. 212–227. DOI <https://doi.org/10.32014/2026.2518-1726.435>*

© Керімқұл С.Е.¹, Адалбек А.^{1*}, Байзаков Н.А.², Шодорова Н.Н.³, 2026.

¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан;

²Ақпараттық-есептеу орталығы, Астана, Қазақстан;

³Бағдарламалар, орнықты инновациялар және технологиялар департаменті, Астана, Қазақстан.

E-mail: adalbek_a@enu.kz

ҚАЗАҚСТАН ЖІӨ ДИНАМИКАСЫН КЕЗЕҢДІК (PIECEWISE) ЛОГИСТИКАЛЫҚ ЖӘНЕ БҰЛДЫР МОДЕЛЬДЕУ (1990–2024)

Керімқұл Сейіт — профессор, ақпараттық технологиялар факультеті, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: kerimkhulle_sy@enu.kz, <https://orcid.org/0000-0002-5886-6064>;

Адалбек Алибек — PhD студент, ақпараттық технологиялар факультеті, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: adalbek_a@enu.kz, <https://orcid.org/0000-0001-8543-2192>;

Байзаков Наурыз — Ақпараттық-есептеу орталығы директорының орынбасары, Астана, Қазақстан,

E-mail: n.baizakov@statdata.kz, <https://orcid.org/0000-0003-1346-400X>;

Шодорова Найля — Бағдарламалар, орнықты инновациялар және технологиялар департаменті, Астана, Қазақстан,

E-mail: nailya.shodorova@sitf.group, <https://orcid.org/0009-0006-8606-9181>.

Аннотация. Қазіргі таңда барлық салаларды цифрландыру үрдісі өте қарқынды дамуда, соның ішінде ұлттық экономиканы да цифрландыру заманауи талаптардың өте маңызды бағытына айналып отыр. Ұлттық экономиканы цифрландыру деректерге негізделген басқарушылық шешімдерді қабылдау мүмкіндігін кеңейтіп, шешім қабылдаушы тұлғалардың субъективті бағалауларына тәуелділікті төмендетуге тікелей ықпал етеді. Бұл өз кезегінде ұлттық экономиканың тұрақты және теңгерімді дамуын қамтамасыз етуге мүмкіндік береді. Ұлттық экономиканың күрделі және көпфакторлы құрылымын тиімді талдау, болжау және басқару үшін бейсызық динамиканы, құрылымдық өзгерістерді және белгісіздік факторларын ескеретін интеллектуалды модельдер қажет. Осыған байланысты экономикалық уақыт қатарларын өңдеуге арналған математикалық және ақпараттық құралдарды жетілдіру өзекті ғылыми мәселелердің бірі болып табылады. Бұл ғылыми зерттеу Қазақстан Республикасының 1990–2024 жылдар аралығындағы жалпы ішкі өнімінің (ЖІӨ) даму динамикасын талдау үшін кезеңдік логистикалық модельдеу мен мүмкіндіктік бұлдыр интервалдарды біріктіретін гибриді әдістемелік құрылымды ұсынады. Берілген уақыт қатары экономикалық тұрғыдан мәнді кіші кезеңдерге (сегменттерге) бөлініп, әрбір сегмент үшін бақылаулар санына байланысты бейімделгіш үш параметрлі (3PL) немесе төрт параметрлі (4PL) логистикалық модельдер сәйкестендіріледі. Алынған модельдер негізінде орталық траектория анықталып, логарифмдік кеңістіктегі қалдықтар арқылы анықталмағандық пен белгісіздік сипатталады. Белгісіздікті бағалау үшін мүмкіндіктік тәсіл қолданылып, алдын ала берілген сенімділік деңгейінде қамтуды қамтамасыз ететін бұлдыр интервалдар құрылады. Әрбір уақыт нүктесі үшбұрышты бұлдыр сан (ҮБС) ретінде ұсынылып, кейін сегмент деңгейінде агрегатталады. Барлық сегменттер бойынша алынған нәтижелер сегмент ұзындығы бойынша өлшенген тәсіл арқылы біріктіріліп, жаһандық үшбұрышты бұлдыр сан қалыптастырылады. Оның центроиді белгісіздік жағдайындағы ЖІӨ-нің ұзақ мерзімді интегралды бағасын береді. Алынған нәтижелер ұсынылған тәсілдің макроэкономикалық уақыт қатарларындағы құрылымдық үзілістерді, бейсызық даму траекторияларын және белгісіздіктің таралуын тиімді сипаттайтынын көрсетеді. Сонымен қатар, авторлар ұсынған гибриді әдістемелік құрылым цифрландыру талаптарын қанағаттандыру мақсатында интеллектуалды ақпараттық жүйелердің жеке модулі ретінде интеграцияланып, басқа экономикалық және аналитикалық платформаларға масштабталуы мүмкін. Ұсынылған модель экономикалық талдау, болжау және цифрлық экономиканы басқару міндеттерінде қолдануға перспективалы құрал болып табылады.

Түйін сөздер: ЖІӨ модельдеуі, бөліктік логистикалық регрессия, бұлдыр сандар, ықтималды белгісіздік, Қазақстан экономикасы, интеллектуалды жүйелер

Қаржыландыру. Л.Н. Гумилев атындағы Еуразия ұлттық университеті қызметкерлерінің бұл жұмысы Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитетінің қаржылық қолдауымен жүзеге асырылады (Грант № AP26196337).

© Керимкулов С.Е.¹, Адалбек А.^{*1}, Байзаков Н.А.²,
Шодорова Н.Н.³, 2026.

¹Евразийский национальный университет имени Л.Н. Гумилева,
Астана, Казахстан;

²Информационно-вычислительный центр, Астана, Казахстан;

³Департамент программ, устойчивых инноваций и технологий,
Астана, Казахстан.

*E-mail: adalbek_a@enu.kz

КУСОЧНО-ЛОГИСТИЧЕСКОЕ И НЕЧЕТКОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ ВВП КАЗАХСТАНА (1990–2024)

Керимкулов Сейт — профессор факультета Информационных технологий, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: kerimkhulle_sye@enu.kz, <https://orcid.org/0000-0002-5886-6064>;

Адалбек Алибек — PhD-студент факультета информационных технологий, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: adalbek_a@enu.kz, <https://orcid.org/0000-0001-8543-2192>;

Байзаков Наурыз — заместитель директора Информационно-вычислительного центра, Астана, Казахстан,
E-mail: n.baizakov@statdata.kz, <https://orcid.org/0000-0003-1346-400X>;

Шодорова Найля — Департамент программ, устойчивых инноваций и технологий, Астана, Казахстан,
E-mail: nailya.shodorova@sitf.group, <https://orcid.org/0009-0006-8606-9181>.

Аннотация: *Актуальность.* В условиях ускоренной цифровизации экономики возрастает потребность в интеллектуальных моделях анализа, прогнозирования и управления сложными макроэкономическими процессами. Цифровизация народного хозяйства расширяет возможности принятия управленческих решений на основе данных и снижает зависимость от субъективных экспертных оценок. Для эффективного анализа динамики валового внутреннего продукта необходимы математические и информационные инструменты, учитывающие нелинейность, структурные изменения и факторы неопределенности. *Цель.* Разработать гибридную методологическую основу для анализа динамики ВВП Казахстана за 1990–2024 годы на основе кусочно-логистического моделирования и вероятностных нечетких интервалов. *Методы.* В исследовании предложен подход, сочетающий периодическое логистическое моделирование и вероятностные нечеткие интервалы. Временной ряд ВВП разделен на

экономически значимые подпериоды, для каждого из которых в зависимости от количества наблюдений строятся адаптивные трехпараметрические (3PL) или четырехпараметрические (4PL) логистические модели. На основе полученных моделей определяется центральная траектория, а неопределенность и остаточные значения описываются в логарифмическом пространстве. Для оценки неопределенности применяется вероятностный подход и формируются нечеткие интервалы, обеспечивающие охват на заданном уровне доверия. Каждая точка времени представляется в виде треугольного нечеткого числа (TFN), которое далее агрегируется на уровне сегмента. *Результаты и выводы.* Результаты, полученные для всех сегментов, объединяются с использованием подхода, взвешенного по длине сегмента, для формирования глобального треугольного нечеткого числа. Его центроид обеспечивает долгосрочную интегральную оценку ВВП в условиях неопределенности. Полученные результаты показывают, что предложенный подход эффективно характеризует структурные разрывы, нелинейные траектории развития и распространение неопределенности в макроэкономических временных рядах. Гибридная методологическая структура может быть интегрирована в качестве отдельного модуля интеллектуальных информационных систем и масштабирована на другие экономические и аналитические платформы. Предложенная модель является перспективным инструментом для экономического анализа, прогнозирования и управления цифровой экономикой.

Ключевые слова: моделирование ВВП, кусочно-логистическая регрессия, нечеткие числа, вероятностная неопределенность, экономика Казахстана, интеллектуальные системы, макроэкономические временные ряды

Kіріспе. Қазіргі таңда макроэкономикалық жүйелердің даму мен өсу динамикасын зерттеу барысында бейсызық модельдер мен гибридітік аналитикалық тәсілдер өте кеңінен қолданылуда. Әсіресе жалпы ішкі өнім (ЖІӨ) сияқты интегралды көрсеткіштер экономикалық өсудің күрделі, кезеңдік және құрылымдық өзгерістерге тәуелді сипатын нақты көрсетеді. Осыған байланысты экономикалық өсуді сипаттауда бөліктік логистикалық өсу модельдері, уақыт қатарларын болжаудың бұлдыр әдістері және нақты деректерге негізделген гибридітік тәсілдер маңызды орын алады (Kwasnicki, 2013; Jayathilaka et al., 2022; Chakamera & Pisa, 2021; Saidi et al., 2020; Oancea et al., 2021; Rogachev et al., 2023; Elbatal et al., 2025).

Әдеби шолу. Логистикалық өсу модельдері экономикалық жүйелердің шектеулі ресурстар жағдайындағы эволюциясын нақты сипаттауда тиімді құрал ретінде қарастырылады. Сонымен қатар, соңғы зерттеулерде ұлттық экономикалық көрсеткіштерді болжауда классикалық статистикалық әдістермен қатар бұлдыр логикаға негізделген модельдер де өте кеңінен қолданылып келеді, себебі олар белгісіздік пен толық емес ақпарат

жағдайында сенімді әрі нақты нәтижелер береді (Oancea et al., 2021; Rogachev et al., 2023). Сонымен бірге экономикалық өсім, логистика және макроэкономикалық көрсеткіштер арасындағы өзара байланыстарды зерттеу экономикалық жүйелердің күрделі табиғатын көрсетеді (Jayathilaka et al., 2022; Saidi et al., 2020).

Дегенмен, көптеген зерттеулерде уақыт қатарлары біртекті деп қарастырылып, құрылымдық үзілістер мен кезеңдік өзгерістер жеткілікті деңгейде ескерілмейді. Бұл өз кезегінде модельдердің дәлдігіне әсер етеді. Осыған байланысты кезеңдік (piecewise) модельдеу және әрбір кезең үшін жеке бейсызық модельдерді қолдану қажеттілігі туындайды. Сонымен қатар, экономикалық деректердегі анықталмағандық пен белгісіздікті сипаттау үшін мүмкіндіктік (possibilistic) және бұлдыр тәсілдерді енгізу өзекті ғылыми мәселе болып табылады (Rogachev et al., 2023; Elbatal et al., 2025).

Осы ғылыми зерттеуде Қазақстан Республикасының ЖІӨ динамикасын талдау үшін кезеңдік логистикалық модельдеу мен бұлдыр интервалдарды біріктіретін интеллектуалды модуль ұсынылады. Ұсынылған тәсіл экономикалық өсудің бейсызық табиғатын, кезеңдік құрылымын және анықталмағандық пен белгісіздік факторларын бір мезгілде ескеруге мүмкіндік береді.

Осылайша, дәстүрлі регрессиялық модельдер мұндай күрделі динамиканы жеткілікті деңгейде сипаттай алмайды. Осыған байланысты соңғы жылдары:

- логистикалық және сигмоидалық өсу модельдері;
 - бұлдыр (fuzzy) және мүмкіндіктік (possibilistic) тәсілдер;
 - кезеңдік (piecewise) модельдеу
- сияқты әдістер белсенді қолданылуда.

Бұл ғылыми жұмыстың мақсаты – Қазақстанның 1990–2024 жылдар аралығындағы ЖІӨ деректерін кезеңдерге бөліп, әр кезеңге логистикалық модельдер қолдану және нәтижелерді мүмкіндіктік бұлдыр интервалдар мен үшбұрышты бұлдыр сандар (ҮБС) арқылы біріктіріп талдау.

Материалдар мен әдістер. Бұл зерттеуде Қазақстан Республикасының 1990–2024 жылдар аралығындағы жалпы ішкі өнімінің (ЖІӨ) жылдық уақыт қатары пайдаланылды. Деректер Дүниежүзілік банк (World Bank) ашық статистикалық базасынан алынған және әрі қарай кезеңдік логистикалық модельдеу мен бұлдыр интервалдарды құру үшін өңделді.

Деректер және кезеңдеу. Бұл зерттеуде Қазақстан Республикасының ағымдағы АҚШ долларымен есептелген ЖІӨ деректері (1990–2024 жж.) пайдаланылды. Уақыт қатары экономикалық тұрғыдан мағыналы сегменттерге бөлінді. Кезеңдеу процедурасы екі еселену қағидасына негізделді, яғни әрбір сегмент ЖІӨ мәні бастапқы деңгейінен шамамен екі есеге артқан алғашқы уақыт нүктесімен аяқталды. Мұндай тәсіл экономикалық өсудің қарқындылығы мен құрылымдық өзгерістерін табиғи түрде айқындауға мүмкіндік береді.

Аталған критерийге сәйкес уақыт қатары келесі кезеңдерге бөлінді:

– 1990–2005: өтпелі кезең және экономикалық құлдыраудан кейінгі баяу қалпына келу;

– 2006–2011: мұнай бағасының өсуімен байланысты жедел экономикалық кеңею;

– 2012–2024: салыстырмалы тұрақты, бірақ сыртқы факторларға сезімтал өсу кезеңі.

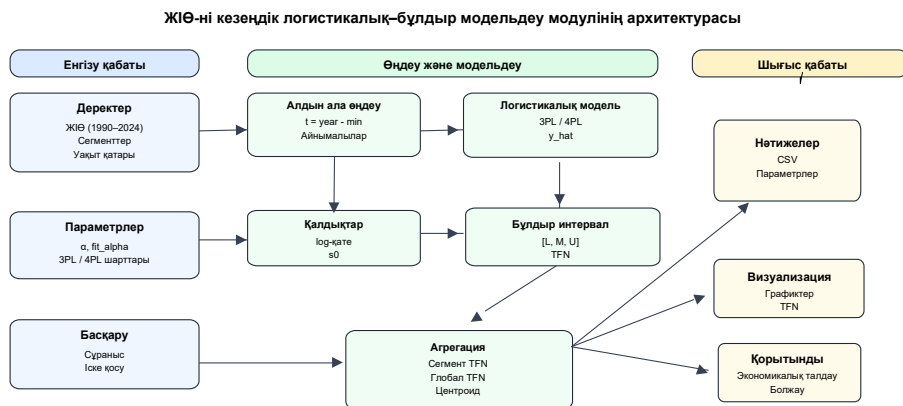
Екі еселену шегіне негізделген кезеңдеу әдісі әрбір сегменттің ішкі экономикалық біртектілігін қамтамасыз етеді және логистикалық модель параметрлерін сенімді бағалауға мүмкіндік береді. Сонымен қатар, бұл тәсіл уақыт қатарындағы құрылымдық үзілістерді формалды түрде анықтауға және әрбір кезеңді дербес экономикалық режим ретінде қарастыруға жағдай жасайды. Әрбір анықталған сегмент үшін жеке логистикалық модель сәйкестендіріліп, кейінгі талдау үшін қолданылды.

Ұсынылған әдіснама классикалық бұлдыр регрессия (fuzzy regression) және ықтималды регрессия (possibilistic regression) теорияларына негізделген және оларды кезеңдік логистикалық модельдеумен біріктіреді. Бұл тәсілдің теориялық негізі алғаш рет бұлдыр регрессия моделін ұсынған Танака және оның әріптестерінің еңбектеріне барып тіреледі, онда модель параметрлері бұлдыр сандар ретінде қарастырылады (Tanaka et al., 1982). Кейін бұл тәсіл possibilistic регрессия бағытында дамытылып, белгісіздік ықтималдық емес, мүмкіндіктік сипатта қарастырылды (Tanaka & Hayashi, 1989).

Осы жұмыста аталған теориялар кезеңдік (piecewise) регрессиямен біріктірілді. Мұндай тәсіл уақыт қатарындағы құрылымдық өзгерістерді автоматты түрде немесе алдын ала анықталған сегменттер арқылы есепке алуға мүмкіндік береді. Бұл бағыттағы маңызды зерттеулердің бірі ретінде бұлдыр кезеңдік регрессия (fuzzy piecewise regression) моделін ұсынған Ю және оның әріптестерінің жұмысы қарастырылады (Yu et al., 2001).

Ұсынылған модульде әрбір сегмент үшін логистикалық өсу моделі сәйкестендіріліп, кейін логарифмдік қалдықтар негізінде мүмкіндіктік интервалдар құрылады. Әрбір уақыт нүктесі үшбұрышты бұлдыр сан ретінде ұсынылып, сегменттік және глобалдық деңгейде агрегатталады. Бұл тәсіл уақыт қатарын тек нүктелік бағалау арқылы емес, белгісіздік интервалымен бірге сипаттауға мүмкіндік береді.

Осы әдіс интеллектуалды жүйе құрамындағы модуль ретінде жүзеге асырылады және макроэкономикалық уақыт қатарларын бейсызық және белгісіздік жағдайында талдауға бағытталған. Модульдің жұмыс логикасы үш негізгі кезеңнен тұрады: деректерді енгізу, өңдеу және нәтижелерді шығару (Сурет 1).



Сурет 1 – Қазақстанның ЖІӨ динамикасын бөліктік логистикалық және ықтималдық тұрғыдан бұлдыр модельдеуге арналған интеллектуалды жүйе модулінің архитектурасы (1990–2024)

Енгізілетін мәліметтер қабаты. Жүйенің бірінші қабаты бастапқы деректер мен модель параметрлерін қабылдауға арналған. Бұл қабатқа келесі ақпараттар енгізіледі:

- Макроэкономикалық уақыт қатары: 1990–2024 жылдар аралығындағы Қазақстанның ЖІӨ мәндері (ағымдағы АҚШ долларымен);
- Сегментация құрылымы: уақыт қатары алдын ала экономикалық интерпретацияға сәйкес бірнеше кезеңдерге бөлінеді;
- Модель параметрлері:
 - α – бұлдыр интервал деңгейі;
 - fit_alpha – қамту деңгейін анықтайтын параметр;
- Жүйелік басқару сигналдары: қолданушының сұранысы немесе аналитикалық сценарий.

Бұл қабат интеллектуалды жүйенің сыртқы интерфейсімен байланысып, есептеу процесінің бастапқы шарттарын анықтайды.

Өңдеу және модельдеу қабаты. Екінші қабат жүйенің негізгі есептеу ядросы болып табылады және бірнеше өзара байланысты модульдерден тұрады:

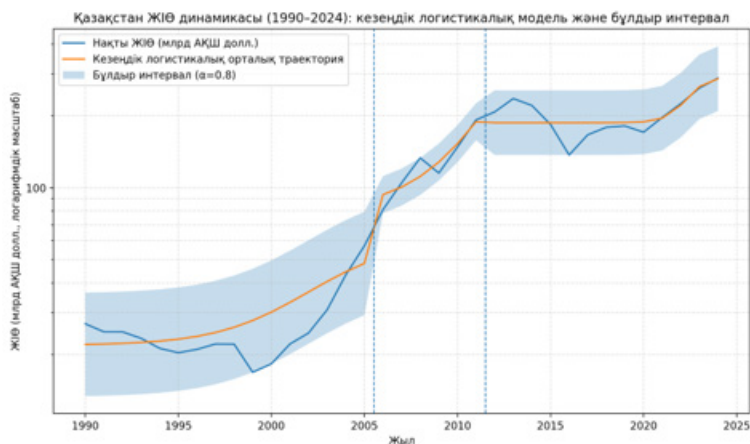
- (a) Деректерді алдын ала дайындау;
- (b) Кезеңдік логистикалық сәйкестендіру;
- (c) Қалдықтарды және белгісіздікті бағалау;
- (d) Бұлдыр интервал және үшбұрышты бұлдыр сандар (ҮБС);
- (e) Агрегация және интегралды бағалау;
- (f) Итерациялық тексеру механизмі.

Шығыс және қорытынды нәтижелер қабаты. Соңғы қабатта есептеу нәтижелері бірнеше форматта ұсынылады:

- (a) Құрылымдалған деректер;
- (b) Параметрлік бағалар;

(с) Визуализация.

Нәтижелер. Ұсынылған кезеңдік логистикалық-бұлдыр модель нәтижелері Қазақстан Республикасының 1990–2024 жылдар аралығындағы жалпы ішкі өнімінің (ЖІӨ) динамикасын кешенді түрде сипаттауға мүмкіндік берді. Нәтижелердің негізгі көрінісі Сурет 2-те ұсынылған, онда ЖІӨ-нің нақты мәндері, кезеңдік логистикалық модель арқылы алынған орталық траектория траектория және оған сәйкес бұлдыр интервал көрсетілген.



Сурет 2 – Қазақстан Республикасының ЖІӨ динамикасы (1990–2024): кезеңдік логистикалық модель және мүмкіндіктік бұлдыр интервал

2-суретте 1990–2024 жылдардағы Қазақстанның ЖІӨ динамикасы бөліктік логистикалық көрініспен және логарифмдік доменде салынған ықтималдық дәлізбен бірге көрсетілген. Негізгі әдіснамалық қадам – экономикалық тұрғыдан түсіндірілетін ережені қолдана отырып, өсу режимдерін (сегменттерін) анықтау: режим ЖІӨ деңгейі s жылынан басталып, бірінші жылы Y_s жылы аяқталады, осылайша $Y_t \geq 2Y_s$. Бұл екі еселенген шекті сегменттеу үш кезеңді береді:

- 1-кезең (1990–2005): алғашқы екі еселену тек 2005 жылы ғана болады, бұл екі еселенудің ұзаққа созылуын (15 жыл) білдіреді.
- 2-кезең (2006–2011): екінші екі еселенуге 2011 жылға қарай (5 жыл) қол жеткізіледі, бұл жылдам кеңею циклін көрсетеді.
- 3-кезең (2012–2024): екі еселену шартына 2024 жылға дейін қол жеткізілмейді, бұл жоғары деңгейлі, бірақ баяу және тұрақсыз режимді көрсетеді.

Бұл сегменттеу тек статистикалық ғана емес, сонымен қатар экономикалық тұрғыдан да маңызды: ол әрбір режимді екі еселік масштабты өсуге жету үшін қажетті уақытпен байланыстырады, осылайша кеңею жылдамдығын байқалатын құрылымдық ерекшелік ретінде қорытындылайды. Маңыздысы, бірінші және үшінші кезеңдер макроэкономикалық турбуленттілік

кезеңдеріне ие, ал екінші кезең салыстырмалы түрде тұрақты жоғары өсу аралығын көрсетеді.

ЖІӨ траекторияларының бөліктік логистикалық көрінісі. Әрбір анықталған кезеңде ЖІӨ сызықтық емес өсу үлгілерін (S-тәрізді динамика) анықтау үшін логистикалық типтегі траекториямен жуықталады. Әрбір сегменттегі орталық жол төрт параметрлі логистикалық (4PL) арқылы беріледі,

$$\hat{Y}(t) = A + \frac{K - A}{1 + \exp[-r(t - t_0)]},$$

немесе қысқа сегменттер үшін шамадан тыс параметрлеуді болдырмау үшін сегменттің минимумында A бекітілген үш параметрлі логистика (3PL):

$$\hat{Y}(t) = A_{fixed} + \frac{K - A_{fixed}}{1 + \exp[-r(t - t_0)]},$$

мұнда A – төменгі шекара (floor), K – жоғарғы асимптота, r – өсу қарқыны және t_0 – иілу уақыты (inflection time). Бөлшектік (piecewise) тәсіл (Сурет 2) Қазақстанның ЖІӨ эволюциясын бірыңғай тегіс процесс ретінде сипаттауға болмайтынын түсіндіреді: керісінше, ол режимге тән динамиканы әртүрлі «орталықтары» мен қарқындарымен көрсетеді (Кесте 1). Мәндер: ЖІӨ – млрд АҚШ доллары; r – жылдық қарқын; t_0 – календарлық жыл (инфлексия уақыты); көбейткіш m ; $\alpha = 0.5$ деңгейіндегі дәліздің мультипликаторы:

Кесте 1 — Қазақстан ЖІӨ динамикасының кезеңдік логистикалық модель параметрлері және бұлдыр интервал сипаттамалары (1990–2024)

Regime (years)	n (years)	Doubling lag	Logistic spec	A (bn)	K (bn)	r	t_0 (year)	Multiplier m ($\alpha = 0.5$)
R1: 1990–2005	16	15	4PL	21.90	61.20	0.407	2003.26	1.648
R2: 2006–2011	6	5	3PL (A fixed)	81.00 (fixed)	963.13	0.453	2015.34	1.196
R3: 2012–2024	13	≥ 12 (not reached)	4PL	187.53	292.75	1.756	2022.41	1.366

Ескерту: R2 (2006–2011) сегменті қысқа шығын 3PL логистикасы K және t_0 -ды үлгі сыртында “экстраполяциялау” мүмкін; бұл режим үшін ең интерпретацияланатын параметр – r және “band” тарлығы (көбейткіш).

1-режим (1990–2005) өтпелі кезеңге және ерте қалпына келуге сәйкес келеді. ЖІӨ 1990 жылдардың басында төмендеді және 1990 жылдардың аяғында баяу болып қала береді, 1998 жыл шамасында Ресей қаржы дағдарысының аймақтық әсерлеріне сәйкес айқын бұзылулар байқалды. Логистикалық сәйкестік 2005 жылға қарай алғашқы екі еселенумен

аякталатын тегіс қалпына келу жолын қамтамасыз етеді. Ұзақ екі еселену кешігуі экономиканың өтпелі қысқаруды өтеу және тұрақты номиналды кеңеюді орнату үшін он жылдан астам уақыт қажет болғанын көрсетеді.

2-режим (2006–2011) ең жылдам екі еселену кідірісін көрсетеді (5 жыл). Орталық жол 2006 жылғы базалық деңгейден 2011 жылғы деңгейге дейін күрт көтеріледі. Бұл режим 2009 жылы айтарлықтай ауытқуды (әлемдік қаржы дағдарысын) қамтиды, онда байқалған ЖІӨ өсу жолынан төмен түсіп, содан кейін тез қалпына келеді. Режимнің 2011 жылға қарай екі еселенуге қол жеткізуі осы кезеңде жоғары тұрақтылықты және шоктан кейінгі қалпына келтірудің күшті мүмкіндігін көрсетеді.

3-режим (2012–2024) күшті құбылмалылықпен және толық емес екі еселенумен сипатталатын жоғары деңгейлі режимді білдіреді. Байқалған ЖІӨ 2013 жылы шыңына жетеді, 2014–2016 жылдары төмендейді (мұнай бағасының құлдырауы және сыртқы күйзелістер, соның ішінде валюта бағамын түзету), содан кейін біртіндеп қалпына келеді. 2020 жыл шамасында тағы бір үзіліс байқалады (COVID-19 күйзелісі), содан кейін ЖІӨ 2021–2024 жылдары күрт өседі. Жақында болған күрт өсуге қарамастан, 2012 жылғы базалық деңгейден екі еселену шарты 2024 жылға қарай орындалмайды, бұл «жылдам масштабтаудан» «баяулау эпизодтарымен ауытқып отыратын өсуге» ауысуды көрсетеді, бұл кеңею жаһандық жағдайлармен және тауар цикліне тәуелділікпен шектелетін жетілген номиналды деңгейге сәйкес келеді.

Жалпы алғанда, 2-сурет бөліктік (piecewise) логистикалық орталықтың әрбір режимде түсіндірілетін «құрылымдық үрдіс» (structural trend) ретінде әрекет ететінін растайды, ал 1998, 2009, 2014–2016 және 2020 жылдардағы ауытқулар ЖІӨ-ні режим жолынан уақытша ығыстыратын ірі макроэкономикалық күйзелістерге сәйкес келеді.

Логарифмдік домендері (log domain) мүмкіндікті (possibilistic) айқын емес дәліз. Модельдеу құбырының негізгі үлесі логистикалық орталықтың айналасында мүмкіндікті (possibilistic) айқын емес жолақты құру болып табылады (2-сурет). Ықтималдық (probabilistic) болжамдарды (мысалы, қалыпты қалдықтар немесе тұрақты дисперсия) таңудың орнына, белгісіздік логарифмдік түрлендірілген кеңістіктегі абсолютті ауытқулардан алынған ықтималдық үлестірімі ретінде ұсынылады:

$$z_t = \ln(Y_t), \hat{z}_t = \ln(\hat{Y}_t), e_t = |z_t - \hat{z}_t|.$$

Әрбір режим үшін s_0 масштабтық параметрі таңдалған калибрлеу деңгейінде $\delta = \max_t e_t$: максималды абсолютті логарифмдік қалдық (absolute log-residual) $\alpha_{fit} \in [0,1]$ арқылы калибрленеді:

$$s_0 = \frac{\delta}{1 - \alpha_{fit}}.$$

Содан кейін, берілген α -кесу (α -cut) үшін (2-суретте көрсетілген), логарифмдік кеңістіктегі бұлдыр дәліз келесідей анықталады:

$$|z - \hat{z}| \leq (1 - \alpha)s_0.$$

Бастапқы шкала бойынша қайта оралу мультипликативті белгісіздік дәлізін береді:

$$Y \in [\hat{Y} \exp(-(1 - \alpha)s_0), \hat{Y} \exp((1 - \alpha)s_0)].$$

Бұл форма макроэкономикалық деректер үшін өте қолайлы, себебі күйзелістер көбінесе мультипликативті түрде әрекет етеді (валюта бағамының өзгеруі, сауда шарттарының әсері), ал логарифмдік масштабтағы дәліз абсолютті ауытқуларға емес, салыстырмалы (пайыздық) белгісіздікке сәйкес келеді. Берілген есепті іске асыруда ($\alpha = \alpha_{fit} = 0.5$) дәліз құрылым арқылы бақыланатын мәндерді толық шектеуге қол жеткізеді, яғни барлық бақылаулар таңдалған деңгейде бұлдыр диапазонда болады. Экономикалық тұрғыдан алғанда, бұл әрбір режимде сенімді «тұрақты дәлізді» қамтамасыз етеді: логистикалық орталыққа жақын жылдарды трендке жақын (құрылымдық тұрғыдан тұрақты) деп түсіндіруге болады, ал дәліз шекараларына жақындаған жылдар күйзелісті немесе ерекше күйзелістерді көрсетеді (мысалы, 2009, 2015–2016, 2020).

Дәліз ендерін режимдер бойынша салыстыру қосымша түсінік береді: ерте өтпелі режим ең кең салыстырмалы бұлыңғырлықты көрсетеді (терең құрылымдық өзгерістер мен тұрақсыздықты көрсетеді), 2006–2011 жылдардағы кеңею режимі ең тар жолақты тудырады (неғұрлым үйлесімді өсу), ал 2012 жылдан кейінгі режим қайтадан кеңейген белгісіздік көрсетеді (соққыға сезімтал жетілу фазасы). Демек, бұлыңғыр жолақ жай ғана визуализация емес, режимге тәуелді тұрақтылық индикаторы болып табылады.

Бұлыңғыр дәліздерден үшбұрышты бұлыңғыр сандарға дейін: режим деңгейіндегі және жаһандық (global) көрсеткіштер. Салыстырмалы талдау үшін дәлізді іске қосу үшін жыл сайын t үшбұрышты бұлыңғыр санға бейнеленеді:

$$\hat{Y}_t = (L_t, M_t, U_t),$$

мұнда

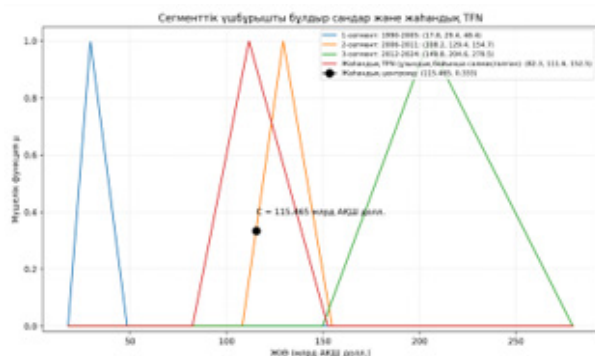
$$M_t = \hat{Y}_t, L_t = \hat{Y}_t \exp(-(1 - \alpha)s_0), U_t = \hat{Y}_t \exp((1 - \alpha)s_0).$$

Осылайша, логистикалық орталық модальды (ең шындыққа жақын) (most plausible) M_t мәніне айналады, ал α -кесілген дәліз шекаралары сол/оң жақ

тіректер (L_t, U_t) болады. 2-суретте әрбір режимдегі жылдық үшбұрыш бұлдыр сандарды орташалау арқылы алынған әрбір режим үшін бір агрегаттық үшбұрыш бұлдыр сандар көрсетілген (жылына тең салмақтар):

$$\tilde{Y}_{reg} = \left(\frac{1}{n} \sum_{t \in reg} L_t, \frac{1}{n} \sum_{t \in reg} M_t, \frac{1}{n} \sum_{t \in reg} U_t \right).$$

Режимнің үшбұрыш бұлдыр сандары (3-сурет) айқын реттілікті көрсетеді: 1-режим ең төменгі модальды деңгейге және ең кең салыстырмалы белгісіздікке ие, 2-режим таралуды тар таралумен жоғары деңгейлерге ауыстырады, ал 3-режим ең жоғары модальды деңгейді, бірақ ең кең абсолютті таралуды береді, бұл номиналды шаманың жоғарылығына және соққыға негізделген дисперсияға сәйкес келеді.



Сурет 3 – Қазақстан ЖІӨ динамикасындағы белгісіздік: мүмкіндіктік бұлдыр интервалдың графикалық көрінісі

Әрі қарай, 1990-2024 жылдар аралығындағы кезеңді қорытындылау үшін жаһандық (global) үшбұрыш бұлдыр саны есептеледі. Екі толықтырушы агрегация қағидаты ақпараттық болып табылады:

1. n_j (әр режим жылдары) режимінің ұзақтығын пайдаланып режимнің үшбұрыш бұлдыр сандарынан құрылған, ұзындығы бойынша үлестірілген (length-weighted) құрылымдық жаһандық (global) үшбұрыш бұлдыр сандар:

$$\tilde{Y}_G^{(len)} = \left(\frac{\sum_j n_j L_j}{\sum_j n_j}, \frac{\sum_j n_j M_j}{\sum_j n_j}, \frac{\sum_j n_j U_j}{\sum_j n_j} \right).$$

2. Үшбұрыш бұлдыр сандар үшбұрышының ауданына пропорционалды салмақтары бар жылдық үшбұрыш бұлдыр сандарынан тікелей құрылған белгісіздікке сезімтал (аудан бойынша салмақталған) (area-weighted) жаһандық (global) ҰБС:

$$w_t = \frac{U_t - L_t}{2}, \tilde{Y}_G^{(area)} = \left(\frac{\sum_t w_t L_t}{\sum_t w_t}, \frac{\sum_t w_t M_t}{\sum_t w_t}, \frac{\sum_t w_t U_t}{\sum_t w_t} \right).$$

3-суретте режимдік ҰБС және ұзындық бойынша өлшенген жаһандық ҰБС көрсетілген.

Үшбұрыш бұлдыр сандар центроиды арқылы тепе-теңдік деңгейі және тұрақты дәліз. Үшбұрыш бұлдыр сандарынан бірыңғай айқын индикатор алу үшін үшбұрышты мүшелік функциясының геометриялық центроидын қолданамыз. $(L, 0)$, $(M, 1)$, $(U, 0)$ төбелері бар кез келген үшбұрыш бұлдыр сандары $\tilde{Y} = (L, M, U)$ үшін оның центроиды:

$$Y^* = \frac{L + M + U}{3}, \mu^* = \frac{1}{3}.$$

3-суретте центроид кара нүкте ретінде көрсетілген. Біз Y^* -ті модельденген режим(дер)ге сәйкес келетін анық емес тепе-теңдік деңгейі (немесе «орталық тұрақты деңгей») ретінде түсіндіреміз, ал $[L, U]$ таңдалған α -деңгейде тұрақты дәлізді қамтамасыз етеді. Бұл жылдарды түсіндірудің жұмыс ережесін береді:

– **Тепе-теңдікке жақын:** Y_t мәні Y^* мәніне жақын (немесе дәлірек айтқанда, M мәніне жақын) болуы режимнің құрылымдық үрдісімен сәйкестікті көрсетеді.

– **Стресс/тенгерімсіздік:** L немесе U маңындағы Y_t экономиканың тұрақты дәліз шекарасында жұмыс істеп тұрғанын көрсетеді, бұл көбінесе ерекше күйзелістерге немесе қызып кету фазаларына байланысты.

– **Дәлізден тыс аномалия:** егер таңдалған α кезінде $Y_t \notin [L, U]$ болса, бұл қолайлы ықтималдық белгісіздік шеңберінен тыс ерекше ауытқуды білдіреді (қазіргі калибрлеуде шеңбер шектеуді қамтамасыз ету үшін таңдалған).

Қазақстан ЖІӨ-нің сегменттік және жаһандық деңгейлер бойынша үшбұрышты бұлдыр сандары және олардың центроидтық бағалары Кесте 2-де ұсынылған.

Кесте 2 — Қазақстан ЖІӨ үшін үшбұрышты бұлдыр сандар және центроид көрсеткіштері, $\alpha = 0.5$

Regime / Global	TFN (L, M, U) in bn current US\$	Centroid Y^* (bn)	Weighting/ aggregation
R1 (1990–2005)	(17.85, 29.40, 48.44)	31.90	Equal-years within regime
R2 (2006–2011)	(108.16, 129.36, 154.71)	130.74	Equal-years within regime
R3 (2012–2024)	(149.75, 204.57, 279.47)	211.26	Equal-years within regime
GLOBAL (length-weighted)	(82.32, 111.60, 152.47)	115.46	Regime length (n_years)
GLOBAL (area-weighted)	(122.34, 165.98, 225.86)	171.39	Yearly TFN area $(U - L)/2$

Ескерту: GLOBAL (length-weighted) – уақыт бойынша “құрылымдық орташа” деңгей; GLOBAL (area-weighted) – белгісіздігі (дәліз ені) үлкен жылдарды көбірек ескеретін “тәуекел-сезімтал” деңгей.

Талқылау. Ұсынылған модельдеу нәтижелері Қазақстан ЖІӨ динамикасының құрылымдық және бейсызық сипатын нақты көрсетеді. Екі еселену шегіне негізделген сегменттеу экономикалық тұрғыдан интерпретацияланатын өсу режимдерін анықтауға мүмкіндік берді: екі еселену уақытының 15 жылдан 5 жылға дейін қысқаруы 2000 жылдардағы жедел өсімді сипаттаса, 2012 жылдан кейін екі еселенудің болмауы баяу және тұрақсыз өсу фазасына өтуді білдіреді. Кезеңдік логистикалық модель әр режим үшін бейсызық өсу траекториясын сипаттап, 1998, 2009, 2014–2016 және 2020 жылдардағы ауытқуларды уақытша сыртқы соққылар ретінде түсіндіреді. Сонымен қатар, логарифмдік кеңістіктегі қалдықтарға негізделген мүмкіндіктік бұлдыр интервал белгісіздікті ықтималдық болжамдарға тәуелсіз түрде сипаттап, оның ені режим тұрақтылығының маңызды көрсеткіші ретінде қарастырылады.

Әрі қарай бұлдыр интервалдарды үшбұрышты бұлдыр сандарға және олардың центроидтарына түрлендіру ЖІӨ-нің интегралды бағаларын алуға мүмкіндік береді. Ұзындық бойынша және аумақ бойынша өлшенген жаһандық үшбұрыш бұлдыр сандардың айырмашылығы «орташа деңгейдің» агрегация тәсіліне тәуелді екенін көрсетеді, яғни құрылымдық және тәуекел-сезімтал бағалаулардың айырмашылығы анықталады. Алынған нәтижелер заманауи зерттеулермен сәйкес келеді: логистикалық модельдер экономикалық динамиканы сипаттауда тиімді (Ray, 2024), ал бұлдыр регрессия белгісіздік жағдайында сенімді бағалаулар береді (Pandelara et al., 2022). Сонымен қатар, экономикалық жүйелердегі бейсызық байланыстар (Li et al., 2021) және fuzzy тәсілдердің шешім қабылдаудағы маңызы (Profiroiu et al., 2025) ұсынылған әдістің қолданбалы құндылығын растайды.

Қорытынды. Бұл жұмыста Қазақстан Республикасының 1990–2024 жылдар аралығындағы ЖІӨ динамикасын талдау үшін кезеңдік логистикалық модельдеу мен мүмкіндіктік бұлдыр интервалдарды біріктіретін жаңа гибриді тәсіл ұсынылды. Ұсынылған әдіс уақыт қатарын құрылымдық сегменттерге бөліп, әрбір кезең үшін бейімделгіш логистикалық модельдерді қолдану арқылы экономикалық өсудің бейсызық сипатын нақты сипаттауға мүмкіндік береді.

Зерттеу нәтижелері көрсеткендей, екі еселену шегіне негізделген сегменттеу экономикалық даму кезеңдерін анықтауда тиімді құрал болып табылады. Сонымен қатар, логарифмдік қалдықтарға негізделген мүмкіндіктік бұлдыр интервал белгісіздікті сенімді сипаттап, классикалық ықтималдық тәсілдерге балама ретінде ұсынылады. Үшбұрышты бұлдыр сандар мен олардың центроидтарын қолдану ЖІӨ-нің интегралды бағаларын алуға және экономикалық жүйенің ұзақ мерзімді тепе-теңдік деңгейін анықтауға мүмкіндік береді.

Ұсынылған тәсілдің негізгі артықшылығы – құрылымдық үзілістерді, бейсызық өсуді және белгісіздікті бір мезгілде есепке алуында. Бұл оны макроэкономикалық талдау, болжау және шешім қабылдау міндеттерінде

қосымша құрал ретінде қолдануға үлкен мүмкіндік береді. Болашақ зерттеулерде ұсынылған модельді салалық деңгейде кеңейту, сондай-ақ жасанды интеллект әдістерімен біріктіру оның қолданбалы мүмкіндіктерін одан әрі арттыра алады.

References

- Kwasnicki W. (2013) Logistic growth of the global economy and competitiveness of nations. *Technological Forecasting and Social Change*, 80(1). — P. 50–76 pp. <https://doi.org/10.1016/j.techfore.2012.07.007>. (in Eng.)
- Jayathilaka R., Jayawardhana C., Embogama N., Jayasooriya S., Karunarathna N., Gamage T., & Kuruppu N. (2022) Gross domestic product and logistics performance index drive the world trade: A study based on all continents. *PLOS ONE*, 17(3), e0264474 p. <https://doi.org/10.1371/journal.pone.0264474>. (in Eng.)
- Chakamera C., & Pisa N.M. (2021) Associations Between Logistics and Economic Growth in Africa. *South African Journal of Economics*, 89(3). — P. 417–438 pp. <https://doi.org/10.1111/saje.12272>. (in Eng.)
- Saidi S., Mani V., Mefteh H., Shahbaz M., & Akhtar P. (2020) Dynamic linkages between transport, logistics, foreign direct investment, and economic growth: Empirical evidence from developing countries. *Transportation Research Part A: Policy and Practice*, 141. — P. 277–293 pp. <https://doi.org/10.1016/j.tra.2020.09.020>. (in Eng.)
- Oancea B., Pospíšil R., Jula M.N., & Imbrișcă C.-I. (2021) Experiments with Fuzzy Methods for Forecasting Time Series as Alternatives to Classical Methods. *Mathematics*, 9(19), 2517 p. <https://doi.org/10.3390/math9192517>. (in Eng.)
- Rogachev A.F., Simonov A.B., Ketko N.V., & Skiter N.N. (2023) Fuzzy Algorithmic Modeling of Economics and Innovation Process Dynamics Based on Preliminary Component Allocation by Singular Spectrum Analysis Method. *Algorithms*, 16(1). — 39 p. <https://doi.org/10.3390/a16010039>. (in Eng.)
- Elbatal I., Sarwar M., Jamal F., Daniyal M., Hussain Z., & Ben Ghorbal A. (2025) Modelling on gross domestic product annual growth rate data by using time series, machine learning, and probability models. *Journal of Radiation Research and Applied Sciences*, 18(2), 101481 p. <https://doi.org/10.1016/j.jrras.2025.101481>. (in Eng.)
- Tanaka H., Uejima S., & Asai K. (1982) Linear Regression Analysis with Fuzzy Model. *IEEE Transactions on Systems, Man, and Cybernetics*, 12. — P. 903–907 pp. <https://doi.org/10.1109/TSMC.1982.4308925>. (in Eng.)
- Tanaka H., & Hayashi I. (1989) Possibilistic Linear Regression Analysis for Fuzzy Data. *European Journal of Operational Research*, 40, 389–396 pp. [https://doi.org/10.1016/0377-2217\(89\)90431-1](https://doi.org/10.1016/0377-2217(89)90431-1). (in Eng.)
- Yu J.R., Tzeng G.-H., & Li H.-L. (2001) General fuzzy piecewise regression analysis with automatic change-point detection. *Fuzzy Sets and Systems*, 119(2). — P. 247–257 pp. [https://doi.org/10.1016/S0165-0114\(98\)00384-4](https://doi.org/10.1016/S0165-0114(98)00384-4). (in Eng.)
- Li X., Chen F., & Wang W. (2021) Impact of Logistics Development on Economic Growth: An Empirical Research from Guangdong Province in China. *Complexity*, 2021. — 9950935 p. <https://doi.org/10.1155/2021/9950935>. (in Eng.)
- Pandelara D., Kristjanpoller W., Michell K., & Minutolo M.C. (2022) A fuzzy regression causality approach to analyze relationship between electrical consumption and GDP. *Energy*, 239(Part E). — 122459 p. <https://doi.org/10.1016/j.energy.2021.122459>. (in Eng.)
- Profiroi C.M., Profiroi A.G., Constantin D.-L., Nica I., Delcea C., & Brișcariu M.R. (2025) Assessing regional economic performance: Fuzzy multi-criteria decision making and panel regression approach. *Socio-Economic Planning Sciences*, 100. — 102245 p. <https://doi.org/10.1016/j.seps.2025.102245>. (in Eng.)
- Ray A.K. (2024) Logistic Forecasting of GDP Competitiveness. *Global Economy Journal*, 24(01n04). — 2450008 p. <https://doi.org/10.1142/S2194565924500088>. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 228–247

<https://doi.org/10.32014/2026.2518-1726.436>

IRSTI 89.25

UDC 629.78

© Kulakayeva A.^{1*}, Ashurov A.², Aitmagambetov A.¹,
Ongenybayeva Zh.¹, 2026.

¹ JSC «International Information Technology University», Almaty, Kazakhstan;

² NJSC «L.N. Gumilyov Eurasian National University», Astana, Kazakhstan.

E-mail: a.kulakayeva@iitu.edu.kz

DEVELOPMENT OF MATHEMATICAL MODELS AND CRITERIA FOR THE ADMISSIBILITY OF ORBITAL MANEUVERS OF SPACECRAFT

Kulakayeva Aigul — PhD, Associate Professor of International Information Technology University, Almaty, Kazakhstan,

E-mail: a.kulakayeva@iitu.edu.kz, <https://orcid.org/0000-0002-0143-085X>;

Ashurov Abdikul — Candidate of Physical and Mathematical Sciences, Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan,

E-mail: ashurov_ae@enu.kz, <http://orcid.org/0000-0001-6044-5579>;

Aitmagambetov Altay — Candidate of Technical Sciences, professor of International Information Technology University, Almaty, Kazakhstan,

E-mail: a.aitmagambetov@iitu.edu.kz, <http://orcid.org/0000-0002-7808-5273>;

Ongenybayeva Zhadyra — senior-lecturer of International Information Technology University, Almaty, Kazakhstan,

E-mail: zh.ongenybayeva@iitu.edu.kz, <http://orcid.org/0000-0001-8497-2955>.

Abstract. The relevance of this work is driven by the increasing number of spacecraft and the increasing complexity of autonomous rendezvous, satellite servicing, and space debris avoidance. Existing classical approaches are often ineffective due to stringent engineering constraints requiring rapid solutions in real time. The goal of the study is to develop mathematical models and algorithms for synthesizing feasible maneuvers that ensure the achievement of specified end conditions, fuel economy, and strict adherence to dynamic thrust constraints, energy limits, and geometric attitude constraints. The study is based on classical celestial mechanics methods within the two-body problem. The spacecraft's state is described both in the inertial coordinate system (ECI) and in the local orbital reference frame (RTN). The methodology is based on the sequential application of three models: impulse, gradient, and low-thrust. As a result of the study, a set of mathematical models was developed that establish an analytical relationship

between the impact components in the RTN system and changes in the six Keplerian orbital elements. Optimal points of force application were determined to minimize side perturbations of orbital elements. A significant achievement was the formalization of maneuver admissibility conditions by describing a set of admissible controls. These conditions take into account constraints on the thrust magnitude and rate of change, mass consumption limits based on the Tsiolkovsky formula, and thrust vector orientation requirements. The study resulted in a comprehensive methodology that combines the theoretical foundations of ballistics with practical requirements for spacecraft design. The scientific novelty lies in the development of criteria that enable the transformation of theoretical calculations into a feasible flight plan, taking into account technical constraints. The practical significance of the study lies in the creation of a comprehensive methodology that combines the classical theory of celestial mechanics with real-world engineering requirements for spacecraft.

Keywords: orbital maneuvers, impulse maneuver, gradient maneuver, low-thrust maneuver, Keplerian elements, permissible maneuver, deorbiting, orbit

Financing. *This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan Grant No. AP26103099, titled “Development of deorbiting methods for Kazakhstan satellites using virtual reality technology”*

For citations: Kulakayeva A., Ashurov A., Aitmagambetov A., Ongenbayeva Zh. Development of mathematical models and criteria for the admissibility of orbital maneuvers of spacecraft. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 228–247. DOI <https://doi.org/10.32014/2026.2518-1726.436>

© Кулакаева А.Е.^{1*}, Ашуrow А.Е.², Айтмағамбетов А.З.¹,
Онгенбаева Ж.Ж.¹, 2026.

¹ «Халықаралық ақпараттық технологиялар университеті» АҚ,
Алматы, Қазақстан;

² «Л.Н. Гумилев атындағы Еуразия ұлттық университеті» КеАҚ,
Астана, Қазақстан.

E-mail: a.kulakayeva@iitu.edu.kz

ҒАРЫШ АППАРАТТАРЫНЫҢ ОРБИТАЛЫҚ МАНЕВРЛЕРІНІҢ МАТЕМАТИКАЛЫҚ МОДЕЛЬДЕРІ МЕН РҰҚСАТ КРИТЕРИЙЛЕРІН ӨЗІРЛЕУ

Кулакаева Айгуль — PhD, Халықаралық ақпараттық технологиялар университетінің қауымдастырылған профессоры, Алматы, Қазақстан,

E-mail: a.kulakayeva@iitu.edu.kz, <https://orcid.org/0000-0002-0143-085X>;

Ашуrow Абдикул — физика-математика ғылымдарының кандидаты, Л.Н. Гумилев атындағы Еуразия Ұлттық Университеті, Астана, Қазақстан,

E-mail: ashurov_ae@enu.kz, <http://orcid.org/0000-0001-6044-5579>;

Айтмағамбетов Алтай — техника ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университетінің профессоры, Алматы, Қазақстан,

E-mail: a.aitmagambetov@iitu.edu.kz, <http://orcid.org/0000-0002-7808-5273>;

Онгенбаева Жадыра — Халықаралық ақпараттық технологиялар университетінің аға оқытушысы, Алматы, Қазақстан,

E-mail: zh.ongenbayeva@iitu.edu.kz, <http://orcid.org/0000-0001-8497-2955>.

Аннотация. Жұмыстың өзектілігі ғарыш аппараттары санының артуына және автономды жақындасу, спутниктерге қызмет көрсету және ғарыш қоқыстарынан жалтару міндеттерінің күрделілігіне байланысты. Қолданыстағы классикалық тәсілдер көбінесе нақты уақыттағы жедел шешімді қажет ететін қатаң инженерлік шектеулерге байланысты тиімсіз болып шығады. Зерттеудің мақсаты-берілген соңғы шарттарға қол жеткізуді, отын үнемдеуді және динамикалық тарту шектеулерін, энергия шектеулерін және бағдарлау геометриялық шектеулерін қатаң сақтауды қамтамасыз ететін рұқсат етілген маневрлерді синтездеудің математикалық модельдері мен алгоритмдерін әзірлеу. Ғарыш аппаратының күйі ЕСІ инерциялық координаттар жүйесінде де, жергілікті орбиталық RTN жүйесінде де сипатталады. Әдістеме үш модельді дәйекті қолдануға негізделген: импульстік, градиенттік және төмен тарту. Жұмыс нәтижесінде RTN жүйесіндегі әсер ету компоненттері мен орбитаның алты Кеплер элементінің өзгерістері арасында аналитикалық байланыс орнататын математикалық модельдер кешені жасалды. Орбита элементтерінің жанама бұзылыстарын азайту үшін күш қолданудың оңтайлы нүктелері анықталды. Маңызды жетістік көптеген рұқсат етілген басқаруларды сипаттау арқылы маневрлердің рұқсат етілу шарттарын ресімдеу болды. Мұндай шарттар модуль бойынша шектеулерді және тартудың өзгеру жылдамдығын, Циолковский формуласына негізделген масса ағынының лимиттерін, сондай-ақ тарту векторының бағытына қойылатын талаптарды ескереді. Зерттеудің нәтижесі баллистиканың теориялық негіздерін ғарыш аппараттарының дизайнына практикалық талаптармен біріктіретін кешенді әдістеме болды. Ғылыми жаңалық-техникалық шектеулерді ескере отырып, теориялық есептеуді іске асырылатын ұшу жоспарына айналдыруға мүмкіндік беретін критерийлерді әзірлеу. Зерттеудің практикалық маңыздылығы аспан механикасының классикалық теориясын ғарыш аппараттарына қойылатын нақты инженерлік талаптармен біріктіретін кешенді әдістемені құру болып табылады.

Түйін сөздер: орбиталық маневрлер, импульстік маневр, градиент маневрі, төмен тарту маневрі, Кеплер элементтері, жарамды маневр, деорбитация, орбита

© Кулакаева А.Е.^{1*}, Ашуров А.Е.², Айтмагамбетов А.З.¹,
Онгенбаева Ж.Ж.¹, 2026.

¹ АО «Международный университет информационных технологий»,
Алматы, Казахстан;

² НАО «Евразийский национальный университет им. Л.Н. Гумилева»,
Астана, Казахстан.

E-mail: a.kulakayeva@iitu.edu.kz

РАЗРАБОТКА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ И КРИТЕРИЕВ ДОПУСТИМОСТИ ОРБИТАЛЬНЫХ МАНЕВРОВ КОСМИЧЕСКИХ АППАРАТОВ

Кулакаева Айгуль — PhD, Ассоциированный профессор Международного университета информационных технологий, Алматы, Казахстан,

E-mail: a.kulakayeva@iitu.edu.kz, <https://orcid.org/0000-0002-0143-085X>;

Ашуров Абдикул — кандидат физико-математических наук, Евразийский Национальный Университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: ashurov_ae@enu.kz, <http://orcid.org/0000-0001-6044-5579>;

Айтмагамбетов Алтай — кандидат технических наук, профессор Международного университета информационных технологий, Алматы, Казахстан,

E-mail: a.aitmagambetov@iitu.edu.kz, <http://orcid.org/0000-0002-7808-5273>;

Онгенбаева Жадыра — старший преподаватель Международного университета информационных технологий, Алматы, Казахстан,

E-mail: zh.ongenbayeva@iitu.edu.kz, <http://orcid.org/0000-0001-8497-2955>.

Аннотация. Актуальность работы обусловлена увеличением количества космических аппаратов и возрастающей сложностью задач автономного сближения, обслуживания спутников и уклонения от космического мусора. Существующие классические подходы часто оказываются недостаточно эффективными из-за жестких инженерных ограничений, требующих оперативного решения в реальном времени. Целью исследования является разработка математических моделей и алгоритмов синтеза допустимых маневров, обеспечивающих достижение заданных конечных условий, экономию топлива и строгое соблюдение динамических ограничений по тяге, энергетических лимитов и геометрических ограничений на ориентацию. В основу исследования положены методы классической небесной механики в рамках задачи двух тел. Состояние космического аппарата описывается как в инерциальной системе координат ECI, так и в локальной орбитальной системе RTN. Методология базируется на последовательном применении трех моделей: импульсной, градиентной и низкотяговой. В результате работы разработан комплекс математических моделей, устанавливающих аналитическую связь между компонентами воздействия в системе RTN и изменениями шести кеплеровых элементов орбиты. Определены оптимальные точки приложения силы для минимизации побочных возмущений элементов орбиты. Важным достижением стала формализация условий допустимости маневров посред-

ством описания множества допустимых управлений. Такие условия учитывают ограничения по модулю и скорости изменения тяги, лимиты расхода массы на основе формулы Циолковского, а также требования для ориентации вектора тяги. Итогом исследования стала комплексная методология, объединяющая теоретические основы баллистики с практическими требованиями к конструкции космических аппаратов. Научная новизна заключается в разработке критериев, позволяющих преобразовать теоретический расчет в реализуемый полетный план с учетом технических ограничений. Практическая значимость исследования заключается в создании комплексной методологии, которая объединяет классическую теорию небесной механики с реальными инженерными требованиями к космическим аппаратам.

Ключевые слова: орбитальные маневры, импульсный маневр, градиентный маневр, низкотяговый маневр, кеплеровы элементы, допустимый маневр, деорбитация, орбита

Введение. Актуальность исследований орбитальных маневров связана с ростом числа космических аппаратов (КА) и усложнением задач на орбите, включая автономное сближение и стыковку, обслуживание КА и уклонение от космического мусора. Поэтому переход от традиционного планирования маневров к точным алгоритмам управления движением КА в реальном времени становится особенно важным. Это связано с тем, что при выполнении маневров необходимо учитывать ограничения по тяге, запасу энергии и ориентации бортовой аппаратуры, что требует разработки новых математических моделей и критериев допустимости маневров.

Литературный обзор. Современное моделирование орбитальных маневров КА включает в себя разнообразные методы: от классических аналитических приближений до сложных алгоритмов наведения, работающих в режиме реального времени. В работе (Baranova et al., 2025) приведены результаты исследований параметров копланарного сближения на околокруговых орбитах, где показано, что применение трехэтапного алгоритма (аналитический расчет импульсов, их распределение и оптимизация) позволяет минимизировать суммарную характеристическую скорость при использовании двигателей малой тяги. Для задач формирования группового полета в работе (Wang et al., 2025) предложен аналитический закон наведения на базе концепции «искусственной гравитации», обеспечивающий синтез устойчивых траекторий облета в форме пространственных эллипсов. В исследовании (Fan et al., 2024) продемонстрирована эффективность применения параметризации кривыми Безье для быстрой генерации траекторий межпланетных перелетов с учетом динамических ограничений тяги. В свою очередь, в работе (Sanchez et al., 2020) показано, что использование теории дифференциальной плотности и В-сплайнов позволяет трансформировать дифференциальные уравнения связанного движения в эквивалентную алгебраическую форму, существенно упрощая процесс планирования.

Отдельным важным направлением является учет совокупности динамических и геометрических ограничений при реализации маневров. В работе (Sumarokov et al., 2019) представлен алгоритм BVLS (Bounded Variables Least Squares) для оптимизации конфигурации включения 30 двигателей причаливания и ориентации, обеспечивающий точность стабилизации в условиях жестких ограничений в каждом такте управления. Вопросы соблюдения «запретных конусов» (keep-out zones) для защиты оптических датчиков от засветки рассмотрены в работах (Celani et al., 2021; Ailabouni, et al., 2023), где доказано, что использование функций Слепиана и геометрических методов оптимизации на многообразиях позволяет гарантировать безопасность наведения приборов на яркие небесные тела. Проблема исключения загрязнения чувствительных поверхностей выхлопом двигателей успешно решена в исследовании (Kazi et al., 2025) с помощью смешанно-целочисленного линейного программирования. Для обеспечения мягкой стыковки в работе (Di Cairano et al., 2012) формализованы условия предиктивного управления на вращающихся платформах, учитывающие предельные значения относительных скоростей и ограничения.

Однако остаются нерешенными вопросы, связанные с обеспечением гарантированной вычислительной устойчивости и безопасности в условиях высокой неопределенности параметров и сложной конфигурации препятствий. В работе (Fiori et al., 2024) отмечается, что классические методы на базе искусственных потенциальных полей склонны к попаданию в локальные минимумы, что требует разработки дополнительных алгоритмов сглаживания траекторий. Причиной этого могут быть высокая чувствительность методов оценки вероятности столкновений к ошибкам навигационных данных (Pavanello et al., 2025) и ограниченность допущений о линейности динамики, используемых в концепции инвариантных множеств для обхода препятствий (Frey et al., 2018). Высокая вычислительная стоимость нелинейных моделей связанного движения также ограничивает их применение в бортовых системах реального времени.

Все это позволяет утверждать, что целесообразным является проведение исследования, посвященного разработке математических моделей и алгоритмов синтеза допустимых маневров КА, которые обеспечивали бы выполнение терминальных условий и экономию топлива при строгом соблюдении совокупности динамических ограничений тяги, энергетических лимитов и геометрических запретов на ориентацию в режиме реального времени.

Целью исследования является разработка моделей допустимых маневров КА с учетом комплекса ограничений по тяге, ориентации и расходу топлива для повышения автономности, и безопасности навигационных операций в сложных динамических условиях.

Методы исследования. В данном исследовании качестве теоретической основы данного исследования используются методы классической небесной механики. Базовое движение КА рассматривается в рамках задачи двух тел.

Для математического описания состояния системы применяется инерциальная система координат ECI (Earth Centered Inertial), а для физической интерпретации и задания параметров маневрирования – локальная орбитальная система координат RTN (Radial, Transverse, Normal). Состояние КА в каждый момент времени определяется как векторами положения и скорости, так и набором из шести классических кеплеровых элементов, что обеспечивает наглядное отслеживание изменений формы и ориентации орбиты.

Методология получения результатов основана на последовательном применении трех математических моделей маневрирования. Во-первых, импульсная модель идеализирует мгновенное изменение скорости, что полезно для предварительного баллистического анализа и оценки бюджета ΔV . Во-вторых, градиентная модель описывает маневры конечной длительности, учитывая реальные профили тяги (ступенчатые, линейные или сглаженные), изменение массы КА и движение по орбите непосредственно во время работы двигателя. В-третьих, низкотяговая модель предназначена для моделирования длительных воздействий, когда тяга активна в течение часов или суток, что требует непрерывного контроля программы управления вектором тяги.

Результаты. Положение КА описывается радиус-вектором $\mathbf{r}(t)$, а скорость – вектором $\mathbf{v}(t)$. Эти векторы задаются в ECI, то есть в системе, которая «не вращается» вместе с Землей и удобна для записи уравнений движения (Ivanov et al., 2016; Curtis et al., 2020).

Вектор состояния КА в каждый момент времени определяется следующим образом:

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{r}(t) \\ \mathbf{v}(t) \end{bmatrix}, \quad (1)$$

Если временно игнорировать тягу и возмущения (атмосферу, J_2 , Солнце, Луну и т.д.), то движение описывается задачей двух тел:

$$\dot{\mathbf{r}} = \mathbf{v}, \quad \dot{\mathbf{v}} = -\mu \frac{\mathbf{r}}{r^3}, \quad (2)$$

где μ – гравитационный параметр центрального тела (Земли), $r = \|\mathbf{r}\|$. Это «фон», на который накладывается маневр.

Здесь именно маневр создает добавочное ускорение или мгновенное изменение скорости. Когда мы говорим «дать импульс вдоль скорости» или «дать импульс в сторону нормали к плоскости орбиты», в инерциальных осях это неудобно, потому что направления постоянно меняются. Поэтому в орбитальной механике вводят локальную систему координат, связанную с текущей орбитой в данной точке. Здесь используется орбитальная система RTN:

Эта система определяется через векторы $\mathbf{r}$ и $\mathbf{v}$:

- единичный радиальный вектор:

$$\hat{\mathbf{e}}_R = \frac{\mathbf{r}}{\|\mathbf{r}\|}, \quad (3)$$

- вектор орбитального момента:

$$\mathbf{h} = \mathbf{r} \times \mathbf{v}, \quad \hat{\mathbf{e}}_N = \frac{\mathbf{h}}{\|\mathbf{h}\|}, \quad (4)$$

Направление $\hat{\mathbf{e}}_N$ показывает куда направлена нормаль к плоскости мгновенной орбиты.

Касательное направление определяется так, чтобы получилась правая тройка:

$$\hat{\mathbf{e}}_T = \hat{\mathbf{e}}_N \times \hat{\mathbf{e}}_R, \quad (5)$$

Смысл RTN простой: в любой точке орбиты мы можем разложить любое воздействие (например, импульс) на три физически понятные компоненты: радиальную, касательную и нормальную. Именно в этой системе будем задавать компоненты маневра.

Теперь рассмотрим три типа моделей маневров:

- импульсный;
- градиентный;
- низкотяговый.

Первый и самый распространенный тип модели – импульсная. Здесь предполагается, что воздействие тяги происходит настолько быстро, что за время включения положение $\mathbf{r}$ практически не меняется, а изменяется только скорость. Тогда маневр в момент t_0 записывается как скачок:

$$\mathbf{r}^+ = \mathbf{r}^-, \quad \mathbf{v}^+ = \mathbf{v}^- + \Delta \mathbf{v}, \quad (6)$$

где символы «-» и «+» означают значения до и после импульса. Следующий шаг – задать $\Delta \mathbf{v}$ не случайным образом, а через компоненты в RTN. Поэтому импульс $\Delta \mathbf{v}$ выразим в таком виде:

$$\Delta \mathbf{v} = \Delta V_r \hat{\mathbf{e}}_R + \Delta V_t \hat{\mathbf{e}}_T + \Delta V_N \hat{\mathbf{e}}_N, \quad (7)$$

где ΔV_r – радиальная компонента импульса (вдоль $\hat{\mathbf{e}}_R$), ΔV_t – тангенциальная (касательная) компонента импульса (вдоль $\hat{\mathbf{e}}_T$), ΔV_N – нормальная компонента импульса (вдоль $\hat{\mathbf{e}}_N$). Именно эти три компоненты $\Delta V_r, \Delta V_t, \Delta V_N$ в дальнейшем являются «управляющими параметрами» импульсного маневра.

Чтобы формализовать изменение орбиты после маневра, будем использовать классические кеплеровы элементы: $a, e, i, \Omega, \omega, f$. Здесь a –

большая полуось, e – эксцентриситет орбиты, i – наклонение, Ω – прямое восхождение (иногда долгота) восходящего узла, ω – аргумент перигея, f – истинная аномалия. После маневра эти величины изменятся на величину: $\Delta a, \Delta e, \Delta i, \Delta \Omega, \Delta \omega, \Delta f$.

Влияние импульсов на орбиту определяются следующим образом. Тангенциальный импульс ΔV_t в первую очередь влияет на скорость, а значит и на большую полуось. Поэтому именно ΔV_t рассматривается как основной инструмент для изменения большой полуоси. Нормальный импульс ΔV_N поднимает или опускает плоскость орбиты к плоскости экватора, поэтому он отвечает за изменение наклонения i и долготы узла Ω .

При этом необходимо отметить, что орбитальные элементы связаны между собой не линейно. Поэтому, даже если целимся только в один элемент, неизбежно может измениться другой. В частности: Δa определяется только ΔV_t , Δi и $\Delta \Omega$ определяются только ΔV_N . Однако импульсы в целом меняют и e , и ω , т.е. возникают перекрестные изменения формы и ориентации эллипса. Эта оговорка принципиальна для вопроса допустимых маневров: когда мы вводим ограничения вида $|\Delta e| \leq \Delta e_{\max}$ или $|\Delta \omega| \leq \Delta \omega_{\max}$, их нельзя игнорировать, даже если главная цель маневра изменение a или i .

Итак, в импульсной модели маневр задается вектором Δv , обычно через $\Delta V_t, \Delta V_r, \Delta V_N$ в RTN.

В градиентной модели (маневр конечной длительности с профилем тяги) маневр задается уже не одним скачком, а функциями тяги $T(t)$ и направлением $u(t)$ на интервале времени. Но и там конечный эффект часто оценивают как эквивалентный Δv .

В низкотяговой модели математическая форма та же, что и у конечной длительности, но характерный интервал времени существенно больше, и ограничения по ориентации и тяге действуют во времени, а не в одну точку. Все эти случаи используют одни и те же базовые определения: вектор состояния, RTN-разложение и набор элементов орбиты с их приращениями.

Импульсные маневры. Импульсный маневр – это удобная и очень распространенная идеализация, когда воздействие двигательной установки считается настолько коротким, что за время включения положение КА практически не меняется, а изменится только скорость. Такая модель особенно полезна в задачах баллистического анализа, планирования коррекций, оценок ΔV -бюджета и в алгоритмах обнаружения маневров по орбитальным данным: мы видим скачок параметров орбиты и интерпретируем его как импульс Δv .

Рассмотрим математическую модель импульса. В момент t_0 импульсный маневр записывается в виде выражения (6), а импульс Δv выразим в виде (7). Далее рассмотрим влияние маневров на элементы орбиты. В таблице 1 даны выражения для определения изменений элементов орбиты под действием различных импульсов (Ivanov et al., 2016; Curtis et al., 2020; Gurfil et al., 2006, Montenbruck et al., 2010).

Таблица 1— Изменения элементов орбиты в результате маневров

Вид импульса	Изменение элементов орбиты
Тангенциальный ΔV_t	$\Delta a = \frac{2a^2 v}{\mu} \cdot \Delta V_t$ $\Delta e_t = \frac{2a(e + \cos f)v}{\mu} \cdot \Delta V_t$ $\Delta \omega_t = \frac{2a \sin f v}{e \mu} \cdot \Delta V_t$ $\Delta f = -\Delta \omega_t$
Нормальный ΔV_N	$\Delta i = \frac{\cos u}{v} \cdot \Delta V_N$ $\Delta \Omega = \frac{\sin u}{v \sin i} \cdot \Delta V_N$ $\Delta \omega_N = \frac{\sin u}{v \cos f \operatorname{tg} i} \cdot \Delta V_N$ $\Delta f = -\Delta \omega_N$
Радиальный ΔV_r	$\Delta e_r = \frac{r \sin f}{a v} \cdot \Delta V_r$ $\Delta \omega_r = -\left(2 + \frac{r}{a e} \cos f\right) \cdot \frac{\Delta V_r}{v}$ $\Delta f = -\Delta \omega_r$

Здесь изменения угловых величин даны в радианах; Δe_t , Δe_r – изменения эксцентриситета под действием тангенциального и радиального импульсов, $\Delta \omega_t$, $\Delta \omega_N$, $\Delta \omega_r$ – изменения аргумента перицентра под действием тангенциального, нормального и радиального импульсов. Как видно из этой сводки формул изменение большой полуоси Δa происходит только под действием тангенциального импульса ΔV_t , изменения наклона Δi и долготы восходящего узла $\Delta \Omega$ происходят только под действием нормального импульса ΔV_N . Но подача этих импульсов приведет к изменению и эксцентриситета (e), и аргумента перицентра (ω).

Следующим важным вопросом при выполнении маневров является выбор точки приложения импульсов.

Рассмотрим выбор точки приложения тангенциальных импульсов ΔV_t . Данный импульс выполняется для коррекции большой полуоси. Для изменения этого параметра наиболее эффективно действовать вдоль скорости. На практике используют схему из двух тангенциальных импульсов. ΔV_t разбивается на две части пополам, т.е. $\Delta V_{t1} = \Delta V_{t2} = \Delta V_t / 2$. В точке перигея ($f=0$) подается первая часть импульса ΔV_{t1} . В точке апогея ($f=180$ градусов) подается вторая часть импульса ΔV_{t2} . Это делается для того, чтобы предотвратить увеличение эксцентриситета в результате первого импульса.

Рассмотрим выбор точки приложения нормальных импульсов ΔV_N . Для коррекции наклона Δi эффективным будет точка с $\cos u=1$, т.е. при $u=0$ или $\cos u=-1$, т.е. при $u=180^\circ$. В этих точках $\sin u=0$, соответственно изменение $\Delta \Omega=0$, т.е. мы избежим изменения долготы восходящего узла Ω . А

для эффективного изменения этого элемента необходимо подать нормальный импульс в точках, где $\sin u = 1$ или $\sin u = -1$, т.е. в точках $u = 90^\circ$ или $u = 270^\circ$.

Рассмотрим выбор точки приложения радиальных импульсов ΔV_r . Радиальный импульс на практике используется не часто. Его можно использовать для изменения эксцентриситета (e) и аргумента перигентра (ω). Как видно из вышеуказанных формул, этот импульс целесообразно применить вблизи точки $f = 90^\circ$.

Градиентные маневры. Между идеализацией «мгновенного» импульса и длительным низкотяговым управлением есть промежуточный класс – включения конечной длительности, когда тяга действует заметное время, но все еще речь идет о коротком маневре (минуты, десятки минут), а не о многосуточной низкотяговой спирали. Здесь под термином «градиентные маневры» будем понимать именно такие включения конечной длительности с заданным профилем тяги $T(t)$, в том числе линейным, то есть «нарастающим/спадающим» – отсюда и слово «градиентный» в инженерном смысле (Chobotov, et al., 2009).

Данный класс маневров выделяется отдельно, поскольку в реальных двигательных установках их реализация связана с рядом ограничений:

- тяга не всегда включается ступенькой и мгновенно;
- направление тяги может меняться в процессе наведения;
- за время включения КА успевает пройти заметный участок орбиты;
- масса меняется, а значит меняется и ускорение $\frac{T}{m}$;
- возникает отличие между идеальным Δv и тем, что реально получается в инерциальном пространстве.

В импульсной модели говорили, что скорость скачком меняется на Δv . Это удобно, но подразумевает, что все произошло в одной точке орбиты.

В градиентной модели говорим иначе: в течение времени $[t_0, t_0 + \tau]$ двигатель создает силу тяги $T(t)$, а значит – ускорение $a_T(t)$. Поэтому уже нельзя честно писать $r^+ = r^-$: положение и ориентация орбитальных осей успевают измениться. То есть маневр теперь – это не одно число Δv , а процесс, задаваемый функциями: модуль тяги $T(t) = \|T(t)\|$, направление тяги $u(t)$ (единичный вектор), длительность τ , и, при необходимости, законом изменения тяги (ступенька, линейный градиент, сглаживание).

Теперь рассмотрим уравнение движения для этой модели. Запишем уравнение движения в модели ЕСІ. Тогда:

$$\dot{\mathbf{r}} = \mathbf{v}, \quad (8)$$

$$\dot{\mathbf{v}} = -\mu \frac{\mathbf{r}}{r^3} + \mathbf{a}_T(t), \quad (9)$$

$$\mathbf{a}_T(t) = \frac{T(t)}{m(t)} \mathbf{u}(t), \quad (10)$$

где $u(t)$ – единичный вектор направления тяги, $m(t)$ – текущая масса КА.

Если требуется учитывать расход топлива, используется стандартная модель:

$$\dot{m}(t) = -\frac{T(t)}{I_{sp}g_0}, \quad (11)$$

где I_{sp} – удельный импульс, g_0 – стандартное ускорение свободного падения. Эта связка прямо отражает идею ΔV -бюджета и зависимости маневренности от топлива.

На практике встречаются следующие типовые варианты профиль тяги.

1. Ступенчатое включение:

$$T(t) = T_0, t \in [t_0, t_0 + \tau], \quad (12)$$

Здесь тяга постоянная, а направление чаще всего зафиксировано.

2. Линейный (градиентный) разгон (или спад):

$$T(t) = T_0 + k(t - t_0), t \in [t_0, t_0 + \tau], \quad (13)$$

где k – темп нарастания/убывания тяги. Этот случай хорошо соответствует смыслу слова «градиентный»: тяга меняется по градиенту (линейно), а не скачком.

3. Сглаженный профиль. Это когда профиль $T(t)$ задается по определенной закономерности. Конкретная формула выбирается по требованиям двигательной установки и системы управления ориентацией.

Хотя маневр является процессом конечной длительности, почти всегда удобно оценивать его результат в виде эквивалентного прироста скорости. Векторно это записывается так:

$$\Delta v \approx \int_{t_0}^{t_0 + \tau} a_T(t) dt = \int_{t_0}^{t_0 + \tau} \frac{T(t)}{m(t)} u(t) dt. \quad (14)$$

Низкотяговые маневры. Низкотяговым называют маневры, при которых ускорение от тяги мало по сравнению с гравитационным ускорением, но тяга действует долго – часы, сутки, недели и даже месяцы. По форме уравнений это тот же «маневр конечной длительности», что и в предыдущем подразделе, однако, по сути, это уже другой режим полета: орбита меняется не скачком и не коротким включением, а постепенно, за счет накопления малого ускорения на большом интервале времени.

Именно поэтому в перечне «импульсные – градиентные – низкотяговые» это тип маневра стоит отдельным классом: он требует иной логики планирования, и ограничения по ориентации начинают играть решающую роль, потому что их надо выдерживать непрерывно или почти непрерывно.

В инженерном смысле термин «низкая тяга» не только означает маленькую тягу, но и дает следующие последствия.

1. Ускорение тяги мало, поэтому за небольшой промежуток времени изменение скорости невелико:

$$\|\Delta \mathbf{v}\| \approx \left\| \int \frac{T}{m} \mathbf{u} dt \right\| \quad (15)$$

и если τ мал, то $\Delta \mathbf{v}$ тоже мал. Поэтому низкотяговый маневр всегда требует времени, результат достигается накоплением эффекта за длительное время.

2. Орбита меняется непрерывно, часто по спиральному или квазиспиральному типу (например, при подъеме/спуске по высоте), а не по схеме двух импульсов как у Хоманна.

3. Модель допустимости становится динамической, т.е. недостаточно давать импульс $\Delta \mathbf{V}$, но и нужно гарантировать, что КА способен создавать тягу нужной величины, держать разрешенное направление тяги, и делать это длительное время, не нарушая ориентационных ограничений.

Теперь рассмотрим математическую модель низкотяговых маневров. Здесь уравнения те же, но смысл другой. С точки зрения математики уравнения движения такие же, как и для градиентного маневра конечной длительности (Vallado, et al., 2013; Hu, et al., 2015):

$$\dot{\mathbf{r}} = \mathbf{v}, \quad (16)$$

$$\dot{\mathbf{v}} = -\mu \frac{\mathbf{r}}{r^3} + \frac{T(t)}{m(t)} \mathbf{u}(t), \quad (17)$$

$$\dot{m}(t) = -\frac{T(t)}{I_{sp} g_0}, \quad (18)$$

Но отличие в том, что:

- интервал управления $[t_0, t_1]$ становится длинным;
- $\mathbf{u}(t)$ может целенаправленно меняться на протяжении маневра;
- часто нельзя пренебрегать ни изменением массы, ни поворотом локальных орбитальных осей, ни накоплением ошибок.

И если в импульсном случае маневр можно описывать одним вектором $\Delta \mathbf{v}$, то в случае низкотяговых маневров более корректно говорить: маневр задается как программа управления $(T(t), \mathbf{u}(t))$ на всем интервале.

Формально мы все еще можем записать эквивалентный прирост скорости:

$$\Delta \mathbf{v} = \int_{t_0}^{t_1} \frac{T(t)}{m(t)} \mathbf{u}(t) dt, \quad (19)$$

Однако важно понимать: при длительном управлении это выражение уже не является простой оценкой, по нескольким причинам: направление $u(t)$ выбирается по цели маневра и ограничениям; вектор $u(t)$ может задаваться относительно корпуса, а корпус ориентируется относительно орбиты, которая сама меняется; масса существенно убывает, ускорение растет, и т.д. В итоге, в этих задачах обычно интересует не просто ΔV , а достижимость целевого перехода при заданных ограничениях: возможно ли, имея $T_{\max}$, I_{sp} , ограничения ориентации, перевести КА из одной орбиты в другую.

Чтобы класс низкотяговых маневров не выглядел абстрактно, полезно перечислить типовые цели, которые дальше лягут на формулы допустимости:

- подъем или спуск орбиты по высоте (например, увеличение a или высоты круговой орбиты);
- изменение наклона i и плоскости орбиты;
- фазирование, или изменение положения по орбите;
- комбинированные переводы (при необходимости одновременного изменения и размера орбиты, и ее плоскости. В импульсной баллистике это иногда делают «комбинированным» импульсом, а в низкотяговых маневрах – программой управления с меняющимся направлением тяги).

Теперь рассмотрим вопрос допустимости низкотяговых маневров во времени. Самое важное для нашей задачи состоит в следующем:

- для импульса достаточно проверить: можем ли мы реализовать Δv нужной величины и направления (и в нужной точке орбиты);
- для градиентного маневра уже нужно проверять ограничения на интервале включения $[t_0, t_0 + \tau]$;
- для низкотяговых маневров это становится еще строже: ограничения должны выполняться на длинном интервале $[t_0, t_1]$, возможно с сотнями, или даже тысячами шагов дискретизации.

То есть допустимый низкотяговый маневр – это не один вектор Δv , а класс программ управления $(T(t), u(t))$, которые удовлетворяют ограничениям:

$$0 \leq T(t) \leq T_{\max}, \quad (20)$$

$u(t)$ допускается ориентацией КА (конус, запреты, скорость разворота),

$$m(t) \geq m_{\min}, \quad (21)$$

и, если нужно, ограничения по скорости изменения тяги и по скорости поворота.

Также необходимо рассмотреть формулы допустимых маневров с учетом ограничений по тяге и ориентации.

Смысл термина «допустимый маневр» в инженерной постановке такой: маневр считается допустимым, если существует реальное управление двигателем и ориентацией, которое не:

- превышает возможностей двигательной установки (тяга/мощность/скорость изменения тяги);
- нарушает ограничений по ориентации КА (куда можно направлять тягу и как быстро можно разворачиваться);
- нарушает топливные ограничения и при этом создает требуемый эффект (в импульсной модели – нужный Δv ; в более общей – нужный переход орбиты или заданные приращения элементов).

Далее необходимо задать ограничения математически, затем ввести множество допустимых управлений U , а после этого получим формулы множества допустимых маневров (по сути – множества достижимых Δv или достижимых приращений элементов) для каждого из трех классов (Vallado, et al., 2013; Hu, et al., 2015).

Ограничение по тяге. Начнем с самого базового: двигательная установка не может выдавать тягу больше некоторого максимума. Поэтому для всех неимпульсных моделей (конечной длительности и низкой тяги) мы вводим ограничение:

$$0 \leq T(t) \leq T_{max}, \quad (22)$$

Если двигательная установка или система питания не позволяет мгновенно менять тягу, добавляют ограничение на скорость изменения тяги:

$$|\dot{T}(t)| \leq \dot{T}_{max}, \quad (23)$$

Ограничение по топливу и ΔV -бюджет. Даже если тяга по модулю допустима, маневр может быть невозможен из-за недостатка топлива. В классической форме это выражается через связь тяги, расхода массы и удельного импульса:

$$\dot{m}(t) = -\frac{T(t)}{I_{sp} g_0}, \quad (24)$$

Интегрально это дает ограничение на достижимую суммарную скорость изменения (« ΔV -бюджет»), которое обычно записывают через формулу:

$$\Delta v_{\text{топ}} \leq I_{sp} g_0 \ln \left(\frac{m_0}{m_f} \right), \quad (25)$$

где m_0 – масса до маневров, m_f – масса после, $m_f \geq m_{\min}$. Смысл $\Delta v_{\text{топ}}$ в том, что – это максимально доступный запас изменения скорости (ΔV -budget), который КА может получить за счет имеющегося топлива. Такая связка как

раз и используется для контроля ΔV -бюджета и допустимости маневра по топливу.

Важно подчеркнуть: это ограничение работает вне зависимости от того, импульсный маневр или низкотяговый. Разница лишь в том, что в импульсной модели мы обычно задаем Δv как итоговый вектор, а в конечной длительности (низкотяговый) Δv получается интегралом ускорения. Но ограничение по топливу одинаково ограничивает достижимый эффект.

Ограничение по ориентации. Следующим ключевым моментом является ограничение по ориентации. В реальном КА вектор тяги не может быть направлен куда угодно: он связан с геометрией двигателя и ориентацией КА.

Введем обозначения:

- $b(t)$ – единичный вектор разрешенной оси в системе координат корпуса (например, продольная ось сопла или ось, вдоль которой должна быть направлена тяга);

- $u(t)$ – единичный вектор направления тяги в инерциальной системе, то есть

$$a_T(t) = \frac{T(t)}{m(t)} u(t), \quad \|u(t)\| = 1, \quad (26)$$

Если двигатель можно отклонять, то $u(t)$ может отличаться от $b(t)$, но в пределах допустимого угла. Наиболее универсальная математическая запись такого ограничения – конус:

$$u(t) \cdot b(t) \geq \cos \theta_{max}, \quad (27)$$

где θ_{max} – максимально допустимое отклонение направления тяги от оси $b(t)$. Это неравенство означает, что угол между $u(t)$ и $b(t)$ не должен превышать θ_{max} . То есть возможные направления тяги образуют конус вокруг оси $b(t)$.

Множество допустимых управлений U . Теперь мы можем собрать все вместе и формально определить, что такое допустимое управление.

Будем считать, что управление – это пара функций $(T(t), u(t))$ на заданном интервале времени. Тогда множество допустимых управлений:

$$U = \{(T(t), u(t)) : 0 \leq T(t) \leq T_{max}, \\ \|u(t)\| = 1, u(t) \cdot b(t) \geq \cos \theta_{max}, m(t) \geq m_{min}\}, \quad (28)$$

и при необходимости добавляются ограничения:

$$|\dot{T}| \leq \dot{T}_{max}, \quad \|\omega\| \leq \omega_{max}, \quad \|\dot{\omega}\| \leq \alpha_{max}, \quad (29)$$

Допустимые импульсные маневры. В импульсной модели маневр задается вектором Δv (или компонентами $\Delta V_r, \Delta V_p, \Delta V_N$).

а) Ограничение по ориентации для импульса. Направление импульса Δv должно быть достижимо ориентацией КА в момент маневра. В терминах конуса это записывается так же, как и для непрерывной тяги:

$$\frac{\Delta v}{\|\Delta v\|} \cdot b(t_0) \geq \cos \theta_{\max}, \quad (30)$$

То есть направление импульса должно лежать внутри допустимого конуса вокруг оси b в момент t_0 .

б) Ограничение по тяге для импульса через эквивалентную длительность. Импульсная модель сама по себе не содержит $T_{\max}$, поэтому вводят физически осмысленную связь: любой реальный импульс есть результат конечного включения длительностью τ . Тогда:

$$\Delta v = \int_{t_0}^{t_0+\tau} \frac{T(t)}{m(t)} u(t) dt, \quad (31)$$

Если мы хотим получить простую верхнюю оценку, можно записать:

$$\|\Delta v\| \leq \int_{t_0}^{t_0+\tau} \frac{T_{\max}}{m(t)} dt \approx \frac{T_{\max}}{m} \tau \quad (\text{если } m \gg \text{const}), \quad (32)$$

Эта оценка задает верхний физический предел: при данной массе и максимальной тяге невозможно получить импульс больше, чем $T_{\max} \tau / m$ за время τ .

Итак, множество допустимых импульсных маневров можно сформулировать следующим образом:

$$M_{\text{imp}} = \left\{ \Delta v: \frac{\Delta v}{\|\Delta v\|} \cdot b(t_0) \geq \cos \theta_{\max}, \right. \\ \left. \|\Delta v\| \leq \int_{t_0}^{t_0+\tau} \frac{T_{\max}}{m(t)} dt, \|\Delta v\| \leq \Delta v_{\text{топ}} \right\}, \quad (33)$$

где $\Delta v_{\text{топ}} = I_{sp} g_0 \ln \left(\frac{m_0}{m_f} \right)$. Как было сказано выше, $\Delta v_{\text{топ}}$ – это максимально доступный запас изменения скорости (Δv -budget), который КА может получить за счет имеющегося топлива.

Допустимые градиентные маневры. Здесь маневр – это управление $(T(t), u(t))$ на интервале $[t_0, t_0+\tau]$ с заданным профилем тяги (ступенька, линейный и т.д.). Результат маневра выражается через интеграл:

$$\Delta v = \int_{t_0}^{t_0+\tau} \frac{T(t)}{m(t)} u(t) dt, \quad (34)$$

Маневр допустим, если существует управление $(T, u) \in U$ на этом интервале. Тогда множество допустимых градиентных маневров:

$$M_{\text{grad}} = \left\{ \Delta v: \Delta v = \int_{t_0}^{t_0+\tau} \frac{T(t)}{m(t)} u(t) dt, (T, u) \in U \right\}, \quad (35)$$

Если профиль $T(t)$ считается заданным (например, линейный градиент), то ограничение $0 \leq T(t) \leq T_{\max}$ превращается в проверку параметров профиля (чтобы нигде не превысить $T_{\max}$) и, при наличии, $|\dot{T}| \leq \dot{T}_{\max}$ (чтобы градиент не был слишком крутым).

Важное отличие от импульса в том, что ориентационное ограничение $u(t) \cdot b(t) \geq \cos \theta_{\max}$ должно выполняться для всех $t \in [t_0, t_0 + \tau]$, а не только в одной точке.

Допустимые низкотяговые маневры. Низкотяговый маневр отличается от градиентного не формой уравнений, а масштабом времени: управление задается на длинном интервале $[t_0, t_1]$, и направление тяги может целенаправленно меняться по программе.

Результат в виде эквивалентного Δv :

$$\Delta v = \int_{t_0}^{t_1} \frac{T(t)}{m(t)} u(t) dt, \quad (36)$$

И тогда множество допустимых низкотяговых маневров:

$$M_{LT} = \left\{ \Delta v: \Delta v = \int_{t_0}^{t_1} \frac{T(t)}{m(t)} u(t) dt, (T, u) \in U \text{ на } [t_0, t_1] \right\}, \quad (37)$$

Здесь наиболее строгим становится именно ориентационное условие: даже если по тяге и топливу маневр возможен, он может оказаться недостижимым из-за того, что КА не может долго удерживать требуемое направление тяги (например, из-за требований на ориентацию солнечных батарей или антенн).

Закключение. В результате данного исследования разработана комплексная методология моделирования орбитальных маневров, объединяющая классическую теорию небесной механики с практическими инженерными требованиями к КА. Использование трех различных моделей – импульсной, градиентной и низкотяговой – обеспечивает необходимую гибкость расчетов: от оперативной оценки затрат топлива на коррекцию орбиты до детального планирования многосуточных перелетов. Важным элементом методологии является разложение управляющих воздействий в локальной системе координат RTN, что позволяет напрямую связать работу двигательной установки с изменениями параметров орбиты, таких как большая полуось, наклонение и долгота восходящего узла. Анализ показал, что теоретическая эффективность маневра напрямую зависит от

выбора оптимальных точек приложения силы, таких как перигей, апогей или узлы орбиты, для минимизации нежелательных возмущений других элементов. Главным результатом работы стала формализация условий допустимости маневров, позволяющая преобразовать математический расчет в реальный полетный план. Учет ограничений, связанных с максимальной тягой двигателя, предельным запасом топлива (рассчитанным на основе формулы Циолковского) и допустимым конусом ориентации КА, гарантирует исполнимость маневра бортовыми системами. Такой комплексный подход позволяет не только рассчитывать требуемые приращения скорости, но и предварительно исключать технически нереализуемые для конкретного КА сценарии, обеспечивая высокую надежность и точность планирования космических миссий.

References

- Baranova A.A., & Oliviob A.P. (2025) Modified algorithm for calculating the parameters of maneuvers of coplanar meeting of spacecraft in a near-circular orbit using low-thrust engines. *Vestnik RUDN. Engineering Researches*, 26(2). — P.113–126 (in Eng.). <https://doi.org/10.22363/2312-8143-2025-26-2-113-126>
- Wang X., Hu M., Xin C., & Zheng S. (2025) A novel analytical approach for spacecraft fly-around formation design with a low-thrust maneuver. *Aerospace*, 12(5). — 361 p. (in Eng.). <https://doi.org/10.3390/aerospace12050361>
- Fan Z., Ke W., Qi J., Huo M., Qi N., Cheng F., ... & Xie S. (2024) Initial trajectory design of low-thrust spacecraft considering attitude constraints. *Scientific Reports*, 14(1). — 20246 p. (in Eng.). <https://doi.org/10.1038/s41598-024-70830-8>
- Sanchez J. C., Gavilan F., Vazquez R., & Louembet C. (2020) A flatness-based predictive controller for six-degrees of freedom spacecraft rendezvous. *Acta Astronautica*, 167. — P. 391–403 (in Eng.). <https://doi.org/10.1016/j.actaastro.2019.11.026>
- Sumarokov A. V., & Tyrnov P. A. (2019) Algorithm for spacecraft angular and translational motion control with use of orientation thrusters. *Vestnik MGTU im. N. E. Bauman. Instrument Engineering Series*, (6). — P. 30–40 (in Rus.). DOI: 10.18698/0236-3933-2019-6-30-40
- Celani F., & Bruni R. (2021). Minimum-time spacecraft attitude motion planning using objective alternation in derivative-free optimization. *Journal of Optimization Theory and Applications*, 191(2). — P. 776–793 (in Eng.). <https://doi.org/10.1007/s10957-021-01834-x>
- Ailabouni D., Meister A., & Spindler K. (2023) Attitude maneuvers avoiding forbidden directions. *Astrodynamics*, 7(3). — P. 351–362 (in Eng.). <https://doi.org/10.1007/s42064-023-0163-7>
- Kazi S. R., Nagarajan H., Hijazi H., & Wozniak P. (2025) Optimal trajectory planning for space object tracking with collision-avoidance constraints. In *Proceedings of the 2025 IEEE Conference on Control Technology and Applications (CCTA)*. — P. 564–570. IEEE (in Eng.). <https://doi.org/10.48550/arXiv.2506.22797>
- Di Cairano S., Park H., & Kolmanovsky I. (2012) Model predictive control approach for guidance of spacecraft rendezvous and proximity maneuvering. *International Journal of Robust and Nonlinear Control*, 22(12). — P. 1398–1427 (in Eng.). <https://doi.org/10.1002/rnc.2827>
- Fiori S., Rachiglia F., Sabatini L., & Sampaolesi E. (2024) Modeling, simulation and control of a spacecraft: Automated rendezvous under positional constraints. *Aerospace*, 11(3). — 245 p. (in Eng.). <https://doi.org/10.3390/aerospace11030245>
- Pavanello Z., De Maria L., De Vittori A., Maestrini M., Di Lizia P., & Armellin R. (2025) CAMmary: A review of spacecraft collision avoidance manoeuvre design methods. *Acta Astronautica* (in Eng.). <https://doi.org/10.1016/j.actaastro.2025.07.027>
- Frey G. R., Petersen C. D., Leve F. A., Girard A. R., & Kolmanovsky I. V. (2018) Invariance-

based spacecraft relative motion planning incorporating bounded disturbances and minimum thrust constraints. In Proceedings of the 2018 American Control Conference (ACC). — P. 658–663. IEEE (in Eng.). DOI: *10.23919/ACC.2018.8430790*

Ivanov N.M., & Lysenko L.N. (2016) *Ballistika i navigaciya kosmicheskikh apparatov* [Ballistics and navigation of spacecraft]. Moscow: Bauman MSTU Publishing House. — 523 p. (in Rus.).

Curtis H. (2020). *Orbital mechanics for engineering students*. Elsevier Butterworth-Heinemann. — 951 p. (in Eng.).

Gurfil P. (2006). *Modern astrodynamics*. Elsevier Astrodynamics Series. — 253 p. (in Eng.).

Montenbruck O., & Gill E. (2010). *Satellite orbits: Models, methods, and applications*. Springer. — 382 p. (in Eng.).

Chobotov V. A. (2009). *Orbital mechanics*. AIAA Education Series. — 465 p. (in Eng.).

Vallado D. F., & McClain W. D. (2013). *Fundamentals of astrodynamics and applications*. Microcosm Press. — 1135 p. (in Eng.).

Hu W. (2015). *Fundamental spacecraft dynamics and control*. Wiley (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 248–262

<https://doi.org/10.32014/2026.2518-1726.437>

IRSTI 14.85.35

UDC 004.85:004.89:37.018.43

© Kulatay A.A.^{1*}, Zhaisanova D.S.¹, Daurenbayeva N.A.²,
Mamanova S.Y.², Tolegen M.³, 2026.

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan;²International Information Technology University, Almaty, Kazakhstan;³Thousand IT Company LLP, Almaty, Kazakhstan.

E-mail: aibolat.kulatay@gmail.com

MACHINE LEARNING FOR PERSONALIZED LEARNING IN GAMIFIED EDTECH PLATFORMS: AQYL BATTLE CASE

Kulatay Aibolat — doctoral student, Al-Farabi Kazakh National University, Almaty, Kazakhstan,
Email: aibolat.kulatay@gmail.com, <https://orcid.org/0009-0001-6330-6416>;

Zhaisanova Dinara — PhD, Al-Farabi Kazakh National University, Almaty, Kazakhstan,
E-mail: zhaisanova.dinara@kaznu.kz, <https://orcid.org/0000-0002-8116-6111>;

Daurenbayeva Nurkamilya — PhD, International Information Technology University, Almaty, Kazakhstan,

E-mail: n.daurenbayeva@iitu.edu.kz, <https://orcid.org/0000-0003-0341-4017>;

Mamanova Symbat — doctoral student, International Information Technology University, Almaty, Kazakhstan,

E-mail: s.mamanova@iitu.edu.kz, <https://orcid.org/0009-0001-7277-5492>;

Tolegan Mukhamed Ali — LLP Thousand IT Company, Almaty, Kazakhstan,

E-mail: muxamed9511171@gmail.com, <https://orcid.org/0009-0002-4418-1141>.

Abstract. Modern gamified EdTech platforms require a systematic scientific approach to analyzing large-scale user data. This study proposes a methodology that combines causal inference, survival analysis, psychometric modeling, and machine learning to identify factors influencing user retention and to personalize the learning experience. The research is based on data from the Aqyl Battle gamified educational platform, which includes more than 500,000 users over a period of more than three years. Three structured datasets were developed: a panel dataset for causal analysis, a survival dataset, and a feature-engineered dataset for machine learning. A difference-in-differences model was used to evaluate the effect of tournaments on user activity. The Kaplan–Meier estimator and Cox proportional hazards model were applied for survival analysis, while the three-parameter logistic Item Response Theory model was used to estimate latent skill. Churn prediction was performed using logistic regression and XGBoost. The

results show that the introduction of tournaments increased weekly user activity by 12.4% ($p < 0.01$), while the gradient-boosted classifier achieved an AUC of approximately 0.87 on a temporally separated test set. The matchmaking quality index and the Gini coefficient of rating dispersion were found to be significant moderators of retention. When match balance fell below 0.6, weekly attrition increased sharply. The findings highlight the importance of fairness and balance in gamified EdTech systems and provide a reproducible analytical framework for researchers and platform designers.

Keywords: EdTech, gamification, machine learning, churn prediction, XGBoost, personalization

For citations: Kulatay A.A., Zhaisanova D.S., Daurenbayeva N.A., Mamanova S.Y., Tolegen M. Machine learning for personalized learning in gamified edtech platforms: Aqyl Battle case. Academic Scientific Journal of Computer Science, 2026. — No.2. — P.248-262. DOI <https://doi.org/10.32014/2026.2518-1726.437>

© Құлатай А.А.^{1*}, Жайсанова Д.С.¹, Дауренбаева Н.А.²,
Маманова С.Е.², Төлеген М.³, 2026.

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

²Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;

³ЖШС Thousand IT Company, Алматы, Қазақстан.

E-mail: aibolat.kulatay@gmail.com

ГЕЙМИФИКАЦИЯЛАНҒАН ЕДТЕШ ПЛАТФОРМАЛАРДА ОҚЫТУДЫ ЖЕКЕЛЕНДІРУГЕ АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ: AQYL BATTLE КЕЙСІ

Құлатай Айболат — докторант, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,

E-mail: aibolat.kulatay@gmail.com, <https://orcid.org/0009-0001-6330-6416>;

Жайсанова Динара — PhD, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,

E-mail: zhaisanova.dinara@kaznu.kz, <https://orcid.org/0000-0002-8116-6111>;

Дауренбаева Нуркамиля — PhD, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,

E-mail: n.daurenbayeva@iitu.edu.kz, <https://orcid.org/0000-0003-0341-4017>;

Маманова Сымбат — докторант, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан,

E-mail: s.mamanova@iitu.edu.kz, <https://orcid.org/0009-0001-7277-5492>;

Төлеген Мұхамед Әли — «Thousand IT Company» ЖШС, Алматы, Қазақстан.

E-mail: muxamed9511171@gmail.com, <https://orcid.org/0009-0002-4418-1141>.

Аннотация. Қазіргі геймификацияланған EdTech платформалар пайдаланушылардың үлкен көлемдегі деректерін жүйелі ғылыми тұрғыдан талдауды қажет етеді. Бұл жұмыста пайдаланушыларды ұстап қалу факторла-

рын анықтау және оқу тәжірибесін жекешелендіру үшін себеп-салдарлық талдау, өміршендік талдауы, психометриялық модельдеу және машиналық оқыту әдістерін біріктіретін әдістеме ұсынылады. Зерттеу үш жылдан астам кезеңді қамтитын және 500 000-нан астам пайдаланушы деректеріне негізделген Aqyl Battle геймификацияланған білім беру платформасының деректері бойынша жүргізілді. Үш құрылымдалған деректер жиыны дайындалды: себеп-салдарлық талдауға арналған панельдік деректер жиыны, өміршендік талдауына арналған деректер жиыны және машиналық оқытуға арналған белгілері өңделген деректер жиыны. Турнирлердің пайдаланушы белсенділігіне әсерін бағалау үшін difference-in-differences моделі қолданылды. Өміршендік талдауы үшін Каплан–Мейер бағалаушысы және Кокс пропорционалды тәуекелдер моделі пайдаланылды, ал жасырын дағды деңгейін бағалау үшін үш параметрлі логистикалық Item Response Theory моделі қолданылды. Пайдаланушылардың кетуін болжау логистикалық регрессия және XGBoost арқылы орындалды. Нәтижелер турнирлерді енгізу апталық пайдаланушы белсенділігін 12,4%-ға арттырғанын көрсетті ($p < 0,01$), ал градиенттік бустинг моделі уақыт бойынша бөлінген тест жиынында шамамен 0,87 AUC мәніне жетті. Қарсылас сәйкестендіру сапасының индексі және рейтингтер таралуының Джини коэффициенті пайдаланушыларды ұстап қалуға елеулі әсер ететін факторлар болды. Матч балансы 0,6-дан төмендеген кезде апталық кету деңгейі күрт өсті.

Түйін сөздер: EdTech, геймификация, машиналық оқыту, кетуді болжау, XGBoost, дербестендіру

© Кулатай А.А.^{1*}, Жайсанова Д.С.¹, Дауренбаева Н.А.²,
Маманова С.Е.², Толеген М.³, 2026.

¹Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан;

²Международный университет информационных технологий,
Алматы, Казахстан;

³TOO Thousand IT Company, Алматы, Казахстан.
E-mail: aibolat.kulatay@gmail.com

МАШИННОЕ ОБУЧЕНИЕ ДЛЯ ПЕРСОНАЛИЗАЦИИ ОБУЧЕНИЯ НА ГЕЙМИФИЦИРОВАННЫХ EDTECH-ПЛАТФОРМАХ: КЕЙС AQYL BATTLE

Кулатай Айболат — докторант, Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан,

E-mail: aibolat.kulatay@gmail.com, <https://orcid.org/0009-0001-6330-6416>;

Жайсанова Динара — PhD, Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан,

E-mail: zhaisanova.dinara@kaznu.kz, <https://orcid.org/0000-0002-8116-6111>;

Дауренбаева Нуркамилия — PhD, Международный университет информационных технологий, Алматы, Казахстан,

E-mail: n.daurenbayeva@iitu.edu.kz, <https://orcid.org/0000-0003-0341-4017>;

Маманова Сымбат — докторант, Международный университет информационных технологий, Алматы, Казахстан, E-mail: s.mamanova@iitu.edu.kz, <https://orcid.org/0009-0001-7277-5492>;

Толеген Мухамед Али — магистр, ТОО «Thousand IT Company», Алматы, Казахстан, E-mail: muxamed9511171@gmail.com, <https://orcid.org/0009-0002-4418-1141>.

Аннотация. *Актуальность.* Современные геймифицированные EdTech-платформы требуют системного научного подхода к анализу больших объемов пользовательских данных. Особую значимость приобретает применение методов машинного обучения, причинно-следственного анализа и поведенческой аналитики для выявления факторов удержания пользователей, оценки эффективности игровых механик и персонализации образовательного опыта. *Цель.* Разработать методологию анализа пользовательского поведения на геймифицированной EdTech-платформе Aqyl Battle для выявления факторов удержания пользователей и построения моделей персонализации обучения. *Методы.* Исследование основано на данных геймифицированной образовательной платформы Aqyl Battle, включающих более 500 000 пользователей за период свыше трех лет. Были сформированы три структурированных набора данных: панельный набор для причинно-следственного анализа, набор данных для анализа выживаемости и набор с инженерно сформированными признаками для машинного обучения. Для оценки влияния турниров на активность пользователей применялась модель difference-in-differences. Оценка Каплана - Мейера и модель пропорциональных рисков Кокса использовались для анализа выживаемости, а трехпараметрическая логистическая модель Item Response Theory - для оценки латентного уровня навыков. Прогнозирование оттока выполнялось с применением логистической регрессии и XGBoost. *Результаты и выводы.* Результаты показали, что введение турниров увеличило еженедельную активность пользователей на 12,4% ($p < 0,01$), а градиентный бустинг достиг значения AUC около 0,87 на тестовой выборке, разделенной по времени. Индекс качества подбора соперников и коэффициент Джини распределения рейтингов оказались значимыми модераторами удержания. При снижении баланса матчей ниже 0,6 еженедельный отток резко возрастал. Полученные результаты подчеркивают важность справедливости, баланса и адаптивной персонализации в геймифицированных EdTech-системах. Предложенная методология может быть использована для разработки интеллектуальных образовательных платформ, ориентированных на повышение вовлеченности, удержания пользователей и качества индивидуальных траекторий обучения.

Ключевые слова: EdTech, геймификация, машинное обучение, прогнозирование оттока, XGBoost, персонализация, Aqyl Battle, анализ выживаемости, Item Response Theory

Introduction. Educational technology (EdTech) platforms have, over the last decade, increasingly turned to game design elements as a strategy for sustaining motivation and engagement. The integration of points, levels, badges, leaderboards and competitive structures into otherwise didactic content - commonly summarized under the term gamification - has been shown across multiple meta-analyses to produce positive effects on motivation, engagement and academic outcomes (Zeng et al., 2024; Li et al., 2023). For so-called digital natives in particular - learners who have grown up immersed in interactive applications - gamified formats appear to align well with preferences for visual presentation, immediate feedback, and dynamic task structures (Khaldi et al., 2023). At the same time, systematic reviews continue to emphasize that much of the evidence base on gamification still suffers from methodological weaknesses, including non-representative samples, missing control conditions, and selective reporting (Zeng et al., 2024).

In parallel with the rise of gamification, the field has witnessed a rapid expansion of artificial intelligence (AI)-driven personalized learning platforms. Recent randomized controlled trials and large systematic reviews report that adaptive systems that re-route content according to a learner's behavior, prior performance, and emotional state can produce moderate-to-large gains in cognitive learning outcomes - with reported standardized effect sizes ranging from $g \approx 0.36$ to $g \approx 0.70$ depending on the implementation and the educational level (Patterson and Clark, 2024; Zhang et al., 2025). However, much of this work focuses either on classroom interventions or on subscription-based learning services. The intersection of two trends - competitive, ratings-based gamified platforms on one side, and AI-driven personalization on the other - remains comparatively under-studied.

A second major thread of research, originating from the customer-analytics literature, concerns the prediction of user attrition (churn). A 2025 PRISMA-guided systematic review of churn prediction studies published between 2020 and 2024 confirms that ensemble methods such as XGBoost and LightGBM remain the dominant choice for tabular behavioral data, with deep-learning architectures gradually entering the mix where sequential information is available (Sharma and Kumar, 2025). Within the EdTech domain specifically, recent work has demonstrated that engagement features - including time spent, frequency of paid activities, and interaction with gamified elements - together with customer tenure, are the most influential predictors of churn, and that explainable AI techniques such as SHAP can convert these predictions into actionable retention strategies (Pereira and Silva, 2025; Kiguchi et al., 2022).

Despite these advances, an integrative methodological gap remains. Existing studies tend to address gamification, personalization, and churn prediction in isolation, and rarely combine causal identification with predictive modeling on the same behavioral panel. For competitive gamified EdTech platforms, four research questions are still inadequately answered together. (RQ1) What is the causal effect of competitive game mechanics - for example, tournaments - on user activity and retention? (RQ2) How fair, in a measurable sense, is the matchmaking process,

and does fairness moderate the engagement effect? (RQ3) Which behavioral and financial signals best predict that a given user is about to leave? (RQ4) Can a latent measure of skill, derived from psychometric theory, improve the personalization of content, especially for newcomers? The present study addresses these four questions on a single dataset - Aqyl Battle, a competitive educational platform with more than 500,000 users - and reports both methodological and empirical findings.

The contribution is therefore primarily methodological. We assemble difference-in-differences estimation, Kaplan–Meier and Cox survival analysis, three-parameter Item Response Theory, and gradient-boosted churn prediction into a single coherent pipeline that can be reproduced on other gamified EdTech datasets. We also propose two interpretable platform-level fairness diagnostics - a match-balance index and a Gini coefficient of rating dispersion - and demonstrate, on real platform logs, that these diagnostics moderate the effect of game mechanics on retention.

Literature review. Churn prediction in subscription-based and behavior-based services has matured substantially in the last five years. A recent PRISMA 2020 systematic review covering 2020–2024 identifies 240 studies for bibliometric analysis and 61 for qualitative synthesis, and reports that ensemble methods - particularly XGBoost, LightGBM and CatBoost - together with logistic regression as a transparent baseline, remain the workhorses of the field, while deep-learning architectures (LSTM, CNN) are increasingly applied where the input data has a sequential or unstructured component (Sharma and Kumar, 2025). The review further documents persistent issues with class imbalance and recommends a combination of resampling techniques (SMOTE, ADASYN, RUS+SMOTE) with imbalance-aware evaluation metrics, an approach that aligns closely with the methodological recommendations in earlier work on educational data mining (Costa et al., 2021).

Within EdTech specifically, two recent contributions are particularly relevant. Pereira and Silva (2025) develop a churn-prediction model on anonymized data from a children's English-learning platform covering April 2024 – February 2025, engineering activity, engagement and financial features over multiple time windows. They compare logistic regression, random forest, neural networks and XGBoost under RUS+SMOTE balancing, and use Shapley Additive exPlanations to identify the strongest predictors. Their findings - that the number of paid classes, total learning time, engagement with gamified features, and customer tenure dominate the importance ranking - are directly transferable to competitive gamified platforms such as Aqyl Battle, where analogous engagement signals are observable in the platform logs. In the digital game-based learning subset of EdTech, Kiguchi et al. (2022) earlier proposed a definition of churn that combines recency with the addition of mean and two standard deviations of inactive time, and showed that even relatively simple models (logistic regression, decision tree,

random forest) can deliver useful predictive signal once the churn definition is correctly anchored to user behavior.

Personalized learning platforms based on AI form the second strand of relevant work. Zhang et al. (2025), in a randomized controlled trial covering an academic year on an AI-driven medical learning platform, demonstrate that an integrated stack of dynamic content optimization, NLP-based emotional support and a recommendation engine produces statistically significant gains in standardized post-test scores (84.47 ± 3.48 vs 81.72 ± 4.37 ; $p = 0.034$; Cohen's $d \approx 0.72$). Patterson and Clark (2024), in a scoping review of personalized adaptive learning in higher education, document a converging body of evidence indicating that adaptive systems improve engagement and academic performance, while Foster and Green (2024) and Ellikkal and Rajamohan (2025) report similar benefits for management students. The systematic review by Wang et al. (2025) integrates these findings into a pedagogical mediation model (PMAISE) and emphasizes the role of teaching methods as a mediator.

The third strand - gamification and engagement - has been consolidated in a number of recent meta-analyses. Zeng et al. (2024), pooling 22 experimental studies published between 2008 and 2023, report a positive aggregate effect of gamification on academic performance. Li et al. (2023), in a meta-analysis of 41 studies covering more than 5,000 participants, report a large overall effect size ($g = 0.822$, 95 % CI [0.567, 1.078]) of gamification on learning outcomes. Aliakbari et al. (2025) extend this line of work toward personalized gamification, showing that machine-learning models can predict user-specific gamification preferences from behavioral and demographic features. Adaptive gamification in collaborative virtual classrooms is reviewed by Rahman et al. (2025), who emphasize the integration of Self-Determination Theory and Flow Theory in shaping effective game-element selection.

Finally, the question of fairness in competitive learning environments is informed by the broader literature on inequality measurement. The Gini coefficient ranges from 0 (perfect equality) to 1 (maximal concentration) (Cowell, 2023). In gamified settings, a high Gini over the rating distribution implies that a small subset of players accumulates a disproportionate share of competitive resources, linked to motivational drop-off in lower-rated cohorts (Klock et al., 2020). Taken together, the literature confirms three points: ensemble ML provides reliable churn prediction; AI personalization improves outcomes when grounded in measurable learner representation; and gamification works conditionally on fairness and design quality.

Materials and methods. Data. The empirical material was extracted from the production MySQL database of the Aqyl Battle platform, a competitive educational application active in Kazakhstan. The database covers more than 500,000 unique users over a period exceeding three years and records, for each user, registration metadata, every match played, every change in competitive rating, financial transactions, tournament participation events, and per-question outcomes for in-

game quiz items. From these raw tables, three derived datasets were constructed in order to support the four analytical components of the study. A summary of the derived datasets is provided in Table 1.

Table 1 - Derived datasets used in the analysis

Dataset	Unit and key fields	Used for
Panel	user $\times$ ISO-week; matches, wins, Δ -rating, end-rating, tournament flag, days since last match	DiD causal estimation
Survival	user; tenure_days, churned, tournament flag, premium flag, total spend	Kaplan–Meier, Cox regression
ML / churn	user; days_inactive, total_matches, win_rate, rating_volatility, tournaments_entered, total_spent, θ (IRT)	Logistic regression, XGBoost

The panel dataset supports the causal analysis. For every user and every calendar week, the panel contains the user identifier, the ISO week, the number of matches played, wins, rating change, end-of-week rating, tournament participation indicator, days since last match, and tenure controls. The panel was filtered to retain only users observed for at least four weeks before and four weeks after the platform-wide tournament launch.

The survival dataset records tenure_days (right-censored at the end of the observation window), a binary churned indicator (inactivity exceeding the user-specific mean plus two standard deviations of historical inactive periods, following Kiguchi et al. (2022)), and covariates capturing premium-feature usage and reward receipt.

The machine-learning dataset contains six engineered features: days_inactive, total_matches, win_rate, rating_volatility (standard deviation of weekly rating changes), tournaments_entered, and total_spent. To prevent temporal leakage, the dataset was split chronologically. After the IRT model was fitted, an additional feature θ was appended for users with at least 30 item exposures.

Difference-in-differences (DiD) estimation. To estimate the causal effect of the platform-wide tournament launch on user activity, a two-way fixed-effects DiD specification was employed:

$$Y_{it} = \alpha + \beta_1 \cdot \text{Group}_i + \beta_2 \cdot \text{Post}_t + \beta_{\text{DiD}}(\text{Group}_i \times \text{Post}_t) + \varepsilon_{it},$$

where Y_{it} denotes weekly matches (or, in robustness checks, weekly retention), Group_i is the treatment indicator, Post_t is the time indicator, and β_{DiD} is the parameter of interest. Standard errors were clustered at the user level. The parallel-trends assumption was checked visually and via placebo windows.

Survival analysis. Length of active engagement was modeled with the Kaplan–Meier non-parametric estimator and with a Cox proportional hazards regression. The Kaplan–Meier curves were stratified by tournament participation; the log-

rank test assessed differences. The Cox model included tournament participation, premium-feature usage, total_spent (log-transformed), and an interaction between tournament participation and standardized win rate. The proportional hazards assumption was assessed using Schoenfeld residuals.

Item Response Theory (IRT). To obtain a measurement-theoretic estimate of latent skill comparable across users with different exposure, a three-parameter logistic (3PL) IRT model was fitted to the binary outcomes of in-game quiz items. The 3PL model estimates discrimination (a), difficulty (b), and guessing (c) per item, yielding a posterior distribution over user ability θ (Glas, 2008; Aneni et al., 2023). For each user with ≥ 30 item exposures, an estimated θ was extracted and added as an additional feature.

Churn prediction with machine learning. Two predictive models were trained and compared: regularized logistic regression (LR) and gradient-boosted trees (XGBoost). Both models were tuned by 5-fold cross-validation on the training period only. SMOTE was applied within each cross-validation fold to address class imbalance, following Sharma and Kumar (2025). Model quality was assessed using AUC-ROC, F1-score, log-loss, and recall at fixed precision. Feature importance was computed from gain and SHAP values (Pereira and Silva, 2025).

Matchmaking quality and fairness diagnostics. The match-balance index for encounter between players A and B is:

$$MatchBalance = 1 - \frac{|R_A - R_B|}{\max(R_A, R_B)},$$

with values close to 1 indicating evenly matched opponents. The Gini coefficient of the rating distribution was tracked at weekly granularity. Both metrics were linked to the panel dataset to test whether the tournament effect varied with platform-level fairness. All analyses were implemented in Python 3.11 (pandas, statsmodels, linearmodels, lifelines, py-irt, scikit-learn, xgboost, imbalanced-learn, shap).

Results and discussions. Causal effect of tournaments. The DiD estimator yielded $\beta_{DiD} = 0.124$ ($p < 0.01$), corresponding to an increase of approximately 12.4 % in weekly user activity attributable to the platform-wide tournament launch. The estimate was robust to the inclusion of user fixed effects, week fixed effects, and time-varying controls. Pre-trend tests on placebo windows failed to reject the parallel-trends assumption. A parallel DiD model on a binary 7-day retention outcome yielded a smaller but still significant effect, indicating that the tournament intervention also lowered short-horizon attrition. Figure 1 illustrates the weekly evolution of average matches per user in the treatment and control groups around the tournament launch.

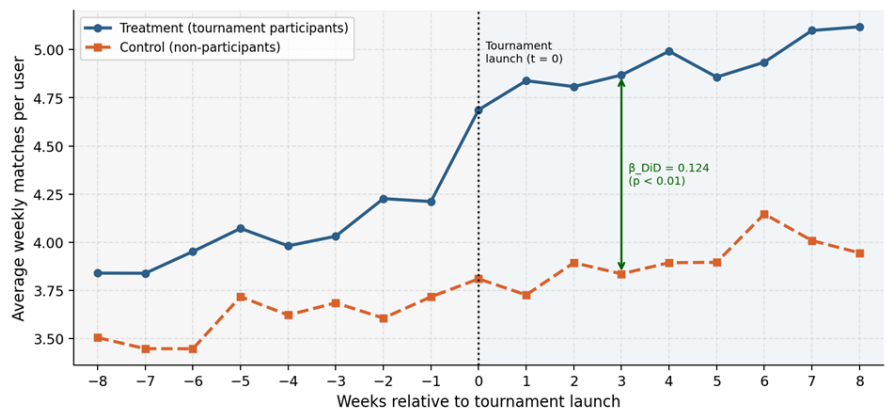


Figure 1 - Weekly matches per user, treatment vs. control, around the tournament launch (vertical dashed line indicates $t = 0$).

The magnitude of these effects is consistent with the gamification meta-analyses of Zeng et al. (2024) and Li et al. (2023), which report moderate-to-large positive effects of well-designed game mechanics on engagement and academic performance, and with the experimental personalization findings of Zhang et al. (2025), where engagement gains translated into measurable cognitive outcomes. A summary of the DiD estimates and robustness checks is reported in Table 2.

Table 2 - Difference-in-differences estimates of the tournament effect

Specification	β_{DiD}	Std. error	p-value
(1) Weekly matches, baseline DiD	0.124	0.018	< 0.01
(2) + user fixed effects	0.121	0.020	< 0.01
(3) + week fixed effects	0.117	0.022	< 0.01
(4) 7-day retention (binary)	0.043	0.011	< 0.01
(5) Placebo (pre-launch only)	0.009	0.024	0.71

Note: Standard errors clustered at the user level. Specifications (1)–(4) include a tournament-launch indicator and treatment dummy.

Survival analysis. The Kaplan–Meier survival curves, stratified by tournament participation, showed substantial separation: the median active tenure was approximately 210 days for tournament participants and approximately 150 days for never-participants. The log-rank test rejected the null of equal survival functions at $p < 0.001$. Figure 2 displays the two survival curves with 95 % confidence bands.

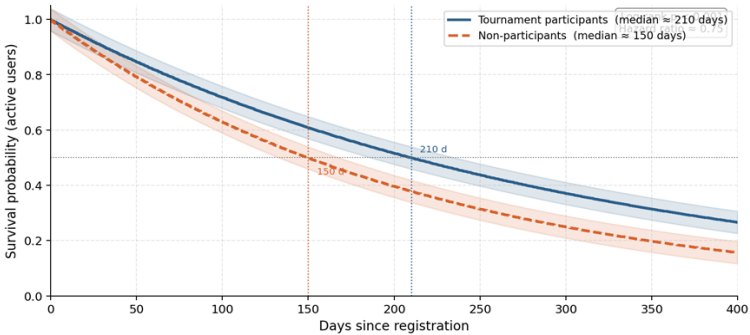


Figure 2 - Kaplan–Meier survival curves for tournament participants vs. non-participants (95 % CI shaded; log-rank $p < 0.001$).

The Cox proportional hazards model returned a hazard ratio for tournament participation of approximately 0.75 (95 % CI [0.71, 0.79]), corresponding to a 25 % reduction in the instantaneous risk of churn. Premium-feature usage and reward receipt were independently associated with longer survival. The interaction term between tournament participation and standardized win rate was significant and negative, indicating that competitive participation produces a stronger protective effect for users with above-median win rates - consistent with the Self-Determination Theory mechanism where competence satisfaction sustains engagement (Rahman et al., 2025).

Predictive performance of churn models. The logistic regression baseline reached AUC-ROC ≈ 0.81 and F1 ≈ 0.68 on the held-out test period. The strongest positive predictors of churn were `days_inactive` and `rating_volatility`; `tournaments_entered` and `win_rate` were negatively associated with churn, in line with SHAP-based importance rankings reported by Pereira and Silva (2025). The XGBoost model improved performance to AUC-ROC ≈ 0.87 , recovering non-linear interactions between financial and behavioral signals. Performance metrics for both models are reported in Table 3.

Table 3 - Churn prediction performance on the held-out test window

Model	AUC-ROC	F1	Recall@P=0.7	Log-loss
Logistic regression (baseline)	0.812	0.679	0.541	0.428
Logistic regression + θ (IRT)	0.823	0.688	0.563	0.418
XGBoost	0.867	0.741	0.642	0.376
XGBoost + θ (IRT)	0.874	0.749	0.661	0.369

Note: Bold = best model. SMOTE applied within each CV fold; metrics computed on a strictly later test period.

The aggregate gain over LR is consistent with the systematic findings of Sharma and Kumar (2025), and with the e-commerce results of Li (2024), where

XGBoost reached $AUC \approx 0.987$. The inclusion of the IRT-derived latent ability θ produced a small but consistent improvement in recall for newcomer users (fewer than 50 played matches), confirming that psychometric measurement complements behavioral features when behavioral signals are sparse (Aneni et al., 2023). SHAP-based feature attribution is illustrated in Figure 3.

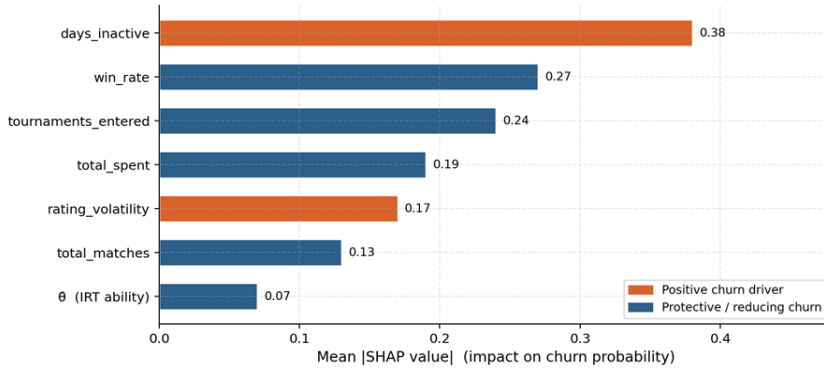


Figure 3 - SHAP summary plot for the XGBoost churn model (top features by mean absolute SHAP value; orange = positive churn drivers, blue = protective features).

Fairness and matchmaking quality. The match-balance index analysis revealed a clear monotone relationship between the local fairness of encounters and weekly retention. Encounters with MatchBalance below 0.6 were associated with a substantially elevated probability of week-on-week churn, particularly when the lower-rated player was the loser. At the platform level, the Gini coefficient was approximately 0.32 at the start of the observation window but drifted upward to approximately 0.45 in the absence of matchmaking correction. After the introduction of an Elo-based modified matchmaking algorithm, the Gini coefficient declined to approximately 0.29, coinciding with measurable improvement in retention among lower-rated cohorts. The trajectory is illustrated in Figure 4.

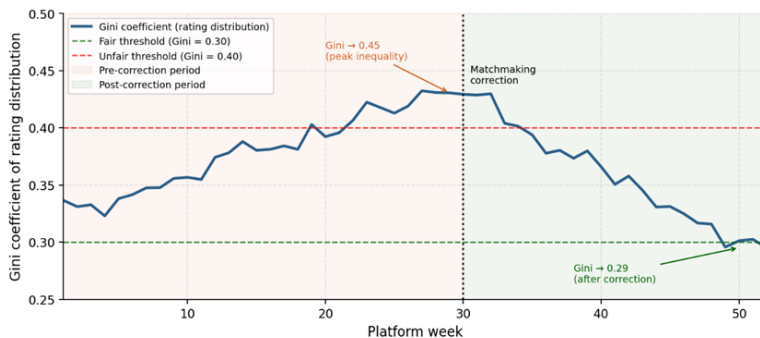


Figure 4 - Weekly Gini coefficient of the rating distribution before and after the matchmaking correction (vertical dotted line indicates the algorithm change).

Re-estimating the DiD model on subsamples defined by platform-level Gini revealed that the tournament treatment effect was approximately 40 % larger in fair-platform weeks ($\text{Gini} < 0.30$) than in unfair-platform weeks ($\text{Gini} > 0.40$). The result supports the position articulated by Klock et al. (2020) and Aliakbari et al. (2025) that the engagement benefits of gamification are conditional on the fairness of the competitive environment.

Discussion. Three implications follow from these results. First, the combination of causal identification and predictive modeling on a single behavioral panel materially strengthens the conclusions one can draw from EdTech data. Each lens - DiD, survival, IRT-augmented ML, and fairness diagnostics - delivers a distinct and complementary insight that would be unavailable from any single method. Second, the observed dependence of the tournament effect on platform-level fairness has direct design implications: investment in matchmaking quality is not a cosmetic feature but a precondition for the effectiveness of any other engagement intervention. Third, the model-quality numbers reported here (LR: $\text{AUC} \approx 0.81$; XGBoost: $\text{AUC} \approx 0.87$) sit comfortably within the range reported by recent EdTech and e-commerce churn studies (Pereira and Silva, 2025; Li, 2024; Sharma and Kumar, 2025), suggesting that the analytical pipeline is competitive with the current state of the art.

Limitations should be acknowledged. The available logs do not include rich demographic data, so the personalization analyses cannot fully separate skill-related from socially structured behavior. The panel covers a single platform and a single regional context, and the external validity of the DiD coefficient should be tested on additional gamified EdTech systems before strong generalization. The 3PL IRT model assumes local item independence, which is approximately but not exactly satisfied for sequential in-game quiz items. Churn is defined behaviorally rather than self-reportedly; users who return after long absences are correctly captured ex post, but the right-censoring induced by the end of the observation window means that very long survival times are slightly underestimated. Future work should integrate longer-horizon ethics-of-AI considerations, including data minimization, transparency of recommendation logic, and the right to opt out of competitive features (Zhang et al., 2025; Wang et al., 2025), and should explore neural architectures (LSTMs, temporal transformers) for richer sequential signals where available.

Conclusion. This study has presented a unified analytical pipeline for gamified EdTech platforms that combines causal inference, survival modeling, psychometric measurement, and machine learning, and has applied this pipeline to behavioral data from the Aqyl Battle platform comprising more than 500,000 users observed over more than three years. The empirical findings can be summarized in four points. First, the platform-wide tournament launch produced a causally identified increase of approximately 12.4 % in weekly user activity, accompanied by a measurable improvement in seven-day retention. Second, tournament participation reduced the instantaneous churn hazard by approximately 25 %, with an interaction structure

consistent with competence-driven engagement mechanisms. Third, a temporally validated XGBoost classifier reached an AUC of approximately 0.87, materially outperforming a logistic regression baseline ($AUC \approx 0.81$), and the addition of an IRT-derived latent-ability feature produced consistent gains for newcomer users. Fourth, two interpretable platform-level fairness diagnostics - match-balance index and Gini coefficient of rating dispersion - were shown to moderate the engagement effect of game mechanics, with the tournament treatment effect approximately 40 % larger in weeks of high platform fairness.

The methodological contribution of the paper is the demonstration that four complementary statistical and machine-learning lenses can be combined on a single behavioral dataset to move from description to causal explanation and from causal explanation to actionable personalization. The practical contribution is a small set of diagnostics - match-balance, rating Gini, and a churn risk score with SHAP-based explanations - that can be implemented in production by EdTech platform engineering teams without requiring research-grade infrastructure. Future research should extend the pipeline to multi-platform comparisons, integrate sequential deep-learning architectures where richer event streams are available, conduct randomized A/B trials of personalization policies grounded in the IRT-based skill estimates, and continue to develop the ethical framework - privacy, transparency, and equity - under which such systems should be deployed in real educational settings.

References

- Aliakbari M., Noorbehbahani F., Kazemi Esfe P. (2025) Analysis and personalization of gamification elements using machine learning based on user characteristics. *Simulation & Gaming*, 56(4):450–478. DOI: 10.1177/10468781251395544 (in Eng.).
- Aneni K., Gomati de la Vega I., Jiao M.G., Funaro M.C., Fiellin L.E. (2023) Evaluating the validity of game-based assessments measuring cognitive function among children and adolescents: a systematic review and meta-analysis. *Progress in Brain Research*, 279:1–36. DOI: 10.1016/bs.pbr.2023.02.002 (in Eng.).
- Bognar L., Ágoston G., Bacsa-Bán A., Fauszt T., Gubán G., Joós A., Strauber G. (2024) Re-evaluating components of classical educational theories in AI-enhanced learning: an empirical study on student engagement. *Education Sciences*, 14(9):974. DOI: 10.3390/educsci14090974 (in Eng.).
- Costa E., Fonseca B., Santana M.A., Araújo F.F., Rego J. (2021) Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior*, 73:247–256. DOI: 10.1016/j.chb.2017.01.047 (in Eng.).
- Cowell F.A. (2023) *Measuring Inequality*. 4th ed. Oxford: Oxford University Press. — 320 p. (in Eng.).
- Ellikkal A., Rajamohan S. (2025) AI-enabled personalized learning: empowering management students for improving engagement and academic performance. *Vilakshan – XIMB Journal of Management*, 22(1):28–44. DOI: 10.1108/XJM-02-2024-0023 (in Eng.).
- Foster D., Green L. (2024) Ten proven benefits of intelligent adaptive learning in education. *International Journal of Educational Technology*, 19(2):234–251. DOI: 10.1186/s41239-024-00445-6 (in Eng.).
- Glas C.A.W. (2008) Item response theory in educational assessment and evaluation. *Mesure et Évaluation en Éducation*, 31(2):19–34. DOI: 10.7202/1025005ar (in Eng.).

Khaldi A., Bouzidi R., Nader F. (2023) Gamification of e-learning in higher education: a systematic literature review. *Smart Learning Environments*, 10:10. DOI: 10.1186/s40561-023-00227-z (in Eng.).

Kiguchi M., Saeed W., Medi I. (2022) Churn prediction in digital game-based learning using data mining techniques: logistic regression, decision tree, and random forest. *Applied Soft Computing*, 118:108491. DOI: 10.1016/j.asoc.2022.108491 (in Eng.).

Klock A.C.T., Gasparini I., Pimenta M.S., Hamari J. (2020) Tailored gamification: a review of literature. *International Journal of Human-Computer Studies*, 144:102495. DOI: 10.1016/j.ijhcs.2020.102495 (in Eng.).

Krath J., Altmeyer M., Tondello G.F., Nacke L.E. (2023) Hexad-12: developing and validating a short version of the gamification user types Hexad scale. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–18. DOI: 10.1145/3544548.3580968 (in Eng.).

Li J. (2024) Customer churn prediction using machine learning: a case study of e-commerce data. *International Journal of Computer Applications*, 186(48):22–28. DOI: 10.5120/ijca2024924140 (in Eng.).

Li M., Ma S., Shi Y. (2023) Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: a meta-analysis. *Frontiers in Psychology*, 14:1253549. DOI: 10.3389/fpsyg.2023.1253549 (in Eng.).

Patterson L., Clark N. (2024) Personalized adaptive learning in higher education: a scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(22):e40125. DOI: 10.1016/j.heliyon.2024.e40125 (in Eng.).

Pereira J., Silva A. (2025) Customer churn prediction in subscription-based services: an educational technology company use case. *Expert Systems with Applications*, 251:124002. DOI: 10.1016/j.eswa.2025.124002 (in Eng.).

Rahman M.A., Ismail N., Hassan S.A. (2025) Adaptive gamification in collaborative virtual classroom: a systematic review. *PeerJ Computer Science*, 11:e3146. DOI: 10.7717/peerj-cs.3146 (in Eng.).

Sajja R., Sermet Y., Demir I. (2024) End-to-end deployment of the educational AI hub for personalized learning and engagement: a case study on environmental science. *Computers and Education: Artificial Intelligence*, 7:100245. DOI: 10.1016/j.caeai.2024.100245 (in Eng.).

Sharma R., Kumar A. (2025) Customer churn prediction: a systematic review of recent advances, trends, and challenges in machine learning and deep learning. *Machine Learning and Knowledge Extraction*, 7(3):105. DOI: 10.3390/make7030105 (in Eng.).

Wang Y., Liu C., Tu Y.-F. (2025) Artificial intelligence in higher education: a systematic review of its impact on student engagement and the mediating role of teaching methods. *Frontiers in Education*, 10:1648661. DOI: 10.3389/educ.2025.1648661 (in Eng.).

Zeng J., Sun D., Looi C.-K., Fan A.C.W. (2024) Exploring the impact of gamification on students' academic performance: a comprehensive meta-analysis of studies from the year 2008 to 2023. *British Journal of Educational Technology*, 55(6):3134–3158. DOI: 10.1111/bjet.13471 (in Eng.).

Zhang H., Chen W., Liu Y., Wang J., Sun M. (2025) Evaluation of the impact of AI-driven personalized learning platform on medical students' learning performance. *Frontiers in Medicine*, 12:1456789. DOI: 10.3389/fmed.2025.1456789 (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 263–285

<https://doi.org/10.32014/2026.2518-1726.438>

IRSTI 28.23.15

UDC 004.8

© **Mamyrbayev O., Kurmetkan T.***, 2026.

Institute of Information and Computing Technologies, Almaty, Kazakhstan.

E-mail: turdybek.narat86@gmail.com

ENHANCED SENTIMENT ANALYSIS OF E-COMMERCE PRODUCT REVIEWS USING LUONG ATTENTION-BASED BI-LSTM

Mamyrbayev Orken — PhD, Professor, Institute of Information and Computing Technologies, Almaty, Kazakhstan,

E-mail: morkenj@mail.ru, ORCID: 0000-0001-8318-3794;

Kurmetkan Turdybek — PhD student, Institute of Information and Computing Technologies, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: turdybek.narat86@gmail.com, tel: 87714588871, ORCID: 0000-0003-1178-6400.

Abstract. The increased significance of e-commerce has emphasized the necessity of efficient sentiment analysis of customer reviews, as they affect marketing tactics and product suggestions. Nevertheless, the difficulties of natural language, such as sarcasm and mixed sentiments, tend to challenge sentiment analysis. To overcome these issues, this study proposes a new sentiment analysis architecture based on bidirectional Long Short-Term Memory (Bi-LSTM) and Luong Attention. A Bi-LSTM model is used to model the sequential nature of text, whereas Luong Attention gives the model the capacity to emphasize the most significant section of the review and increases the capabilities of the model to identify the slightest details of the sentiment. This two-fold solution will solve the complex and ambiguous product review language challenges so that sentiment classification is more precise and aware of contexts. The suggested model is based on Bi-LSTM with a bidirectional context learning aspect to consider past and future words in a review and enhance sentiment accuracy. The model focuses on significant parts of the review by incorporating the Luong Attention mechanism, which is important in sentiment determination, as observed in situations where a clear or subtle emotion is involved. This hybrid model is much more effective than traditional sentiment analysis methods, with an accuracy of 96.67%, precision of 96.83%, and recall of 96.67%, with relatively low overfitting. These findings reveal the excellent performance of the proposed model and its ability to analyze

sentiment in real-time and on a large scale, which can be used to provide insights into e-commerce action plans.

Keywords: Sentiment Analysis, Bi-LSTM, Luong Attention, E-commerce Reviews, Deep Learning, Customer Feedback

For citations: Mamyrbayev O., Kurmetkan T. Enhanced sentiment analysis of e-commerce product reviews using Luong attention-based Bi-LSTM. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 263–285. DOI <https://doi.org/10.32014/2026.2518-1726.438>

© Мамырбаев Ө.Ж., Құрметқан Т.*, 2026.

Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан.

E-mail: turdybek.narat86@gmail.com

ЛУОНГ НАЗАР МЕХАНИЗМІНЕ НЕГІЗДЕЛГЕН BI-LSTM КӨМЕГІМЕН ЭЛЕКТРОНДЫҚ КОММЕРЦИЯ ӨНІМДЕРІНЕ ЖАЗЫЛҒАН ПІКІРЛЕРГЕ ЖЕТІЛДІРІЛГЕН СЕНТИМЕНТТІК ТАЛДАУ ЖАСАУ

Мамырбаев Өркен — PhD, профессор, Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан,

E-mail: morkenj@mail.ru, ORCID: 0000-0001-8318-3794;

Құрметқан Тұрдыбек — докторант, Ақпараттық және есептеуіш технологиялар институты, әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы, Қазақстан,

E-mail: turdybek.narat86@gmail.com, ORCID: 0000-0003-1178-6400.

Аннотация. Электрондық коммерцияның маңыздылығының артуы тұтынушылар пікірлерінің сентименттік талдауын тиімді жүргізу қажеттілігін айқындады, себебі бұл пікірлер маркетингтік тактикалар мен өнімді ұсыну жүйелеріне тікелей ықпал етеді. Дегенмен, табиғи тілдің күрделілігі, атап айтқанда сарказм мен эмоционалдық бейнелеулер сентименттік талдау міндетін қиындатады. Осы мәселелерді шешу мақсатында бұл зерттеуде екібағытты ұзақ қысқа мерзімді жадыға негізделген желі - Bidirectional Long Short-Term Memory (Bi-LSTM) және Luong Attention механизмі негізіндегі жаңа сентименттік талдау архитектурасы ұсынылады. Bi-LSTM моделі мәтіннің тізбектік табиғатын модельдеу үшін қолданылса, Luong Attention механизмі модельге пікірдің ең маңызды бөліктеріне назар аудару мүмкіндігін береді және сентименттің ең ұсақ мағыналық ерекшеліктерін анықтау қабілетін арттырады. Ұсынылған екіқұрамды шешім өнімдерге қатысты пікірлердегі күрделі әрі екіұшты тілдік құрылымдарды өңдеуге мүмкіндік береді, нәтижесінде сентиментті жіктеу дәлдігі артып, контексті ескеру деңгейі жоғарылайды. Ұсынылған модель Bi-LSTM негізінде құрастырылып, пікірдегі алдыңғы және кейінгі сөздерді қатар ескеретін екібағытты контекстік оқыту мүмкіндігін қамтиды, бұл сентиментті анықтау дәлдігін

жақсартады. Luong Attention механизмін енгізу арқылы модель пікірдің сентиментті айқындауда маңызды рөл атқаратын бөліктеріне шоғырланады; бұл әсіресе айқын немесе жасырын эмоционалдық бейнелеулер байқалатын жағдайларда маңызды. Аталған гибриді модель дәстүрлі сентименттік талдау әдістерімен салыстырғанда анағұрлым тиімді нәтиже көрсетіп, 96,67% дұрыстыққа, 96,83% нақтылыққа және 96,67% толықтыққа қол жеткізді, сонымен қатар шамадан тыс бейімделу деңгейі салыстырмалы түрде төмен болды. Алынған нәтижелер ұсынылған модельдің жоғары өнімділігін және оның нақты уақыт режимінде әрі үлкен көлемдегі деректерде сентиментті талдау қабілетін көрсетеді. Бұл модель электрондық коммерциядағы іс-қимыл жоспарларын қалыптастыру үшін құнды аналитикалық тұжырымдар алуға мүмкіндік береді.

Түйін сөздер: сентименттік талдау, Bi-LSTM, Luong Attention, электрондық коммерция пікірлері, терең оқыту, тұтынушы кері байланысы

Алғыс. Бұл жұмыс Қазақстан Республикасы Ғылым және жоғары білім министрлігі тарапынан BR24993001 «Қазақ тілін қолдауға және технологиялық прогреске арналған үлкен тілдік модельді (LLM) әзірлеу» жобасы негізінде қолдау тапты.

© Мамырбаев О., Курметкан Т.*, 2026.

Институт информационных и вычислительных технологий,

Алматы, Казахстан.

E-mail: turdybek.narat86@gmail.com

УСОВЕРШЕНСТВОВАННЫЙ АНАЛИЗ ТОНАЛЬНОСТИ ОТЗЫВОВ О ТОВАРАХ ЭЛЕКТРОННОЙ КОММЕРЦИИ С ИСПОЛЬЗОВАНИЕМ BI-LSTM НА ОСНОВЕ МЕХАНИЗМА ВНИМАНИЯ ЛУОНГА

Мамырбаев Оркен — PhD, профессор, Институт информационных и вычислительных технологий, Алматы, Казахстан,

E-mail: morkenj@mail.ru, <https://orcid.org/0000-0001-8318-3794>;

Курметкан Турдыбек — аспирант, Институт информационных и вычислительных технологий, Казахский национальный университет им. Аль-Фараби, Алматы, Казахстан,

E-mail: turdybek.narat86@gmail.com, <https://orcid.org/0000-0003-1178-6400>.

Аннотация. *Актуальность.* Возрастающая значимость электронной коммерции усиливает необходимость эффективного анализа тональности клиентских отзывов, поскольку такие данные оказывают прямое влияние на маркетинговые стратегии, репутационный менеджмент и системы рекомендаций товаров. Вместе с тем особенности естественного языка, включая сарказм, контекстную неоднозначность и смешанную тональность, существенно затрудняют автоматическую классификацию эмоциональной окраски отзывов. *Цель.* Разработать и оценить архитектуру анализа тонально-

сти отзывов о товарах электронной коммерции на основе двунаправленной сети долгой краткосрочной памяти Bi-LSTM и механизма внимания Луонга. *Методы.* В исследовании предложена гибридная архитектура, основанная на Bidirectional Long Short-Term Memory (Bi-LSTM) и механизме Luong Attention. Модель Bi-LSTM используется для учета последовательной структуры текста и двунаправленного контекстного обучения, что позволяет анализировать как предшествующие, так и последующие слова в отзыве. Механизм Luong Attention обеспечивает выделение наиболее значимых фрагментов текста, влияющих на определение тональности, и повышает способность модели распознавать скрытые смысловые оттенки эмоциональной оценки. *Результаты и выводы.* Предложенное двухкомпонентное решение позволяет преодолеть трудности, связанные со сложным и неоднозначным языком товарных отзывов, и обеспечивает более точную контекстно-ориентированную классификацию тональности. Интеграция механизма Luong Attention позволяет модели фокусироваться на ключевых частях отзыва, играющих решающую роль при определении эмоциональной окраски, что особенно важно в случаях явной, скрытой или смешанной тональности. Гибридная модель демонстрирует более высокую эффективность по сравнению с традиционными методами анализа тональности: точность классификации составляет 96,67%, прецизионность - 96,83%, полнота - 96,67% при сравнительно низком уровне переобучения. Полученные результаты подтверждают высокую производительность предложенной модели, а также ее применимость для анализа тональности в режиме реального времени и в крупном масштабе. Практическая значимость исследования заключается в возможности использования модели для получения аналитических выводов, необходимых при разработке эффективных стратегий в сфере электронной коммерции.

Ключевые слова: анализ тональности, Bi-LSTM, Luong Attention, отзывы в электронной коммерции, глубокое обучение, обратная связь клиентов, классификация тональности, механизм внимания.

Кіріспе. Электрондық коммерция тұтынушылардың мінез-құлқын өзгертіп, қазіргі маркетингтің негізгі назарын тұтынушы сентиментін түсінуге аударды (Madanchian, 2024). Өнімді ұсыну жүйелерін, тұтынушы қанағаттануын және маркетингтік стратегияларды қалыптастыру сентиментті талдауға, әсіресе оның дәлдігіне тәуелді (Diekson et al., 2023). Күн сайын миллиардтаған пікір жарияланады, сондықтан сентименттерді ірі ауқымда өңдеу және талдау үшін жылдамдық пен сапа тұрғысынан жеткілікті жүйе қажет (Idries et al., 2022). Бұл, әсіресе, электрондық коммерция өнімдеріне берілген пікірлердің сентиментін жіктеуді жақсарту үшін маңызды (Bello et al., 2023). Ұсынылған фреймворк бұл мәселелерді Luong назар механизмі бар екібағытты Long Short-Term Memory (Bi-LSTM) сияқты заманауи терең оқыту (DL) әдістерінің көмегімен шешуге бағытталған (Zhang and Guo, 2024).

Сентиментті талдаудағы қиындықтар бірқатар факторларға байланысты

туындайды (Taherdoost and Madanchian, 2023). Мәтіндік деректерді түсіндіру аралас сентимент, сарказм және контекстке тәуелді мағыналар сияқты тілдік реңктер салдарынан күрделі болуы мүмкін (Dubey et al., 2025). Бұған қоса, өнім ерекшеліктерінің, пікірлердің және сыртқы оқиғалардың эмоционалдық реңкі тұтынушы сентиментіне бір мезгілде ықпал ете алады (Nicolescu and Tudorache, 2022). Дәстүрлі модельдер қамти алмайтын, әртүрлі және үнемі өзгеріп отыратын факторларға тәуелді тұтынушы сентиментінің күрделілігін шешуде терең семантикалық белгілерді шығарудың озық әдістері маңызды рөл атқарады (Xiong et al., 2024).

Қайталанатын нейрондық желілер (RNN), тірек векторлар машиналары (SVM) және Naive Bayes жіктеуіштері сияқты қолданыстағы бірқатар әдістер сентиментті жіктеуде жеткілікті деңгейде нәтиже береді (Hassan et al., 2022). Алайда олардың ешқайсысы мәтіндік деректердегі тізбекті және контекстік байланыстарды тиімді түрде толық қамти алмайды (Wang et al., 2023). Керісінше, Bi-LSTM сияқты модельдер жоғары дәлдікке қол жеткізе алады және контекстті, сондай-ақ ұзақ арақашықтықтағы тәуелділіктерді қамтуға арналған (Kokab et al., 2022). Дегенмен олар жиынтық сентиментке ең көбірек әсер ететін пікір бөліктеріне жеткілікті назар аудармайды (Saoualih et al., 2024). Сонымен қатар бұл модельдер сөз маңыздылығына қатысты жіктеуді көбіне елемейді, ал бұл әсіресе нәзік сентимент айырмашылықтарын жіктеуде нәтижеге ықпал етуі мүмкін (Ghatora et al., 2024).

Ұсынылған фреймворк Luong назар механизміне негізделген Bi-LSTM моделін пайдалану арқылы қазіргі тәсілдердің кемшіліктерін азайтады. Бұл терең оқыту моделі Bi-LSTM-нің тізбекті оқытудағы артықшылықтарын қолданады, ал Luong назар механизмі ең маңызды контекстті анықтайды. Мұндай екіжақты тәсіл күрделі артықшылықтар береді: ұзақ қашықтықтағы контекстті жақсы қамту және модельдің мәтіндегі ең релевантты бөліктерге назарын күшейту. Фреймворк сентиментті жіктеуді жақсартып отырып, өнім пікірінің мәтіндік құрылымына және маңызды сөздердің өзара байланысына байланысты қиындықтарға дәл әрі орнықты шешім ұсынады, өйткені өнім пікірлеріндегі сентиментті талдауда бұл екі аспект бірге қарастырылуы тиіс. Бұл жұмысты өзгелерден ерекшелендіретін негізгі белгі - тұтынушы сентиментін тереңірек түсіну үшін негізгі аспектілерді айқындайтын контекстке сезімтал назар механизмімен біріктірілген Bi-LSTM гибриді моделі.

Әдеби шолу. Liu et al. (2025) трансшекаралық электрондық коммерцияның белгілі бір санатындағы ең көп сатылатын тауарларға қатысты 14 078 онлайн пікірді зерттеді. Бұл зерттеуде пікір тақырыптарын, сентимент таралымдарын және олардың уақыт бойынша өзгерісін талдау үшін әлеуметтік желіні талдау мен сентиментті жіктеуге арналған бірнеше Python құралы қолданылды; Зерттеу цифрлық дәуірде трансшекаралық электрондық коммерция маркетингін жетілдіру үшін 4Ps және 4Cs маркетинг теорияларын біріктіруге болатынын ұсынады. Mabrouk et al.

(2021) «пікірлерді жинақтау және зерттеу» әдісін енгізді; H. Li et al. (2023) екі класты модельдің кемшіліктерін көрсетіп, күрделі әрі теңгерімсіз қытай мәтіндерінің сентиментін талдауды одан әрі жақсарту үшін үш класты жіктеу моделін және назар механизмі мен Bi-LSTM қолданатын жаңа диахрондық модельді ұсынды. Nichifor et al. (2023) бизнес пен электрондық коммерция маркетингтік стратегияларында жасанды интеллектті этикалық және тиімді қолдануды зерттеді. Авторлар 18 елде орналасқан 85 онлайн ритейлердің пікірлерін талдап, software-as-a-service (SaaS) негізіндегі табиғи тілді өңдеу және машиналық оқыту (ML) құралдарын пайдалану арқылы 1 687 пікірді қарастырды. Авторлар тұтынушылардың бір жұлдызды пікірлеріндегі бейтарап эмоционалдық күйге қатысты түсінік береді; Wang et al. (2024) CSEB ықпал ету механизмін аралас әдістер парадигмасы арқылы зерттеді. Бастапқы деректер тереңдетілген сұхбаттардан тұрды, ал қосымша деректер пайдаланушы жасаған контентті жинайтын краулерлер арқылы алынып, негізделген теория әдісімен талданды. Li et al. (2024) COVID-19 кезеңінде жаңа азық-түлік электрондық коммерциясы (FFEC) өнімдері мен қызметтеріне қатысты тұтынушы күтулері мен қалауларын зерттеді. Бұл зерттеу онлайн пікірлерден алынған деректермен мәтіндік майнингті қолданады. Guo et al. (2025) пайдаланушы қанағаттануына әсер ететін факторларды зерттеу үшін табиғи тілді өңдеу мен иерархиялық шешім қабылдауды біріктіретін екі кезеңді аналитикалық фреймворк ұсынды. Бірінші кезеңде эмоция моделі смартфон туралы жалған пікірлерді анықтайды, ал шынайы пікірлер қанағаттанумен байланысты тақырыптарды анықтау үшін арнайы әдістермен талданады. Қанағаттанумен байланысты тақырыптар бес негізгі факторға жіктеліп, брендтерді салыстырмалы талдауда қолданылды. Екінші кезеңде өнім пікірлері мен қауымдастықтағы термин жиіліктері салыстырылып, өнім дизайны мен маркетинг мақсаттары үшін маңызды жасырын тұтынушылық пайымдар анықталды. Fici et al. (2024) электрондық коммерциямен салыстырғанда метағалам платформаларындағы тұтынушы мінез-құлқын зерттеп, сатып алудың эмоционалдық және когнитивтік құрамдарын қамту үшін нейроғылыми әдіснаманы қолданды. Зерттеу метағаламның когнитивтік тартылуды арттыратынын көрсетеді, алайда ол салыстырмалы түрде ой еңбегін көбірек талап етеді және дәстүрлі электрондық коммерция деңгейіндегі оң эмоционалдық реакция тудырмайды.

Fan et al. (2022) тұрақты дамуға және брендтің халықаралықтануына жәрдемдесу үшін трансшекаралық электрондық коммерция брендтерін халықаралық бағалаудың индекстік жүйесін құрды. Latent Dirichlet Allocation (LDA) әдісі онлайн пікірлерден индикаторларды шығару үшін қолданылды. Классикалық әлеуетті индикаторларға Long Short-Term Memory (LSTM) оқытуы қолданылды. Kano қағидаттарын пайдалану арқылы ілгерілету стратегияларын анықтау мақсатында трансшекаралық электрондық коммерцияны басқарудың салдарларына сыни талдау жүргізілді. Huang et al. (2022) электрондық коммерция өнімдеріне қатысты қытай тіліндегі

пікірлердің сентиментін талдау мәселесін шешу үшін ERNIE-Bi-LSTM (EBLA) моделін ұсынды. Модель өлшемділікті бейнелеу, сентимент сөздерінің екіұштылығын жою және қытай пікірлеріндегі көпмағыналылық мәселелерін шешу үшін маңыздылығын көрсетті.

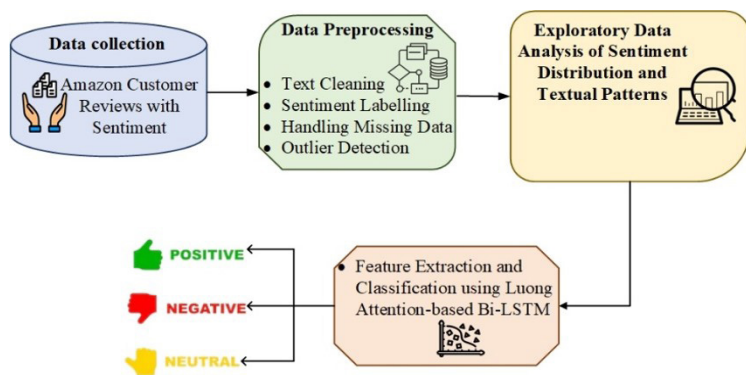
Yang et al. (2022) COVID-19 пандемиясы кезінде жаңа азық-түлік электрондық коммерциясы контекстінде «электрондық коммерция температурасының» тұтынушы қанағаттануына әсерін зерттеді. Бұл электрондық ритейлерлерге қызмет сапасын жақсарту және пандемиядан кейінгі қалпына келу кезеңінде төтенше жағдайларға бейімделу бойынша көзқарастар береді. Yuan et al. (2025) Қытайдың Чжэцзян провинциясындағы мандарин сатылымына электрондық коммерция пікірлерінің әсерін талдап, дәм, қызмет, сапа және бағаға назар аударды. Олар Latent Dirichlet Allocation (LDA) тақырыптық моделін пайдаланып, пікірлерді төрт тақырыпқа - сапа, қызмет, баға және дәм - топтастырды. Әр тақырып бойынша пікірлер SnowNLP арқылы қосымша бағаланып, жалпы оң сентименттер мен кейбір теріс кері байланыстар анықталды. Бұл зерттеу тұтынушы сұранысының әсерін және зерттеушілердің тұтынушы пікірлерін дәлірек түсіну қажеттілігін қарастырады, сондай-ақ болашақ зерттеулерде деректер дәлдігін және тұтынушының сатып алу тәжірибесін жақсартуға көмектесетін неғұрлым жетілдірілген тілдік модельдер мен процедуралық талдауларды қолдануды ұсынады.

Пиева et al. (2022) COVID-19 пандемиясы кезінде онлайн бөлшек сауда сатушыларын талдап, электрондық коммерциядағы тұтынушы мінез-құлқы мен қанағаттануына әсер ететін әртүрлі детерминанттар мен факторларды зерттеді. Зерттеу дәстүрлі және жаңа деректерді талдау әдістерін біріктіріп, сауалнамаларға сүйенді және онлайн пайдаланушылар тұрғысынан тиісті айнымалылар арасындағы байланыстар мен өзара ықпалдастықты зерттеу үшін құрылымдық теңдеулерді модельдеуді (SEM) қолданды. Chen et al. (2022) Қытайдың тікелей эфирдегі электрондық коммерция саясатының әсерін талдап, саясат күшіне енгенге дейінгі және кейінгі қоғамдық көзқарастар мен мәселелерге назар аударды. Сонымен қатар олар онлайн пікірлерді LDA модельдерімен талдап, сентимент полярлығы мен сентимент қарқындылығын бағалады; нәтижесінде 20-40 жас аралығындағы адамдар саясатқа басқа топтарға қарағанда көбірек ықылас білдіргенін, ал әйелдердің ерлерге қарағанда аз ықылас танытқанын анықтады. Анықталған тақырыптар экономикалық трансформация, алаяқтық, контрафактілік тауарлар және құқықтарды қорғау мәселелерімен байланысты.

Электрондық коммерцияның кеңеюі тұтынушы мінез-құлқын, әсіресе өнім пікірлері мен тұтынушы қанағаттануы тұрғысынан, айтарлықтай өзгертті. Тақырыптық модельдеу бойынша зерттеулер едәуір ілгерілегенімен, тілдің күрделілігі, тұтынушы күтулерінің өзгеруі және пандемиялар немесе реттеуші саясат сияқты сыртқы әсерлер тұтынушы мінез-құлқын тиімді қамтуды қиындататын көптеген мәселелерді туындатады.

Материалдар мен әдістер.

1-суретте Amazon тұтынушы пікірлерінің сентиментін талдаудың толық процедурасы көрсетілген. Үдеріс деректерді жинаудан басталады, мұнда пікірлерге сентимент таңбалары алынады. Келесі қадам - деректерді талдауға дайындайтын алдын ала өңдеу; ол мәтінді тазалауды, пікір рейтингтері негізінде сентимент таңбаларын тағайындауды, жетіспейтін деректерді өңдеуді және шеткі мәндерді анықтауды қамтиды. EDA сентимент таралуын зерттеу және пікірлердегі мәтіндік үлгілерді талдау үшін жүргізілді. EDA-дан кейінгі кезең - белгілерді шығару; мұнда маңызды белгілер анықталып, пікір мазмұнына назар аударатын Luong назар механизміне негізделген Bi-LSTM моделі көмегімен сентимент бойынша жіктеледі. Үдерістің нәтижесі пікірлерді оң, теріс және бейтарап санаттарға бөлу арқылы тұтынушы кері байланысы туралы түсінік береді.



Сурет 1 - Ұсынылған жалпы фреймворк

Деректерді жинау.

Деректер жиыны Kaggle платформасында қолжетімді Amazon тұтынушы пікірлерінен құрастырылған және 2013-2019 жылдар аралығын қамтиды. Смартфондар, кітаптар, ноутбуктар және тоңазытқыштар сияқты бірнеше санаттан алынған пікірлерде Category, Review Header, Review Text, Rating және Own_Rating (сентименттік жіктеу) сияқты қосымша белгілер бар. Деректер жиынында 110 000-нан астам пікір қамтылған, ал сентимент рейтингтері оң немесе теріс ретінде таңбаланған. Бұл сентиментті талдау жүргізу үшін құнды ресурс болып табылады, өйткені ол тұтынушылардың сезімдері мен көзқарастарын түсінуге мүмкіндік береді. Деректер жиыны еркін қолжетімді және CC0 1.0 Universal лицензиясы бойынша таралады, яғни оны кез келген мақсатта еркін пайдалануға болады.

Деректерді алдын ала өңдеу.

Алдын ала өңдеу мәтінді тазалауды (кіші әріпке келтіру, арнайы таңбаларды алып тастау, токенизация және лемматизация), рейтингтер негізінде сентиментті таңбалауды (1-2 - теріс, 3 - бейтарап, 4-5 - оң), сондай-ақ жетіспей-

тін деректерді мәндерді толықтыру (мода немесе орта мән арқылы) не толық емес жазбаларды алып тастау арқылы өңдеуді қамтиды. Бұл қадамдар модельдің таза, бірізді және толық деректерде оқытылуын қамтамасыз етіп, сентиментті дәл болжауға мүмкіндік береді.

Мәтінді тазалау - шикі мәтіндік деректерді дайындаудағы маңызды алдын ала өңдеу қадамы. Бірізді мәтін тазалау тәсілінің құндылығы - қажетсіз элементтерді алып тастауға және бүкіл мәтінді стандарттауға мүмкіндік беруінде. Үдерістің алғашқы қадамы мәтінді кіші әріпке келтіру болып табылады, яғни «Apple» және «apple» сияқты сөздер баламалы ретінде қарастырылуы үшін мәтіндегі барлық таңба кіші әріпке түрлендіріледі. Математикалық түрде s жолы үшін бұл түрлендіру (1)-теңдеуде берілген:

$$T_{\text{cleaned}} = \text{lowercase}(T) \quad (1)$$

Кіші әріпке келтіруден кейін арнайы таңбалар мен сандар жойылады. Бұл сентиментті бағалауда мағыналық жүктемесі жоқ барлық әріптік-сандық емес таңбаларды (мысалы, пунктуация мен цифрлар) алып тастауды білдіреді. Математикалық түрде оны (2)-теңдеу түрінде жазуға болады:

$$T_{\text{cleaned}} = \text{remove}(T, \{ \text{special characters, digits} \}) \quad (2)$$

Келесі кезеңде стоп-сөздерді жою «the», «is» және «and» сияқты жиі кездесетін, бірақ пайдалы ақпараты аз сөздерді сүзгіден өткізеді. Бұл амал T токендерінен S_{stop} стоп-сөздер жиынын алып тастау ретінде (3)-теңдеуде сипатталады:

$$T_{\text{cleaned}} = T - S \quad (3)$$

Осы тазалаудан кейін токенизация тазартылған мәтінді жеке токендер немесе сөздер жиынына түрлендіреді; олар кейін дербес өңделе алады.

Сентиментті таңбалау.

Сентиментті талдау әрбір пікірге оның сандық рейтингі негізінде сентимент таңбасын беру арқылы болжау мәселесін үш класты жіктеу мәселесіне түрлендіреді. r пікір рейтингі болсын, мұнда $r \in \{1, 2, 3, 4, 5\}$. Тағайындалған S сентимент таңбасы (4)-теңдеу бойынша анықталады:

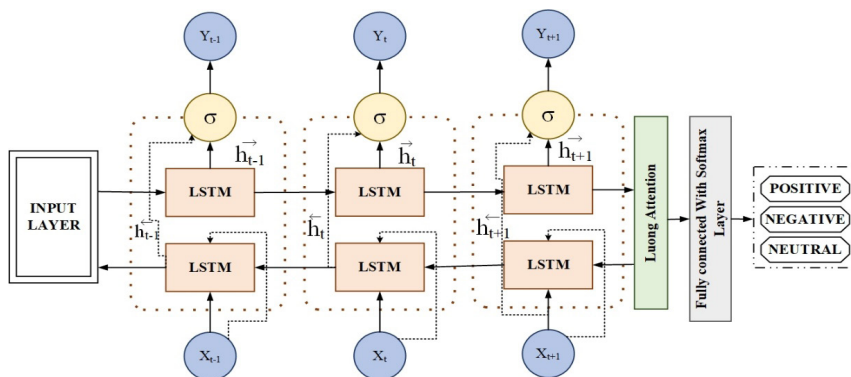
$$S = \begin{cases} \text{Negative} & \text{if } r \in \{1,2\} \\ \text{Neutral} & \text{if } r = 3 \\ \text{Positive} & \text{if } r \in \{4,5\} \end{cases} \quad (4)$$

Осылайша, рейтингі екіге тең пікір теріс, рейтингі үшке тең пікір бейтарап, ал рейтингі беске тең пікір оң ретінде жіктеледі. Бұл сәйкестендіру сандық

рейтингтерді модельдерді оқыту және бағалау үшін тиісті категориялық сентимент таңбаларына проекциялаудың айқын әрі бірізді әдісін қамтамасыз етеді.

Luong назар механизміне негізделген Bi-LSTM арқылы белгілерді шығару және жіктеу.

Сентиментті талдау Luong назар механизмімен бірге қолданылатын LSTM қабаттарына негізделеді. Кіріс - өнім туралы пікір сияқты тізбекті деректерді қабылдайтын кіріс қабаты. Бұл ақпарат бірнеше LSTM қабаты арқылы өңделеді. Әрбір LSTM қабаты әр уақыт қадамында жасырын күйін жаңарту арқылы кіріс мәтінінің тізбекті тәуелділіктерін шығарады. Luong назар механизмі LSTM қабаттарының соңында кіріс тізбегінің ең маңызды бөліктеріне шоғырлану үшін іске асырылады, бұл модельге релевантты сөздер немесе сөз тіркестеріне ең жоғары салмақ тағайындауға мүмкіндік береді. Соңында модель LSTM және назар механизмі арқылы алынған контекстке сүйене отырып, кірістің сентиментін - оң, теріс немесе бейтарап - анықтау үшін толық байланысқан қабатты және Softmax белсендіру функциясын қолданады (2-сурет).



Сурет 2 - Bi-LSTM архитектурасы

Кіріс қабаты. Мәтіндік деректердегі сөздік ендірулер немесе атрибуттар жиыны кіріс қабатына беріледі. Мұнда x_t - t қадамындағы ендіру, ал T - пікірдегі сөздер саны. Кіріс тізбегі әдетте Word2Vec, GloVe немесе BERT сияқты сөздік ендірулер арқылы қалыптастырылады, бұл x_t мәнінің t қадамындағы сөздің ендіруі немесе жоғары өлшемді векторлық көрінісі екенін білдіреді. Сондықтан кіріс тізбегін (5)-теңдеу түрінде жазуға болады:

$$X = \{x_1, x_2, \dots, x_T\}, \text{ where } x_t \in \mathbb{R}^d \quad (5)$$

Мұнда $\mathbb{R}^d$ әрбір сөз үшін d -өлшемді сөздік ендіру кеңістігін білдіреді.

Алға жылжу қақпалары. Алға өту кезінде LSTM кіріс деректерінің тізбегін қабылдап, оларды солдан оңға қарай бірізді түрде өңдейді. t уақыт қадамында

ол ағымдағы x_t кірісі және алдыңғы h_{t-1} жасырын күйі негізінде h_t жасырын күйін жаңартады. Алға өтудің негізгі элементтеріне кіріс қақпасы, ұмыту қақпасы, шығыс қақпасы және жад ұяшығының кандидат күйі жатады.

Кіріс қақпасы.

LSTM-де кіріс қақпасы t уақыт қадамындағы жаңа кірістің ұяшық күйіне қаншалықты енгізілетінін анықтайды. Ол жадыға түсетін жаңа ақпарат көлемін реттеп, тек релевантты кіріс элементтерінің сақталуын қамтамасыз етеді. Қақпа 0 мен 1 аралығында мән беретін sigmoid белсендіру функциясын қолданады, ол кіріс пен h_{t-1} жасырын күйінің үлесін сүзгіден өткізеді. Бұл (6)-теңдеуде көрсетілген:

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (6)$$

мұнда W_i - салмақ матрицасы, b_i - ығысу, ал σ - sigmoid белсендіру функциясы. Бұл үдеріс жад ұяшығын жаңарту үшін жаңа кірістің қандай үлесі қажет екенін анықтауға мүмкіндік береді және модельдің ұзақ мерзімді тәуелділіктерді үйрену қабілетін жақсартады.

Ұмытылу қақпасы. LSTM ұмыту қақпасы алдыңғы жадтың (C_{t-1} ұяшық күйінің) қай элементтері сақталуы және қайсысы ұмытылуы тиіс екенін анықтауда өте маңызды. Ол 0 мен 1 аралығында мән тудыратын sigmoid белсендіру функциясын қолданады; 0-ге жақын мән ұмытуды, ал 1-ге жақын мән сақтауды білдіреді (7):

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (7)$$

мұнда әрбір қабат W_f салмақ матрицасын, b_f ығысу векторын және әдетте sigmoid болып табылатын σ функциясын қамтиды. Осылайша LSTM жад бірлігі уақытқа тәуелді берілген ақпарат үшін релевантты немесе релевантты емес деп саналатын мәліметті сақтай да, қабылдамай да алады.

Шығыс қақпасы.

LSTM-де шығыс қақпасы ағымдағы ұяшық күйін ескере отырып, жасырын күйдің шығуын анықтайды. Оның қосымша қызметі жадтың тек мәнді бөліктерін келесі уақыт қадамына немесе келесі қабатқа өткізу. Шығыс қақпасы алдыңғы h_{t-1} жасырын күйі мен ағымдағы x_t кірісінің сызықтық түрлендіруіне sigmoid белсендіру функциясын қолдану арқылы 8-теңдеу бойынша есептеледі:

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (8)$$

мұнда W_o - салмақ матрицасы, b_o - ығысу мүшесі, ал σ - шығысты 0 мен 1 аралығына масштабтайтын sigmoid функциясы. Қақпалаудың бұл механизмі ұяшық ақпаратының таңдаулы түрде ашылуын қамтамасыз етеді.

Жад ұяшығының кандидат күйі. LSTM-дегі жад ұяшығының кандидат күйі t уақыт қадамында ұяшық күйіне қосылуы мүмкін жаңа әлеуетті ақпаратты есептейді. Қолданылатын белсендіру функциясы - мандерді -1 мен $+1$ аралығына сығатын гиперболалық тангенс ($\tanh$); сондықтан жаңартылған мән оң да, теріс те болуы мүмкін. Кіріс қақпасымен бірге бұл кандидат мән жаңа ақпараттың ұяшық күйіне қандай фактор бойынша қосылатынын анықтайды және 9-теңдеу түрінде жазылады:

$$\hat{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (9)$$

мұнда W_c матрицасында салмақтар сақталады, ал b_c - ығысу мүшесі. Сонымен қатар алдыңғы жасырын күй h_{t-1} , ал x_t - ағымдағы кірісі.

Ұяшық күйін жаңарту. Ұзақ мерзімді тәуелділіктерді сақтау үшін LSTM-дегі ұяшық күйін жаңарту бұрынғы жад пен жаңа ақпаратты біріктіреді. Бұл алдыңғы C_{t-1} ұяшық күйін f_t ұмыту қақпасымен, ал $\hat{C}_t$ кандидат жад ұяшығын i_t кіріс қақпасымен көбейту арқылы орындалады. Ұмыту және кіріс қақпалары 10-теңдеуде көрсетілген:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \hat{C}_t \quad (10)$$

мұнда f_t - C_{t-1} релевантты өткен ақпаратты сақтайды, ал i_t - $\hat{C}_t$ саналы түрде тандалған жаңа ақпаратты қосады. Осы үдерістердің барлығы LSTM-ге жадыны сақтау мен жаңарту арасындағы тепе-теңдікті қамтамасыз етуге мүмкіндік береді, бұл тізбекті тәуелділіктерді тиімді өндеуге жағдай жасайды.

Жасырын күй.

LSTM әр уақыт қадамындағы соңғы шығысты жасырын күйді жаңарту арқылы есептейді: ағымдағы C_t ұяшық күйі o_t шығыс қақпасының реттеуімен көбейтіледі. C_t ұяшық күйі алдымен $\tanh$ белсендіру функциясы арқылы -1 мен 1 аралығына келтіріледі, содан кейін келесі қабаттар үшін басқарылатын мандерге айналады. Нәтиже шығыс қақпасымен элементтік түрде масштабталады; бұл ақпараттың қаншалықты ашылатынын (13)-теңдеуде анықтайды:

$$h_t = o_t \cdot \tanh(C_t) \quad (11)$$

мұнда h_t - t уақытындағы жасырын күй, o_t - шығыс қақпасы, ал C_t - жаңартылған ұяшық күйі. Бұл жасырын күй LSTM бірлігінің тікелей шығысы болып табылады және келесі уақыт қадамына немесе жоғарғы қабатқа кіріс ретінде беріледі, бұл желінің тізбекті өндеуде релевантты контекстік ақпаратты тасымалдауына мүмкіндік береді.

Кері өту қақпасы.

Екібағытты LSTM-дегі кері өту кіріс тізбегін оңнан солға қарай өңдейді, бұл модельге алға өту жалғыз өзі бере алмайтын келешек сөздер немесе

токендерден контекстік ақпаратты қамтуға мүмкіндік береді. Әрбір t уақыт қадамында ол алға өтудегідей қақпа механизмдерін, ұяшық күйін жаңарту және жасырын күй теңдеулерін пайдалана отырып, жеке кері жасырын күйді жаңартады, бірақ олар кері тәртіппен қолданылады. Бұл кері LSTM де кіріс тізбегін кері бағытта өту арқылы өзінің жасырын күйлерін, ұяшық күйлерін және қақпа мәндерін есептейтінін білдіреді. Алға және кері өтулердің шығыстарын біріктіру арқылы Bi-LSTM тізбектің бай әрі толық көрінісін қалыптастырып, өткен және келешек контекстен тәуелділіктерді тиімді қамтиды.

Bi-LSTM.

Bi-LSTM кіріс деректерін алға және кері бағыттарда өңдеу арқылы тізбекті модельдеуді жақсартады. Алға өтуде жасырын күй 12-теңдеудегідей есептеледі:

$$h_t = o_t \cdot \tanh(C_t) \quad (12)$$

мұнда h_t – t уақыт қадамына дейінгі өткен контексті қамтиды. Кері өтуде жасырын күй 13-теңдеумен есептеледі:

$$\hat{h}_t = \hat{o}_t \cdot \tanh(\hat{C}_t) \quad (13)$$

h_t тізбектің соңынан t қадамына дейінгі келешек контексті білдіреді. Әрбір қадамда алға және кері күйлер біріктіріліп, байырақ көрініс (16)-теңдеу арқылы қалыптастырылады:

$$h_t^{BiLSTM} = [h_t, \hat{h}_t] \quad (14)$$

Бұл көрініс өткен және келешек ақпаратты біріктіреді. Біріктірілген жасырын күй кейінгі қабаттарда сентиментті талдау сияқты міндеттер үшін қолданылады және модельге толық контекстік ақпаратты пайдаланып неғұрлым дәл болжам жасауға мүмкіндік береді.

Luong назар қабаты.

Назар механизмі бар Bi-LSTM негізіндегі sequence-to-sequence моделі қолданылады. Кодер Bi-LSTM қабаттары 1 және 2 кіріс тізбектерін қабылдап, тиісінше алға және кері жасырын күйлерді есептейді және біріктірілген жасырын күйді береді. Назар механизмі әр уақыт қадамында α назар салмақтарын есептейді; олар кодер мен декодер жасырын күйлерінің арасындағы сәйкестік бағасы арқылы анықталып, softmax функциясымен 15-теңдеу бойынша нормаланады:

$$\alpha_t = \text{softmax}(a(h_t, s_{t-1})) \quad (15)$$

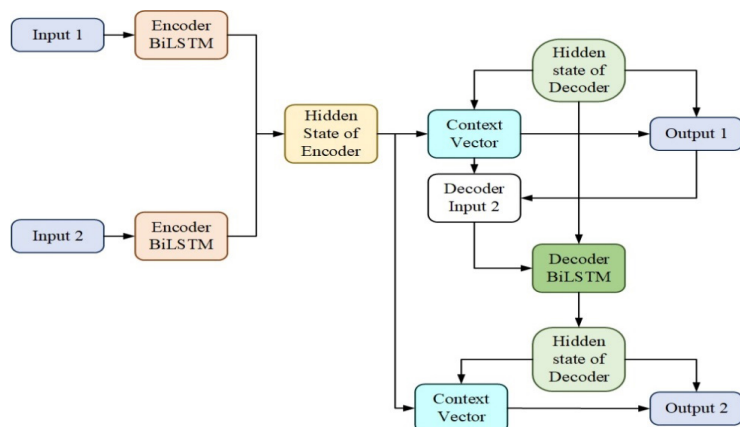
Содан кейін контекст векторы c_t кодердің жасырын күйлерінің салмақталған орташа мәні ретінде 16-тендеу бойынша бағаланады:

$$c_t = \sum_{t=1}^T \alpha_t h_t \quad (16)$$

Бұл контекст векторы жасырын күйлерді тудыратын Decoder Bi-LSTM-ге беріледі. Соңында модель softmax қабаты арқылы y_t шығыс тізбегін (19)-тендеу бойынша қалыптастырады:

$$\hat{y}_t = \text{softmax}(W s_t + b) \quad (17)$$

мұнда W_y - салмақ матрицасы, ал b_y - ығысу мүшесі. Бұл архитектура модельге екібағытты контекстті ұсынуға және 3-суретте көрсетілген назар механизмі арқылы релевантты кіріс ақпаратына шоғырлануға мүмкіндік береді.



Сурет 3 - Luong назар механизмі

Толық байланысқан қабат.

Bi-LSTM негізіндегі модельдегі бұл қабат біріктірілген жасырын күйлерді классқа тән баллдарға түрлендіретін соңғы кезең қызметін атқарады. Біріктірілген h_t жасырын күйі тығыз қабат арқылы өткізіледі, мұнда салмақтар мен ығысулар арқылы сызықтық түрлендіру 18-тендеу бойынша қолданылады:

$$Z = W_{fc} \cdot h_t^{BiLSTM} + b_{fc} \quad (18)$$

мұнда W_{fc} - салмақ матрицасы, b_{fc} - ығысу мүшесі, ал z әр сентимент класы үшін нормаланбаған баллдарды білдіреді. Бұл баллдарды түсіндірілетін ықтималдықтарға түрлендіру үшін softmax белсендіру функциясы 19-тендеу бойынша қолданылады:

$$P(\text{Class} = c) = \frac{e^{z_c}}{\sum_{c'} e^{z_{c'}}} \quad (19)$$

мұнда z_i - i класының балы, ал бөлім барлық ықтимал кластар бойынша нормалауды қамтамасыз етеді. Бұл шығыстардың оң, теріс және бейтарап категориялары бойынша жарамды ықтималдық таралуын қалыптастыруын қамтамасыз етеді. Модель болжанған сентиментті ең жоғары ықтималдыққа ие класты таңдау арқылы анықтайды. Бұл қадам өте маңызды, себебі ол жоғары өлшемді Bi-LSTM көріністерін соңғы жіктеу шешімімен байланыстырып, сентиментті дәл болжауға мүмкіндік береді.

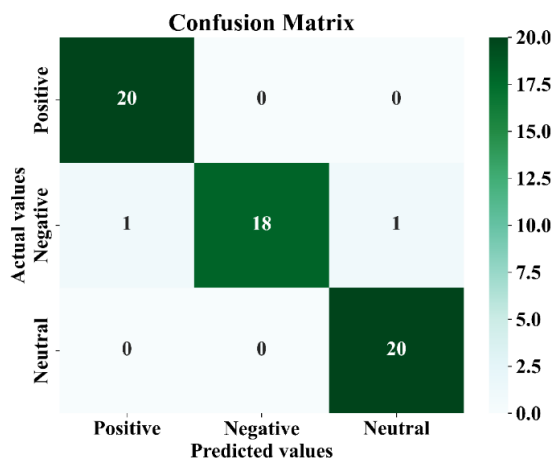
Шығыс қабаты.

Bi-LSTM негізіндегі сентиментті жіктеу моделі үш сентимент категориясының бірін - оң, теріс немесе бейтарап - таңдау арқылы соңғы болжамды шығарады. Толық байланысқан қабат логиттерді есептеп, softmax функциясы оларды әр класс үшін ықтималдықтарға түрлендіргеннен кейін, ең жоғары ықтималдыққа ие класс соңғы шешім ретінде таңдалады. Бұл математикалық түрде 20-тендеуде көрсетілген:

$$\hat{y} = \arg \max_c P(\text{Class} = c) \quad (20)$$

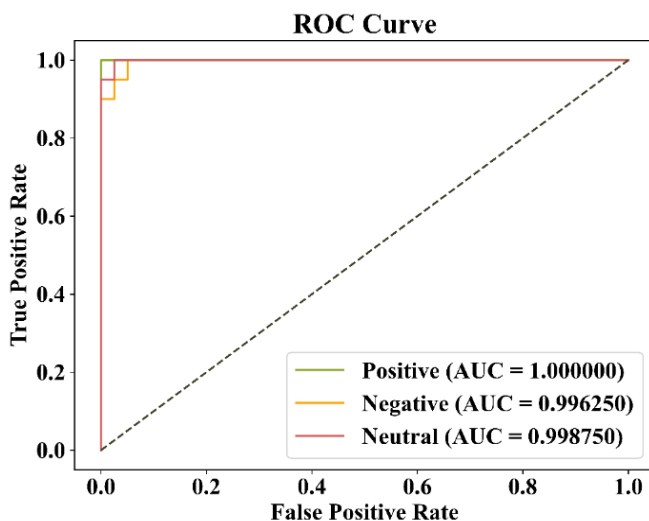
мұнда $\hat{S}$ - болжанған сентимент таңбасы, ал p_i - i класы үшін softmax ықтималдығы. Ең жоғары ықтималдыққа ие класс кіріс мәтінінің сентименті ретінде тағайындалады, бұл айқын және түсіндірілетін жіктеу нәтижесін қамтамасыз етеді.

Нәтижелер және талқылау. Зерттеу жіктеу моделінің үш класс бойынша - оң, теріс және бейтарап - тиімді жұмыс істейтінін дәлелдейді. Модель өте жоғары дәлдікке қол жеткізді, ал теріс мысалдарда аз ғана қате жіктеу байқалды. Штатастыру матрицасы, ROC қисығы және precision-recall қисығы модельдің кластарды өте аз қателікпен ажырататынын көрсетеді. Ол шекті жағдайларда да нақтылық пен толықтық бойынша жақсы нәтиже береді. Сонымен қатар accuracy, precision, recall және F1-score есептеулері де жоғары. Оқыту және валидация метрикалары модельдің деректерге жақсы жалпыланатынын, яғни айтарлықтай артық үйренбейтінін көрсетеді. Эпохалар өткен сайын тұрақты жақсару байқалады.



Сурет 4 - Шатастыру матрицасы

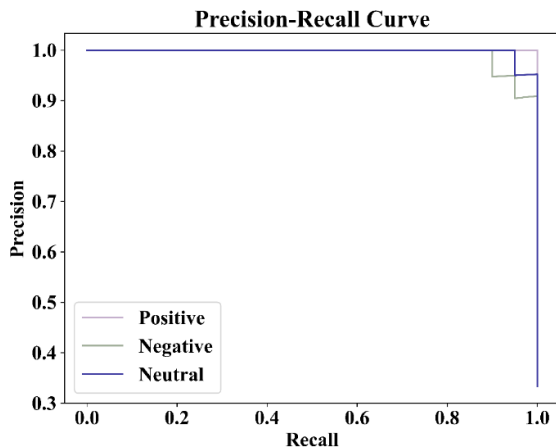
4-сурет үш класка - оң, теріс және бейтарап - арналған модельдің жіктеу нәтижелерін көрсетеді. Модель 20 оң, 18 теріс және 20 бейтарап жағдайды дұрыс анықтады. Қате жіктеулер шамалы болды: бір теріс жағдай оң ретінде, ал бір теріс жағдай бейтарап ретінде жіктелді. Оң немесе бейтарап болжамдарда қате жіктеу болған жоқ. Матрица модель дәлдігінің жоғары екенін көрсетеді, себебі модель бейтарап және оң кластарды жақсы ажыратады, дегенмен теріс класта аздап шатасу бар. Жалпы модель тиімді және қателері өте аз.



Сурет 5 - ROC қисығы

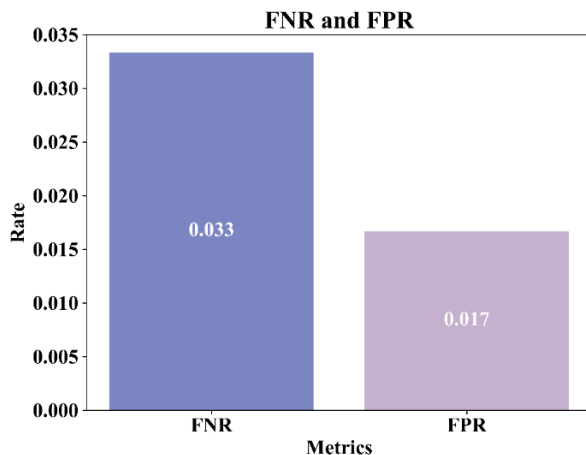
Жасыл қисықтың AUC мәні 1,000000 болды, бұл толық дерлік жіктелу мүмкіндігін білдіреді. Қызғылт сары қисықтың AUC мәні 0,996250 болып,

өте жақсы нәтиже мен аз ғана қателікті көрсетеді. Қызыл қисықтың AUC мәні 0,998750, бұл да өте жоғары көрсеткіш. 5-суретте көрсетілгендей, үш қисықтың барлығы жоғарғы сол жақ бұрышқа жақын орналасқан, бұл модельде шынайы оң үлестің жоғары, ал жалған оң үлестің төмен екенін білдіреді.



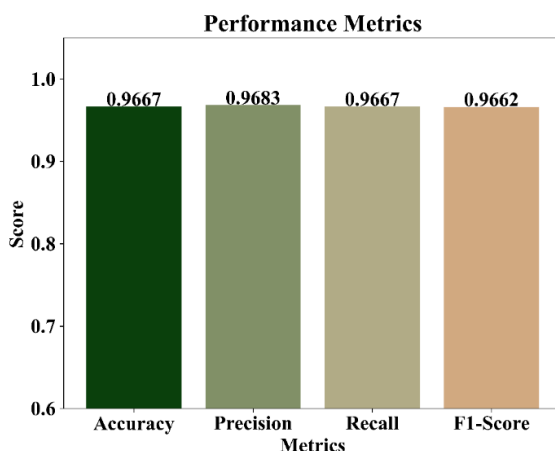
Сурет 6 - Precision-recall қисығы

Барлық үш класс бойынша қисықтар дерлік идеалға жақын және жоғарғы оң жаққа жақын орналасқан, бұл нақтылық пен толықтықтың жоғары екенін көрсетеді. Оң (күлгін қисық), теріс (жасыл қисық) және бейтарап (көк қисық) кластардың барлығында precision мәндері 1,0-ге жуық, ал recall мәндері де соған жақын. Бұл модельдің барлық жағдайларды минималды жалған оң нәтижелермен дұрыс анықтауға жоғары қабілетті екенін білдіреді және 6-суретте көрсетілгендей, жалпы өнімділік өте жоғары.



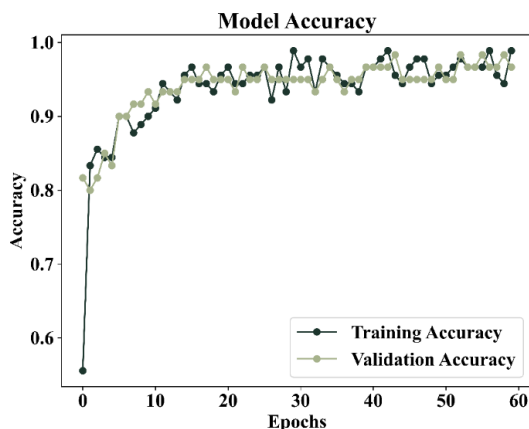
Сурет 7 - FNR және FPR

7-суретте жіктеу моделінің жалған теріс үлесі (FNR) мен жалған оң үлесі (FPR) салыстырылған. Көк бағанмен белгіленген $FNR = 0,033$, бұл нақты оң жағдайлардың шамамен 3,3%-ы теріс ретінде қате жіктелгенін көрсетеді. Күлгін баған $FPR = 0,017$ екенін, яғни нақты теріс жағдайлардың шамамен 1,7%-ы оң ретінде қате жіктелгенін көрсетеді. График жалған теріс үлестің жалған оң үлестен жоғары екенін, бірақ оның да салыстырмалы түрде төмен деңгейде екенін көрсетеді.



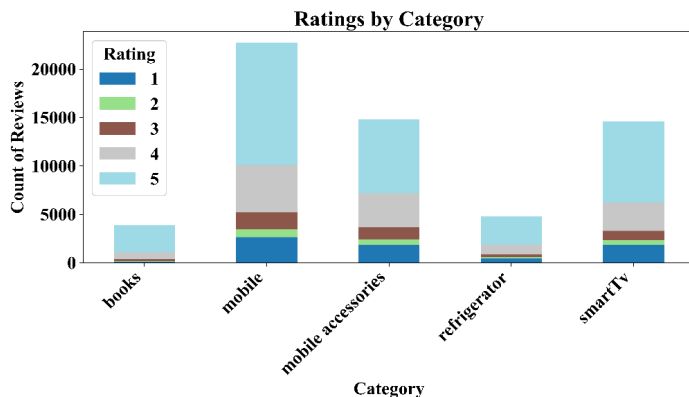
Сурет 8 - Өнімділік метрикалары

Ассурасу мәні 0,9667 болды, бұл болжамдардың 96,67%-ы дұрыс екенін көрсетеді. Precision 0,9683 болды, яғни оң болжамдардың 96,83%-ы дұрыс. Recall = 0,9667, бұл шынайы оң жағдайлардың 96,67%-ы анықталғанын білдіреді. F1-score 0,9662 болды, ол precision мен recall арасындағы теңгерімді сипаттайды. 8-суретте көрсетілгендей, модель барлық маңызды метрикалар бойынша жақсы нәтиже береді, олардың мәндері 1-ге жуық.



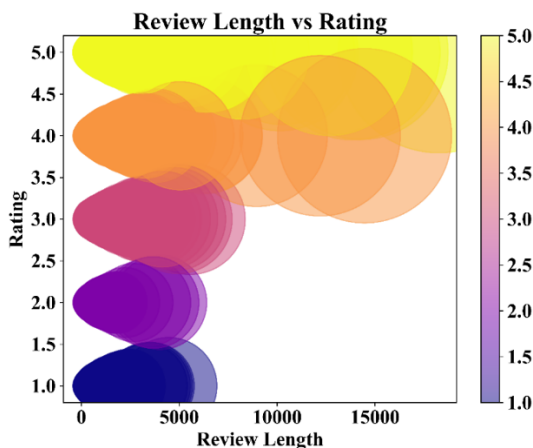
Сурет 9 - Модель дәлдігі

9-суретте модельдің 60 эпоха бойынша оқыту және валидация дәлдігі көрсетілген. Бастапқыда екі дәлдік те төмен болды (шамамен 0,6), бірақ оқыту үдерісі ілгерілеген сайын біртіндеп өсті. Оқыту дәлдігі (қою жасыл) жақсарып, 1,0-ге жақындайды, бұл модельдің жақсы үйреніп жатқанын көрсетеді. Валидация дәлдігі (ашық жасыл) де өсіп, оқытудың соңында шамамен 0,95-ке жетеді. Оқыту және валидация дәлдіктерінің айырмасы аз болды, бұл модельдің жақсы жалпыланғанын және артық үйренбегенін білдіреді.



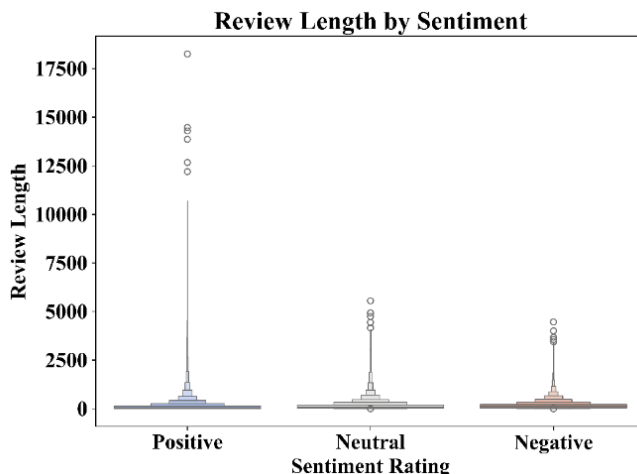
Сурет 10 - Санаттар бойынша рейтингтер

Mobile санатында пікірлер саны ең көп, 20 000-нан астам пікір бар; олардың басым бөлігі 5 жұлдызды (ашық көк) және 4 жұлдызды (қою сұр) рейтингтер. SmartTV пікірлері ең аз, шамамен 4 000, әрі олар негізінен 4 және 5 жұлдызды. Mobile Accessories және Refrigerators санаттарында пікірлер саны орташа деңгейде, ал Mobile Accessories санатында пікірлер арақатынасы анағұрлым теңгерімді, мұнда 3 және 2 жұлдызды пікірлердің үлесі салыстырмалы түрде жоғары. Books санатындағы пікірлер саны ең аз, 5 000-нан төмен, және рейтингтер әртүрлі деңгейлер бойынша біршама біркелкі таралған. 10-суреттегі диаграмма мобильді телефондарға қатысты өнімдердің басқа өнімдерге қарағанда жоғарырақ бағаланатынын көрсетеді.



Сурет 11 - Пікір ұзындығы және рейтинг

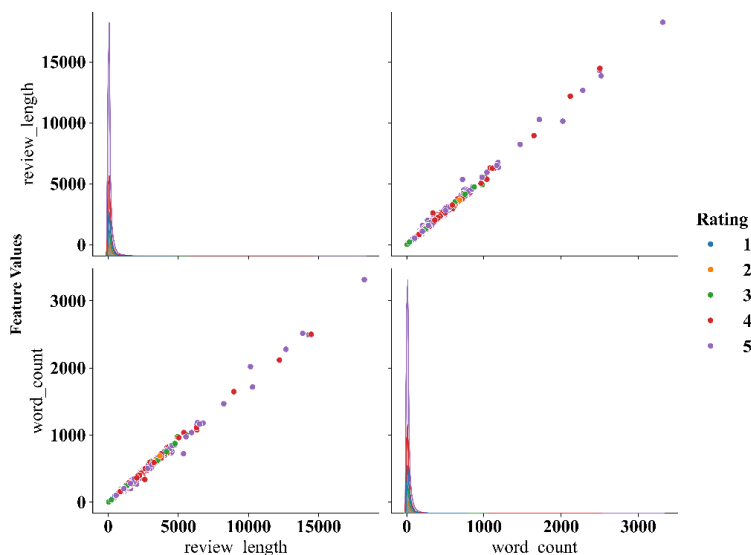
11-суретте x осінде пікір ұзындығы (0-15 000 таңба), ал y осінде рейтинг (1-5) көрсетілген. Көпіршік өлшемі сол санаттағы пікірлер санын білдіреді: үлкен көпіршіктер пікір санының көбірек екенін көрсетеді. Көріп отырғанымыздай, жоғарырақ бағаланған пікірлерде (тиісінше сары және ашық қызғылт түспен белгіленген 4,5 және 5 жұлдыз) пікір ұзындығы ұзағырақ және 5 000 таңбадан асады. Керісінше, төмен бағаланған пікірлер (күлгін және көк түспен белгіленген 1 және 2 жұлдыз) қысқарақ, олардың ұзындығы көбіне 0-5 000 таңба аралығында.



Сурет 12 - Сентимент бойынша пікір ұзындығы

12-суретте оң, бейтарап және теріс деп бөлінген сентимент рейтингтері бойынша пікір ұзындығының таралуы көрсетілген. X осі сентимент рейтингін, ал y осі таңбалар саны түріндегі пікір ұзындығын білдіреді. Оң

пікірлер жағдайында пікірдің медианалық ұзындығы көбірек, ал кейбір пікірлер 17 500 таңбадан асып, шеткі мәндер ретінде белгіленеді. Оң пікірлер кеңірек таралған және ұзын құйрыққа ие, бұл оң пікірлердің әдетте ұзағырақ әрі егжей-тегжейлі болатынын көрсетеді. Бейтарап пікірлер ұзындығы бойынша қысқарак, медианасы шамамен 300 таңба, ал пікірлердің көпшілігі шағын ауқымда орналасып, аз егжей-тегжейлі кері байланысты білдіреді. Теріс пікірлердің үлгісі бейтарап пікірлерге ұқсас, бірақ пікір ұзындығы сәл ұзағырақ; кейбіреулері 1 000 таңбаға дейін жетеді.



Сурет 13 - пікір ұзындығы, рейтинг және сөз саны

Рейтинг пен пікір ұзындығы және сөз саны бойынша шашырау графиктері матрицасы әртүрлі рейтингтер (1-5) үшін пікір ұзындығы мен сөз саны арасындағы байланыстарды көрсетеді (13- сурет). Шашырау графиктері пікір неғұрлым ұзын болса, пікірдегі сөз саны да соғұрлым жоғары болатынын және рейтингтің де жоғарырақ (әсіресе күлгін және қызыл түспен белгіленген 4 және 5 жұлдыз) болатынын көрсетеді. 1 жұлдызды рейтингтері бар пікірлердің (көк түспен) ұзындығы мен сөз саны әдетте қысқарак. Пікір ұзындығы мен сөз саны арасында жоғары корреляция байқалады. Сөз саны көп пікірлер (500-ден 3 000 сөзге дейін) көбіне жоғарырақ бағаланады, әсіресе 4 және 5 жұлдызды рейтингтерде. 1 жұлдызды пікірлер айтарлықтай қысқарак болды және олардың көпшілігі 500 сөзден аз қамтыды. 16-суретте көрсетілгендей, көлемді әрі ұзақ пікірлер, әдетте, жоғары бағамен байланысты болуы ықтимал.

Қорытынды. Бұл мақалада электрондық коммерция өнімдері туралы пікірлердің сентиментін талдау үшін Luong назар механизмін толықтырылған Bi-LSTM архитектурасына негізделген жетілдірілген сентиментті талдау жүйесі ұсынылды. Өнім пікірлеріндегі сентиментті талдауда көпқабатты

мағыналар мен күрделі құрылымдардан тұратын мәтіндік деректерді оқу және пайымдау қиындықтарын шешуде бұл тәсіл тиімді, өйткені ол модельге мәтіндегі ұзақ арақашықтықтағы тәуелділіктерді ескеруге және мәтіннің маңызды бөліктеріне назар аударуға мүмкіндік береді. Нәтижелер ұсынылған фреймворк сентиментті жіктеу дәлдігін арттыруда анағұрлым тиімді екенін және оң, теріс және бейтарап сентименттерді жоғары дәлдік пен толықтықпен ажырататыны бойынша дәстүрлі тәсілдермен салыстырғанда жақсырақ нәтиже беретінін көрсетеді. Модель 96,67% жоғары ассурасу мәнін көрсетті, ал precision және recall тиісінше 96,83% және 96,67% болды. Оқыту мен валидация өнімділігінің тұрақтылығы модельдің жақсы жалпыланатынын және артық үйрену деңгейінің төмен екенін көрсетті. Дегенмен модельді одан әрі жетілдіру үшін қосымша назар механизмдерін немесе доменге тән сипаттамаларды зерттеу неғұрлым нақты түсініктер беруге мүмкіндік береді. Болашақ жұмыстар модельді сарказм немесе контекстке тән эмоциялар сияқты күрделі сентимент вариацияларын өңдеуге бейімдеуге және фреймворкты басқа тілдерге әрі әртүрлі электрондық коммерция деректеріне кеңейтуге бағытталуы мүмкін, бұл оның әмбебаптығын арттырады.

Деректердің қолжетімділігі

<https://www.kaggle.com/datasets/thedevastator/amazon-customer-reviews-with-2013-2019-sentiment>

References

- Bello A., Ng S.-C., and Leung M.-F. (2023) A BERT framework to sentiment analysis of tweets. *Sensors*, 23. — P. 506. <https://doi.org/10.3390/s23010506> (in Eng.).
- Chen T., Tong C., Bai Y., Yang J., Cong G., and Cong T. (2022) Analysis of the public opinion evolution on the normative policies for the live streaming e-commerce industry based on online comment mining under COVID-19 epidemic in China. *Mathematics*, 10. — P. 3387. <https://doi.org/10.3390/math10183387> (in Eng.).
- Dickson Z.A., Prakoso M.R.B., Putra M.S.Q., Syaputra M.S.A.F., Achmad S., and Sutoyo R. (2023) Sentiment analysis for customer review: Case study of Traveloka. *Procedia Comput. Sci.*, 216. — P. 682–690. <https://doi.org/10.1016/j.procs.2022.12.184> (in Eng.).
- Dubey P., Dubey P., and Bokoro P.N. (2025) Unpacking sarcasm: A contextual and transformer-based approach for improved detection. *Computers*, 14. — P. 95. <https://doi.org/10.3390/computers14030095> (in Eng.).
- Fan M., Tang Z., Qalati S.A., Tajeddini K., Mao Q., and Bux A. (2022) Cross-border e-commerce brand internationalization: An online review evaluation based on Kano model. *Sustainability*, 14. — P. 13127. <https://doi.org/10.3390/su142013127> (in Eng.).
- Fici A. et al. (2024) From e-commerce to the metaverse: A neuroscientific analysis of digital consumer behavior. *Behav. Sci.*, 14. — P. 596. <https://doi.org/10.3390/bs14070596> (in Eng.).
- Ghatora P.S., Hosseini S.E., Pervez S., Iqbal M.J., and Shaukat N. (2024) Sentiment analysis of product reviews using machine learning and pre-trained LLM. *Big Data Cogn. Comput.*, 8. — P. 199. <https://doi.org/10.3390/bdcc8120199> (in Eng.).
- Guo P., Li H., and Mo X. (2025) Quantifying post-purchase service satisfaction: A topic–emotion fusion approach with smartphone data. *Big Data Cogn. Comput.*, 9. — P. 125. <https://doi.org/10.3390/bdcc9050125> (in Eng.).
- Hassan S.U., Ahamed J., and Ahmad K. (2022) Analytics of machine learning-based algorithms for text classification. *Sustain. Oper. Comput.*, 3. — P. 238–248. <https://doi.org/10.1016/j.susoc.2022.03.001> (in Eng.).

Huang W., Lin M., and Wang Y. (2022) Sentiment analysis of Chinese e-commerce product reviews using ERNIE word embedding and attention mechanism. *Appl. Sci.*, 12. — P. 7182. <https://doi.org/10.3390/app12147182> (in Eng.).

Idries A., Krogstie J., and Rajasekharan J. (2022) Dynamic capabilities in electrical energy digitalization: A case from the Norwegian ecosystem. *Energies*, 15. — P. 8342. <https://doi.org/10.3390/en15228342> (in Eng.).

Ilieva G., Yankova T., Klisarova S., and Dzhabarova Y. (2022) Customer satisfaction in e-commerce during the COVID-19 pandemic. *Systems*, 10. — P. 213. <https://doi.org/10.3390/systems10060213> (in Eng.).

Kokab S.T., Asghar S., and Naz S. (2022) Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, 14. — P. 100157. <https://doi.org/10.1016/j.array.2022.100157> (in Eng.).

Li H., Lu Y., Zhu H., and Ma Y. (2023) A novel AB-CNN model for multi-classification sentiment analysis of e-commerce comments. *Electronics*, 12. — P. 1880. <https://doi.org/10.3390/electronics12081880> (in Eng.).

Li Y., Shen Z., Zhao C., Chin K.-S., and Lang X. (2024) Understanding customer opinion change on fresh food e-commerce products and services-Comparative analysis before and during COVID-19 pandemic. *Sustainability*, 16. — P. 2699. <https://doi.org/10.3390/su16072699> (in Eng.).

Liu C., Chen T., Pu Q., and Jin Y. (2025) Text mining for consumers' sentiment tendency and strategies for promoting cross-border e-commerce marketing using consumers' online review data. *J. Theor. Appl. Electron. Commer. Res.*, 20. — P. 125. <https://doi.org/10.3390/jtaer20020125> (in Eng.).

Mabrouk A., Redondo R.P.D., and Kayed M. (2021) SEOpinion: Summarization and exploration of opinion from e-commerce websites. *Sensors*, 21. — P. 636. <https://doi.org/10.3390/s21020636> (in Eng.).

Madanchian M. (2024) The role of complex systems in predictive analytics for e-commerce innovations in business management. *Systems*, 12. — P. 415. <https://doi.org/10.3390/systems12100415> (in Eng.).

Nichifor E. et al. (2023) Utilising artificial intelligence to turn reviews into business enhancements through sentiment analysis. *Electronics*, 12. — P. 4538. <https://doi.org/10.3390/electronics12214538> (in Eng.).

Niculescu L., and Tudorache M.T. (2022) Human-computer interaction in customer service: The experience with AI chatbots-A systematic literature review. *Electronics*, 11. — P. 1579. <https://doi.org/10.3390/electronics11101579> (in Eng.).

Saoualih A. et al. (2024) Exploring the tourist experience of the Majorelle Garden using VADER-based sentiment analysis and the latent Dirichlet allocation algorithm: The case of TripAdvisor reviews. *Sustainability*, 16. — P. 6378. <https://doi.org/10.3390/su16156378> (in Eng.).

Taherdoost H., and Madanchian M. (2023) Artificial intelligence and sentiment analysis: A review in competitive research. *Computers*, 12. — P. 37. <https://doi.org/10.3390/computers12020037> (in Eng.).

Wang H., Li J., Wu H., Hovy E., and Sun Y. (2023) Pre-trained language models and their applications. *Engineering*, 25. — P. 51–65. <https://doi.org/10.1016/j.eng.2022.04.024> (in Eng.).

Wang R., Xu S., Li S., and Pang Q. (2024) Research on influence mechanism of consumer satisfaction evaluation behavior based on grounded theory in social e-commerce. *Systems*, 12. — P. 572. <https://doi.org/10.3390/systems12120572> (in Eng.).

Xiong S., Tian W., Si H., Zhang G., and Shi L. (2024) A survey of the applications of text mining for the food domain. *Algorithms*, 17. — P. 176. <https://doi.org/10.3390/a17050176> (in Eng.).

Yang Y., Ma Y., Wu G., Guo Q., and Xu H. (2022) The insights, “comfort” effect and bottleneck breakthrough of “e-commerce temperature” during the COVID-19 pandemic. *J. Theor. Appl. Electron. Commer. Res.*, 17. — P. 1493–1511. <https://doi.org/10.3390/jtaer17040075> (in Eng.).

Yuan L., Liu H., Fu F., Liu Y., Zuo X., and Li L. (2025) Study of Zhejiang tangerine e-commerce reviews based on natural language processing. *Horticulturae*, 11. — P. 151. <https://doi.org/10.3390/horticulturae11020151> (in Eng.).

Zhang X., and Guo C. (2024) Research on multimodal prediction of e-commerce customer satisfaction driven by big data. *Appl. Sci.*, 14. — P. 8181. <https://doi.org/10.3390/app14188181> (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 286–301

<https://doi.org/10.32014/2026.2518-1726.439>

IRSTI: 28.23.37; 20.23.17

UDC 004.912; 004.89

© **Marassulov U.A. *, Kazbekova G., 2026.**Khoja Akhmet Yassawi International Kazakh-Turkish University,
Turkistan, Kazakhstan.

E-mail: marasulov.usen2024@ayu.edu.kz

**TF-IDF-BASED FAKE NEWS DETECTION
IN KAZAKH AND RUSSIAN****Marassulov Usen** — PhD student, Khoja Akhmet Yassawi International Kazakh-Turkish University,
Turkistan, Kazakhstan,E-mail: marasulov.usen2024@ayu.edu.kz, ORCID ID: <https://orcid.org/0009-0008-0801-1229>;**Kazbekova Gulnur** — Candidate of Technical Sciences, Associate Professor, Khoja Akhmet
Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan,E-mail: gulnur.kazbekova@ayu.edu.kz, ORCID ID: <https://orcid.org/0000-0002-2756-7926>.

Abstract. Relevance: The rapid spread of unverified information in digital media and social networks makes automatic fake news detection an important task for natural language processing and machine learning. For Kazakh, this task is particularly relevant because open labeled corpora are limited and ready-made models are not sufficiently adapted to local media discourse. Purpose: The study evaluates TF-IDF-based machine learning models for fake/real text classification in Kazakh and Russian and forms an interpretable baseline for later comparison with transformer-based approaches. Methods: The experiments used a balanced Kazakh-Russian corpus of 1,808 texts, with 452 fake and 452 real texts in each language. Logistic Regression, Linear SVM and Complement Naive Bayes were tested with word-level and character-level TF-IDF features in bilingual, monolingual and cross-lingual scenarios. Model quality was assessed using accuracy, precision, recall, F1-score and confusion matrices. Results: In bilingual, Kazakh-only and Russian-only settings, the best models achieved Macro-F1 of approximately 0.985. Cross-lingual evaluation showed clear directional asymmetry: Kazakh-to-Russian transfer reached Macro-F1 = 0.654, whereas Russian-to-Kazakh transfer reached Macro-F1 = 0.926. Practical significance: The results provide a reproducible and explainable benchmark for Kazakh-Russian fake news detection. They also define a basis for future comparison with multilingual BERT and XLM-RoBERTa under stricter source-based and temporal evaluation settings.

Keywords: fake news, disinformation, Kazakh language, Russian language, TF-IDF, machine learning, cross-lingual classification

For citations: Marassulov U.A., Kazbekova G. TF-IDF-based fake news detection in Kazakh and Russian. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 286–301. DOI <https://doi.org/10.32014/2026.2518-1726.439>

© Марасулов У.А.*, Казбекова Г., 2026.

Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті,
Түркістан, Қазақстан.

E-mail: marasulov.usen2024@ayu.edu.kz

ҚАЗАҚ ЖӘНЕ ОРЫС ТІЛДЕРІНДЕГІ ЖАЛҒАН ЖАҢАЛЫҚТАРДЫ TF-IDF АРҚЫЛЫ АНЫҚТАУ

Марасулов Усен — PhD докторант, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,

E-mail: marasulov.usen2024@ayu.edu.kz, ORCID ID: <https://orcid.org/0009-0008-0801-1229>;

Казбекова Гулнур — техника ғылымдарының кандидаты, қауымдастырылған профессор, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан,
E-mail: gulnur.kazbekova@ayu.edu.kz, ORCID ID: <https://orcid.org/0000-0002-2756-7926>.

Аннотация. Өзектілігі: Цифрлық медиа мен әлеуметтік желілерде тексерілмеген ақпараттың жылдам таралуы жалған жаңалықтарды автоматты түрде анықтауды табиғи тілді өңдеу мен машиналық оқытудың маңызды қолданбалы міндетіне айналдырды. Қазақ тілі үшін бұл мәселе ерекше өзекті, себебі ашық белгіленген корпустар аз, ал дайын модельдер жергілікті медиамәтіндердің тілдік және тақырыптық ерекшеліктеріне толық бейімделмеген. Мақсаты: Зерттеудің мақсаты - қазақ және орыс тілдеріндегі fake/real мәтіндерді жіктеу үшін TF-IDF белгілеріне негізделген машиналық оқыту модельдерін бағалау және кейінгі трансформерлік модельдермен салыстыруға арналған түсіндірілетін baseline қалыптастыру. Әдістері: Экспериментте 1808 мәтіннен тұратын теңгерілген қазақ-орыс корпусы қолданылды: әр тілде 452 fake және 452 real мәтін қамтылды. Logistic Regression, Linear SVM және Complement Naive Bayes модельдері word-level және character-level TF-IDF белгілерімен екітілді, біртілді және кросс-тілдік сценарийлерде салыстырылды. Бағалау accuracy, precision, recall, F1-score және confusion matrix көрсеткіштері арқылы жүргізілді. Негізгі нәтижелер: Екітілді, қазақша және орысша бөлек бағалау режимдерінде ең жақсы модельдер шамамен Macro-F1 = 0.985 нәтижесін көрсетті. Кросс-тілдік бағалау бағытқа тәуелді айырмашылықты анықтады: қазақ тілінен орыс тіліне тасымалдауда Macro-F1 = 0.654, ал орыс тілінен қазақ тіліне тасымалдауда Macro-F1 = 0.926 болды. Практикалық маңызы: Нәтижелер қазақ-орыс медиакөңістігінде жалған жаңалықтарды анықтауға арналған қайталанатын

әрі түсіндірілетін бастапқы өлшем береді және multilingual BERT пен XLM-RoBERTa модельдерімен кейінгі салыстыруларға негіз болады. Бұл өлшем болашақ эксперименттердің сапасын салыстырмалы түрде бағалауға мүмкіндік береді.

Түйін сөздер: жалған жаңалықтар, дезинформация, қазақ тілі, орыс тілі, TF-IDF, машиналық оқыту, кросс-тілдік классификация

© Марасулов У.А.*, Казбекова Г., 2026.

Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан.

E-mail: marasulov.usen2024@ayu.edu.kz

ВЫЯВЛЕНИЕ ЛОЖНЫХ НОВОСТЕЙ НА КАЗАХСКОМ И РУССКОМ ЯЗЫКАХ TF-IDF-МОДЕЛЯМИ

Марасулов Усен — PhD-докторант, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,

E-mail: marasulov.usen2024@ayu.edu.kz, <https://orcid.org/0009-0008-0801-1229>;

Казбекова Гулнур — кандидат технических наук, ассоциированный профессор, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан,

E-mail: gulnur.kazbekova@ayu.edu.kz, <https://orcid.org/0000-0002-2756-7926>.

Аннотация. *Актуальность.* Быстрое распространение непроверенной информации в цифровых медиа и социальных сетях делает автоматическое выявление ложных новостей значимой задачей обработки естественного языка и машинного обучения. Для казахского языка данная проблема особенно актуальна, поскольку открытых размеченных корпусов немного, а готовые модели недостаточно адаптированы к локальному медиадискурсу. *Цель.* Оценить модели машинного обучения на основе TF-IDF для классификации fake/real текстов на казахском и русском языках, а также сформировать объяснимый baseline для последующего сравнения с трансформерными моделями. *Методы.* В эксперименте использован сбалансированный казахско-русский корпус из 1808 текстов: по 452 fake и 452 real текста на каждом языке. Модели Logistic Regression, Linear SVM и Complement Naive Bayes сравнивались с word-level и character-level TF-IDF-признаками в билингвальном, одноязычных и кросс-языковых сценариях. Качество классификации оценивалось по показателям accuracy, precision, recall, F1-score и confusion matrix, что позволило учесть не только общий результат, но и направление ошибок. *Результаты и выводы.* В билингвальном, казахском и русском режимах лучшие модели достигли значения Macro-F1 около 0,985. Кросс-языковая оценка выявила асимметрию переноса: обучение на казахском языке и тестирование на русском дало Macro-F1 = 0,654, тогда как обучение на русском языке и тестирование на казахском достигло Macro-F1 = 0,926. Полученные результаты формируют воспроизводимый и

интерпретируемый ориентир для выявления ложных новостей в казахско-русском медиапространстве и создают основу для дальнейшего сравнения с multilingual BERT и XLM-RoBERTa. Практическая значимость исследования заключается в возможности использования предложенного подхода как базовой модели для мониторинга дезинформации, анализа медиаконтента и разработки интеллектуальных инструментов проверки достоверности текстов на казахском и русском языках.

Ключевые слова: ложные новости, дезинформация, казахский язык, русский язык, TF-IDF, машинное обучение, кросс-языковая классификация, fake news, Macro-F1

Introduction. Online media and social platforms have made the movement of news almost immediate. This speed improves access to information, yet it also helps unverified, distorted or intentionally misleading texts circulate widely. Such content may influence public trust, political and economic choices, and health-related decisions. For that reason, automatic fake news detection is now an applied research task with direct relevance for natural language processing, machine learning and information security. Existing studies mainly address high-resource languages, especially English, where labeled corpora, fact-checking platforms, pretrained models and benchmark datasets are more accessible. Kazakh differs from this setting. Open labeled data are limited, the language has rich morphology, and local misinformation practices are not fully reflected in English-language resources. This gap is consistent with reviews that discuss dataset quality, model limitations and the lack of multilingual and cross-lingual datasets (Harris et al., 2024: 222; Alghamdi et al., 2026: 353), with work on multi-level Kazakh-Russian annotation (Sambetbayeva et al., 2025: 215), and with an early benchmark for Kazakh fake news detection (Telman et al., 2026: 708).

Kazakhstan's information environment is bilingual: Kazakh and Russian are used side by side in news production and consumption. Therefore, fake news detection should not be treated only as a single-language task. It is important to test whether one model can handle both languages in a shared setting and whether signals learned from one language remain useful for fake/real classification in the other. For low-resource languages, such testing is especially valuable because it shows how far a better represented language can support classification when labeled data are scarce.

The purpose of this paper is to estimate the initial capacity of classical machine learning models for Kazakh-Russian fake/real classification and to report the result as an explainable baseline. Two research questions guide the study: how effective are TF-IDF features with linear classifiers on a bilingual Kazakh-Russian corpus, and can a model trained in one language distinguish fake and real texts in another language? The aim is not to present a completed fact-checking product, but to define a transparent reference point for later comparison with transformer-based approaches.

The contribution of the study is fourfold. It uses a balanced experimental dataset containing fake and real texts in Kazakh and Russian. It compares several TF-IDF baselines under the same protocol in monolingual, bilingual and cross-lingual regimes. It also shows that transfer between the two languages is directional: Russian-to-Kazakh transfer performs better than Kazakh-to-Russian transfer. Finally, the interpretation of results explicitly takes into account source bias, topic bias, temporal bias and near-duplicate risks.

Literature review. Automatic fake news detection is not limited to a standard text classification problem. News content, user reactions, propagation networks and source credibility may all affect the decision boundary. Shu et al. (2017: 22) draw attention to news content, user engagement and social context, while Zhou and Zafarani (2020: 109) describe fake news signals through writing style, propagation patterns, false knowledge and source credibility. Benchmark research also shows that model performance depends on the dataset, label quality and evaluation scenario, not only on the selected algorithm (Galli et al., 2022: 237).

Work on the diffusion of misinformation confirms the practical importance of this field. Vosoughi, Roy and Aral (2018: 1146) found on Twitter data that false news may spread faster and more widely than true news. From this perspective, detection models are connected not only with technical classification, but also with information security, media trust and public resilience.

The quality of the dataset is one of the main conditions for reliable evaluation. LIAR supported the analysis of short political statements through multi-level truthfulness labels (Wang, 2017: 422). FEVER connected claim verification with evidence extraction (Thorne et al., 2018: 809), and FakeNewsNet supplemented news texts with social context and spatiotemporal information (Shu et al., 2020: 171). X-Fact extended this direction to multilingual fact-checking (Gupta and Srikumar, 2021: 675). These resources demonstrate that annotation design and corpus structure determine how model scores should be understood.

A separate issue is that high scores may reflect stylistic or source-related differences instead of factual reasoning. Hamed et al. (2023: e20382) connect fake news detection quality with dataset size, label quality, feature representation and data fusion. Thibault et al. (2025: 5801) show that spurious correlations and label noise can reduce generalization. For this reason, the TF-IDF results in the present study are interpreted as evidence of distinguishable signals inside the corpus, not as evidence that factual verification has been solved.

In multilingual settings, evidence-based verification becomes even more important. Dementieva and Panchenko (2021: 310) showed that evidence in another language may improve monolingual fake news detection. The Multiverse study continues this idea by showing that cross-language comparison can serve as an additional explainable signal for classification (Dementieva et al., 2023: 77). This direction fits the Kazakhstani media context, where the same event may be reported in Kazakh and Russian with different wording, source choices and emphases.

The interpretation of evaluation results also depends on data completeness and

cleanliness. Galli et al. compared traditional machine learning and deep learning approaches in a fake news benchmark (Galli et al., 2022: 237). FakeNewsNet combines several information dimensions (Shu et al., 2020: 171), whereas Thibault et al. emphasize the control of label quality, leakage and spurious correlations (Thibault et al., 2025: 5801). Accordingly, the metrics reported in this paper should be treated as baseline measurements that require further checking with source-based, topic-based and temporal splits.

Low-resource and multilingual conditions make these problems harder. Recent reviews name dataset bias, model constraints, limited multilingual data and weak cross-lingual generalization among the key obstacles for fake news detection systems (Harris et al., 2024: 222; Alghamdi et al., 2026: 353). Cross-lingual and cross-domain transfer learning points to the possibility of transferring knowledge from high-resource to lower-resource settings (Providel et al., 2025: 287). De et al. also report the usefulness of multilingual BERT for low-resource fake news classification (De et al., 2021: 9). These findings support the need for carefully designed corpora and evaluation scenarios for Kazakh.

Multilingual transformer models provide an important comparison line. BERT introduced contextual text encoding that improved many NLP tasks (Devlin et al., 2019: 4171), while XLM-RoBERTa achieved strong cross-lingual transfer through large-scale multilingual pretraining (Conneau et al., 2020: 8440). For Kazakh and Russian, Sambetbayeva et al. propose moving beyond a binary fake/real label toward multi-level annotation, including CLAIM, SOURCE, EVIDENCE, DISINFORMATION_TECHNIQUE, AUTHOR_INTENT and TARGET_AUDIENCE (Sambetbayeva et al., 2025: 215). Telman et al. compared TF-IDF, a translation-based cross-lingual approach and XLM-RoBERTa for Kazakh fake news detection (Telman et al., 2026: 708). The present article therefore defines the TF-IDF baseline rather than replacing transformer models.

Overall, the literature shows that a reproducible and transparent baseline for Kazakh-Russian fake/real classification is still necessary. More complex deep learning models can be judged fairly only after the initial level of data quality, source bias and cross-lingual transfer has been measured. This study addresses that preliminary step.

Materials and main methods. Dataset. The experiments were conducted on a bilingual corpus of fake and real texts in Kazakh and Russian. The texts were collected from open sources, cleaned and deduplicated. After deduplication, the complete dataset contained 2,740 records: 452 fake and 518 real Kazakh texts, and 895 fake and 875 real Russian texts. The sources included fact-checking resources, materials documenting misinformation and mainstream news portals. The source_ file field was kept for internal control but was not used as an input feature.

To limit the influence of language and class imbalance, the experimental subset was balanced by selecting 452 texts for each language-class pair. As a result, the final corpus used in the experiments contained 1,808 texts. This design makes model comparison clearer, but it does not reproduce the real distribution of fake

and real content in the media environment. Therefore, the reported values describe performance under controlled laboratory conditions.

Table 1 – Balanced dataset composition by language and class

Language	Fake texts	Real texts	Total
Kazakh	452	452	904
Russian	452	452	904
Total	904	904	1808

Note: Compiled by the authors based on the balanced experimental dataset.

The Fake class was formed from fact-checking and misinformation-related sources, including `factcheck_kz_zhalgan`, `gov_factcheck_fake_claims`, `nofake_new_wp_fake_kk_ru` and `provereno_media_fake_ru`. The Real class was compiled from mainstream news sources such as Egemen, Kazinform, Informburo, Tengrinews, Zakon and user-provided Kazakh news. In this study, the real label indicates the origin of the text from a trusted news source; it does not mean that each item was independently fact-checked. Thus, the experiment measures textual fake/real discrimination rather than absolute truth verification. Model input combined the title, claim and main text fields. Metadata such as ID, `source_file`, URL and date were not provided to the classifier. Empty texts, invalid labels and records with incorrect language values were removed before training. This choice reduces direct source memorization, although indirect source cues may still appear through style, topic, text structure or portal-specific wording.

Deduplication was based on the title, claim and main text fields. Exact duplicate entries were removed, and records with conflicting labels were excluded from the experiment. This step lowers the risk of train/test leakage. Still, lightly edited texts or paraphrases describing the same event may remain in the data; therefore, external testing is needed before extending the conclusions to the wider information space. The two classes were also considered from the viewpoint of label origin. Fake texts often come from refutation or fact-checking-style materials, where denial, explanation and rejection of a source can be frequent. Real texts more often follow editorial news style, official information patterns and portal-specific formatting. These differences provide useful classification signals, but they may also lead the model to learn genre and source features rather than factual truth itself.

Experimental scenarios

Five experimental scenarios were used to evaluate the models.

Table 2 – Description of experimental scenarios

Scenario	Description
bilingual	Joint training and testing on Kazakh and Russian texts
kk_only	Training and testing only on Kazakh texts
ru_only	Training and testing only on Russian texts
cross_kk_to_ru	Training on the Kazakh train split and testing on all Russian texts
cross_ru_to_kk	Training on the Russian train split and testing on all Kazakh texts

Note: Compiled by the authors based on the experimental protocol.

In the bilingual scenario, the train, validation and test split was stratified by both language and label. In `kk_only` and `ru_only`, stratification was based on the label. The approximate proportions were 70% for training, 15% for validation and 15% for testing. In cross-lingual scenarios, the model was trained on the training part of one language and evaluated on all 904 texts of the other language. Random seed = 42 was fixed throughout. Cleaning and balancing were completed before splitting; the validation part was retained for protocol consistency rather than broad hyperparameter search.

Features and models

TF-IDF was used to represent each text as a vector. Two feature configurations were tested. The first used word-level unigrams and bigrams; the second used character-level 3-5 n-grams. Character n-grams are relevant for Kazakh and Russian because they can capture inflectional endings, suffixes, recurring short patterns and orthographic similarity.

The comparison included five baseline configurations.

Table 3 – Compared TF-IDF baseline models

Model name	Feature type	Classifier
<code>word_tfidf_logreg</code>	Word TF-IDF, 1-2 n-grams	Logistic Regression
<code>char_tfidf_logreg</code>	Character TF-IDF, 3-5 n-grams	Logistic Regression
<code>word_tfidf_svm</code>	Word TF-IDF, 1-2 n-grams	Linear SVM
<code>char_tfidf_svm</code>	Character TF-IDF, 3-5 n-grams	Linear SVM
<code>word_tfidf_cnb</code>	Word TF-IDF, 1-2 n-grams	Complement Naive Bayes
Note: Compiled by the authors based on the experimental protocol.		

For Logistic Regression and Linear SVM, `class_weight = balanced` was applied. The word-level TF-IDF vocabulary was limited to 100,000 features, and the character-level vocabulary to 120,000 features. The minimum document frequency was `min_df = 2`; `sublinear_tf` was applied for Logistic Regression and Linear SVM. All experiments were implemented in scikit-learn. The vectorizer and classifier were fitted inside a single Pipeline, so the vocabulary and TF-IDF weights were learned only from the training split; test texts were not involved in feature construction.

TF-IDF was chosen as a baseline because it is simple, fast and relatively interpretable, which is useful in a low-resource setting. At the same time, it does not place Kazakh and Russian words into a shared semantic space. For that reason, cross-lingual performance may depend on shared topics, style, symbols or class-specific lexical patterns rather than on semantic understanding. This limitation is central to the interpretation of the experiment.

Evaluation metrics

Model quality was assessed with Accuracy, Precision, Recall and F1-score. Macro-F1 was used as the main comparison metric because it gives the fake and real classes equal weight. This is important in cross-lingual evaluation, where a model may overpredict one class while Accuracy hides the imbalance. Confusion matrices were added to show the direction and concentration of errors.

Reproducibility and bias control

Reproducibility was supported by fixing seed = 42 for data splitting and stochastic model components. All models were tested under the same scenarios and with the same textual fields. This design improves comparability, although it cannot fully remove topic and source differences that may exist within the corpus.

Three types of bias were considered during interpretation. Source bias may occur because fake and real texts come from different categories of websites. Topic bias appears when the two classes cover different subject areas. Temporal bias can inflate expected generalization when training and test texts are close in time. For these reasons, the high scores are reported as controlled baseline results, not as production-level accuracy.

Results

Table 4 summarizes the aggregated results. Figures 1-3 illustrate the scenario-level Macro-F1 and Accuracy distributions, and Figure 4 shows class-level F1 values. The best performance was observed in monolingual and bilingual regimes. In the bilingual scenario, word_tfidf_svm and char_tfidf_svm produced the same result: Accuracy = 0.9853 and Macro-F1 = 0.9853. In the Kazakh-only setting, word_tfidf_logreg and word_tfidf_svm reached Macro-F1 = 0.9853. In the Russian-only setting, word_tfidf_svm produced the highest score.

Table 4 – Best results by experimental scenario

Scenario	Best model	Accuracy	Macro-F1
bilingual	word_tfidf_svm	0.9853	0.9853
kk_only	word_tfidf_logreg	0.9853	0.9853
ru_only	word_tfidf_svm	0.9853	0.9853
cross_kk_to_ru	word_tfidf_cnb	0.6869	0.6540
cross_ru_to_kk	word_tfidf_cnb	0.9259	0.9257

Note: Compiled by the authors based on the experimental results.

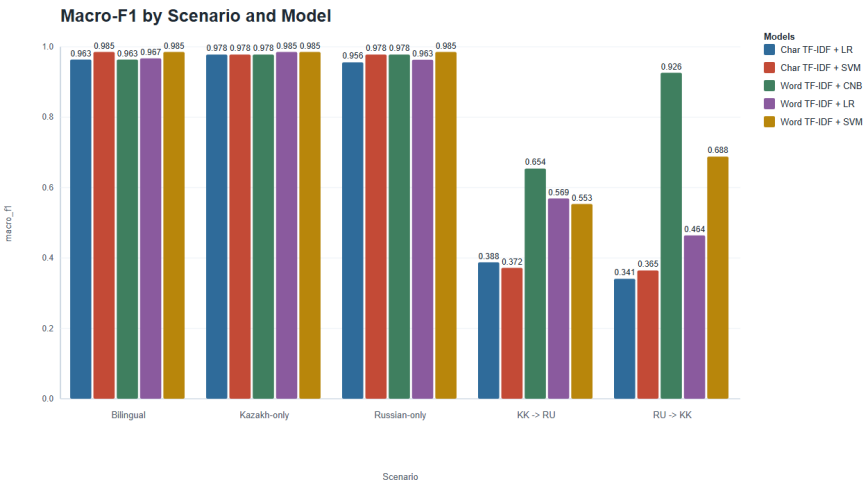


Figure 1 – Macro-F1 scores by experimental scenario

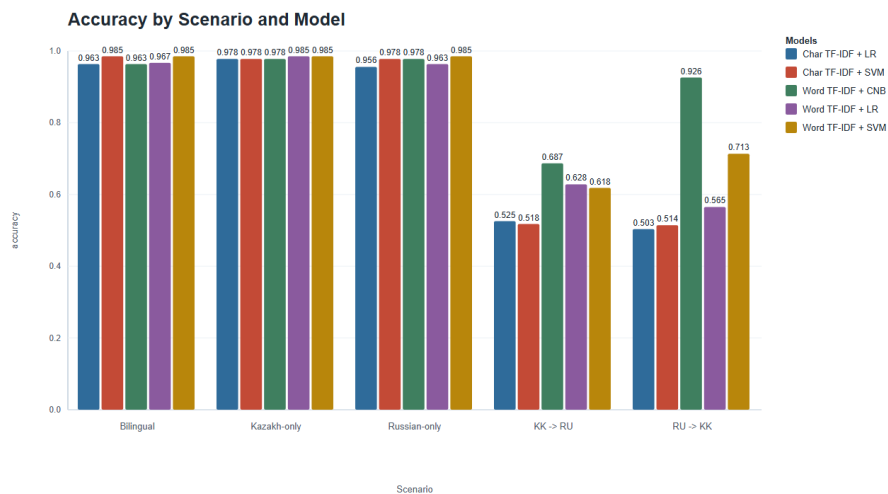


Figure 2 – Accuracy scores by experimental scenario

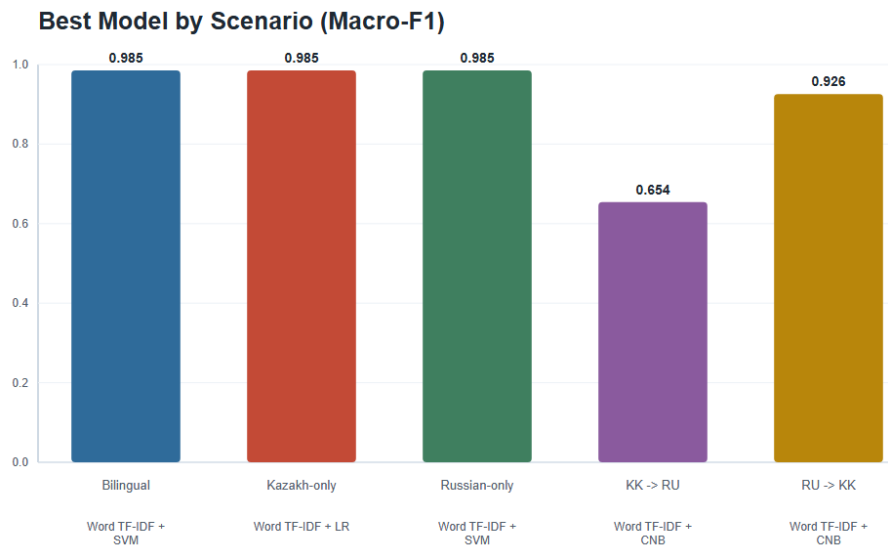


Figure 3 – Macro-F1 scores of the best models in each scenario

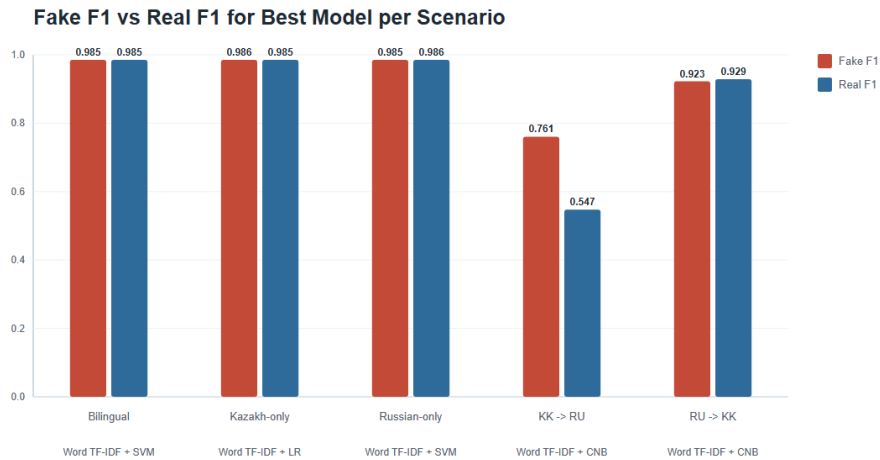


Figure 4 – F1 scores for fake and real classes in the best models

Figure 5 presents the confusion matrix for the bilingual scenario. The best model correctly classified 268 of 272 test texts. In the Fake class, 134 texts were detected correctly and 2 were assigned to the Real class. In the Real class, 134 texts were detected correctly and 2 were assigned to the Fake class. The errors are balanced across the two classes, so no clear class bias is observed in this scenario.

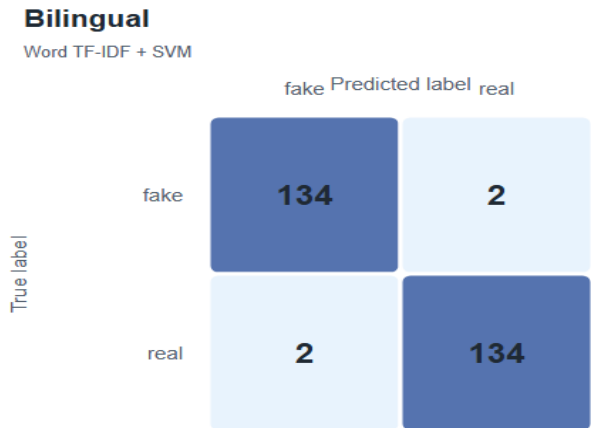


Figure 5 – Confusion matrix of the best model in the bilingual scenario

In the Kazakh-only scenario, the model classified 134 of 136 test texts correctly. All fake texts were detected, while 2 real texts were classified as fake. In the Russian-only scenario, the total number of correct predictions was also 134 of 136, but the error pattern was different: all real texts were retained, and 2 fake texts were classified as real.

Cross-lingual performance was strongly directional. In `cross_kk_to_ru`, the model was trained on the Kazakh training split and tested on Russian texts. The

best result was obtained by word_tfidf_cnb, with Accuracy = 0.6869 and Macro-F1 = 0.6540. The confusion matrix indicates that fake Russian texts were detected well, but real Russian texts were often moved to the fake class: 450 of 452 fake texts were correct, whereas 281 of 452 real texts were misclassified as fake. Figure 6 shows the corresponding matrix.

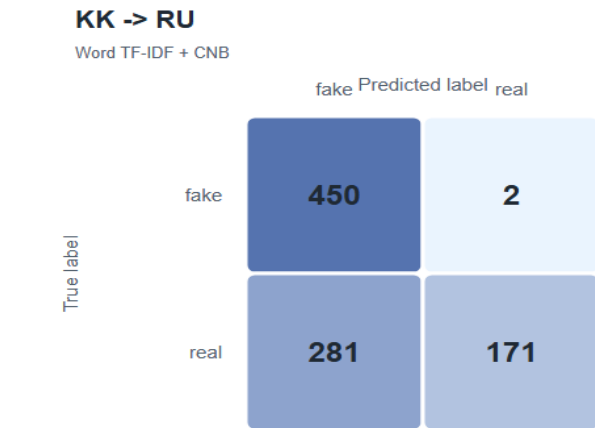


Figure 6 – Confusion matrix of the model trained on Kazakh and tested on Russian

The reverse direction produced a more stable result. In cross_ru_to_kk, the model was trained on the Russian training split and tested on Kazakh texts. The best model was again word_tfidf_cnb, reaching Accuracy = 0.9259 and Macro-F1 = 0.9257. In total, 837 of 904 test texts were classified correctly; this includes 399 of 452 fake Kazakh texts and 438 of 452 real Kazakh texts. Figure 7 presents the corresponding confusion matrix.

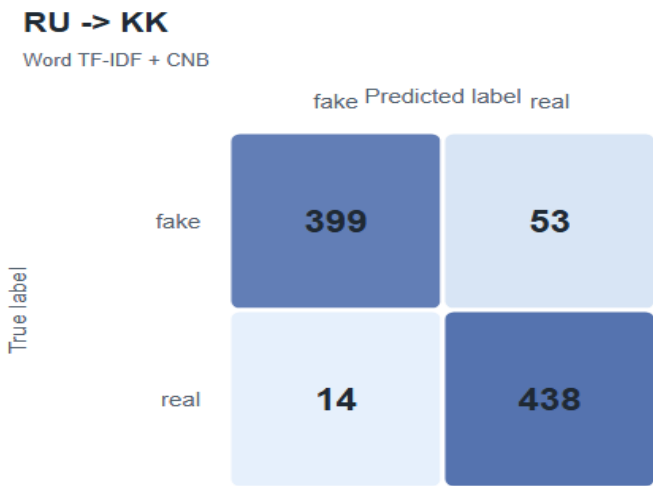


Figure 7 – Confusion matrix of the model trained on Russian and tested on Kazakh

Error analysis

The confusion matrices show where the main errors occur. In monolingual and bilingual regimes, the number of errors is small and the distribution between classes is almost even. In cross_kk_to_ru, the main weakness is the assignment of real Russian texts to the fake class: 281 of 452 real texts were classified as fake. This may mean that fake-related signals learned from Kazakh data overlap with stylistic or topical patterns in real Russian news.

In cross_ru_to_kk, the error distribution is less extreme: 399 of 452 fake texts and 438 of 452 real texts were recognized correctly. One possible explanation is that the Russian training set covers broader fake/real distinctions that are also present in the Kazakh test set. This explanation remains provisional and should be checked with source-based splitting, topic-controlled splitting and manual analysis of misclassified texts.

Discussion. The results indicate that classical TF-IDF models form a strong baseline for fake/real classification in this corpus. Macro-F1 values close to 0.985 in monolingual and bilingual scenarios show that the Kazakh and Russian texts contain clear class-related lexical signals. These signals may come from topic selection, writing style, source format, claim structure or wording typical of fact-checking materials. Thus, the high scores should be understood as evidence of separable textual patterns, not as proof of deep factual understanding.

The bilingual result suggests that Kazakh and Russian texts can be handled within one classification space. This is relevant for Kazakhstan's bilingual information environment. Nevertheless, TF-IDF does not encode contextual meaning; it relies on word and n-gram distributions. Therefore, the result shows that the dataset contains shared lexical and stylistic cues, but it does not prove that the model captured the full semantics of both languages.

The main analytical result is the asymmetry of cross-lingual transfer. Training on Kazakh led to weaker performance on Russian texts, while training on Russian transferred more successfully to Kazakh texts. Several explanations are possible: the Russian training set may be wider in topic and style, the source structure of the Kazakh data may differ from the Russian test set, or Complement Naive Bayes may preserve some lexical likelihood patterns better in cross-lingual settings. These are hypotheses for future ablation and source/topic analysis, not confirmed causal explanations.

This asymmetry requires cautious interpretation. TF-IDF does not build a shared semantic space for Kazakh and Russian, nor does it identify translation equivalents. It works with symbolic, stylistic and topical similarities present in the data. Therefore, the strong cross_ru_to_kk result is best understood as a combination of shared dataset signals and statistical adaptation rather than genuine semantic transfer.

The practical implication has two parts. First, TF-IDF should remain in future experiments as a simple, fast and interpretable comparison point for Kazakh-Russian fake/real classification. Second, multilingual transformer models,

including multilingual BERT and XLM-RoBERTa, should be evaluated under the same protocol to test whether contextual embeddings provide stronger semantic generalization across languages.

Limitations and threats to validity

Several limitations affect the interpretation of the results. First, the Real class was formed from mainstream news portals and not from independently fact-checked texts. Therefore, the model may learn source and style differences rather than truth and falsity. Second, although the corpus is balanced, its size is limited: each language-class pair contains 452 texts, which may not cover the full diversity of topics. Third, structural differences between fact-checking websites and news portals may influence classification quality.

Fourth, the study uses only classical TF-IDF-based models. This is appropriate for the baseline objective, but it does not replace direct comparison with multilingual transformer models. Fifth, temporal and source-based splits were not tested separately; if training and test data are close in time or source structure, future performance may be lower. Sixth, exact duplicates were removed, but near-duplicates or event-level overlap may remain. Seventh, feature weights, important n-grams and SHAP-level explainability were not analyzed in this version.

Despite these limitations, the study offers an initial experimental basis for Kazakh-Russian fake/real classification. Its value is not in presenting a production detector, but in showing the level reached by an explainable baseline, revealing cross-lingual asymmetry and identifying the next steps for future research. The results can support corpus expansion, multi-level annotation and comparison with transformer-based models.

Conclusion. This paper evaluated TF-IDF-based machine learning models for fake/real text classification in Kazakh and Russian across five experimental scenarios. The balanced bilingual corpus contained 1,808 texts with equal language and class proportions. Logistic Regression, Linear SVM and Complement Naive Bayes were compared under a single evaluation protocol, and the results were interpreted as an explainable baseline.

In monolingual and bilingual regimes, the models achieved strong performance: Macro-F1 was approximately 0.985 in the bilingual, *kk_only* and *ru_only* scenarios. This indicates that TF-IDF-based linear models are a strong baseline for Kazakh-Russian fake news detection. At the same time, these scores do not mean that factual verification is solved, because the model may rely on style, topic and source signals. Cross-lingual evaluation showed a directional gap: Kazakh-to-Russian transfer reached Macro-F1 = 0.654, whereas Russian-to-Kazakh transfer reached Macro-F1 = 0.926.

Future work should expand the corpus, apply source-based and temporal splitting, and improve control over near-duplicate and event-level overlap. Multilingual transformer models such as XLM-RoBERTa and multilingual BERT should be compared with the TF-IDF baseline under the same protocol. The binary fake/real label should also be extended with multi-level annotation, including disinformation

technique, author intent, target audience and evidence, to improve the explanatory and applied value of the research.

Conflict of interests

The authors declare no conflict of interest.

Declaration on the use of generative AI tools

Generative AI tools were used only for language editing, readability improvement and technical formatting support under full author supervision. The tools were not used to generate the scientific content, methodology, dataset, experimental results, analysis or conclusions. All scientific content, references, data descriptions, results and conclusions were critically reviewed, verified and approved by the authors.

Acknowledgement and funding

The study was conducted without external funding. The authors have no additional acknowledgements.

References

- Alghamdi J., Lin Y., & Luo S. (2026) Machine learning and deep learning approaches for fake news detection and related topics in multilingual contexts: A systematic literature review. *Multimedia Tools and Applications*, 85, Article 353. <https://doi.org/10.1007/s11042-026-21238-1> (in Eng.)
- Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzman F., Grave E., Ott M., Zettlemoyer L., & Stoyanov V. (2020) Unsupervised cross-lingual representation learning at scale. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. — P. 8440-8451. <https://doi.org/10.18653/v1/2020.acl-main.747> (in Eng.)
- De, A., Bandyopadhyay D., Gain B., & Ekbal A. (2021) A transformer-based approach to multilingual fake news detection in low-resource languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 21(1), Article 9. <https://doi.org/10.1145/3472619> (in Eng.)
- Dementieva D., & Panchenko A. (2021) Cross-lingual evidence improves monolingual fake news detection. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop*. — P. 310-320. <https://doi.org/10.18653/v1/2021.acl-srw.32> (in Eng.)
- Dementieva D., Kuimov M., & Panchenko A. (2023) Multiverse: Multilingual evidence for fake news detection. *Journal of Imaging*, 9(4), Article 77. <https://doi.org/10.3390/jimaging9040077> (in Eng.)
- Devlin J., Chang M.-W., Lee K., & Toutanova K. (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*. — P. 4171-4186. <https://doi.org/10.18653/v1/N19-1423> (in Eng.)
- Galli A., Masciari E., Moscato V., & Sperli G. (2022) A comprehensive benchmark for fake news detection. *Journal of Intelligent Information Systems*, 59. — P. 237-261. <https://doi.org/10.1007/s10844-021-00646-9> (in Eng.)
- Gupta A., & Srikumar V. (2021) X-Fact: A new benchmark dataset for multilingual fact checking. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. — P. 675-682. <https://doi.org/10.18653/v1/2021.acl-short.86> (in Eng.)
- Hamed S.K., Ab Aziz M.J., & Yaakub M.R. (2023) A review of fake news detection approaches: A critical analysis of relevant studies and highlighting key challenges associated with the dataset, feature representation, and data fusion. *Heliyon*, 9(10), Article e20382. <https://doi.org/10.1016/j.heliyon.2023.e20382> (in Eng.)
- Harris S., Hadi H.J., Ahmad N., & Alshara M.A. (2024) Fake news detection revisited: An extensive review of theoretical frameworks, dataset assessments, model constraints, and forward-looking

research agendas. *Technologies*, 12(11), Article 222. <https://doi.org/10.3390/technologies12110222> (in Eng.)

Providel E., Mendoza M., & Solar M. (2025) Cross-lingual cross-domain transfer learning for rumor detection. *Future Internet*, 17(7), Article 287. <https://doi.org/10.3390/fi17070287> (in Eng.)

Sambetbayeva M., Nekessova A., Yerimbetova A., Bayangali A., Kaldarova M., Telman D., & Smailov N. (2025) A multi-level annotation model for fake news detection: Implementing Kazakh-Russian corpus via Label Studio. *Big Data and Cognitive Computing*, 9(8), Article 215. <https://doi.org/10.3390/bdcc9080215> (in Eng.)

Shu K., Mahudeswaran D., Wang S., Lee D., & Liu H. (2020) FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big Data*, 8(3). — P. 171-188. <https://doi.org/10.1089/big.2020.0062> (in Eng.)

Shu K., Sliva A., Wang S., Tang J., & Liu H. (2017) Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1). — P. 22-36. <https://doi.org/10.1145/3137597.3137600> (in Eng.)

Telman D., Yerimbetova A., Sambetbayeva M., & Bolatov B. (2026) Cross-lingual and multilingual approaches to fake news detection in the Kazakh language. *Procedia Computer Science*, 275. — P. 708-715. <https://doi.org/10.1016/j.procs.2026.01.082> (in Eng.)

Thibault C., Tian J.-J., Peloquin-Skulski G., Curtis T.L., Zhou J., Laflamme F., Guan Y., Rabbany R., Godbout J.-F., & Pelrine K. (2025) A guide to misinformation detection data and evaluation. *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. — P. 5801-5809. <https://doi.org/10.1145/3711896.3737437> (in Eng.)

Thorne J., Vlachos A., Christodoulopoulos C., & Mittal A. (2018) FEVER: A large-scale dataset for fact extraction and verification. *Proceedings of NAACL-HLT 2018*. — P. 809-819. <https://doi.org/10.18653/v1/N18-1074> (in Eng.)

Vosoughi S., Roy D., & Aral S. (2018) The spread of true and false news online. *Science*, 359(6380). — P. 1146-1151. <https://doi.org/10.1126/science.aap9559> (in Eng.)

Wang W.Y. (2017) Liar, liar pants on fire: A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. — P. 422-426. <https://doi.org/10.18653/v1/P17-2067> (in Eng.)

Zhou X., & Zafarani R. (2020) A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53(5), Article 109. <https://doi.org/10.1145/3395046> (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 302–316

<https://doi.org/10.32014/2026.2518-1726.440>

IRSTI 05.13.17

UDC 004.912:004.056

© **Omar A.B. *, Mussiraliyeva Sh.Zh., 2026.**

Al-Farabi Kazakh National University, Almaty, Kazakhstan.

E-mail: aiym.omar98@gmail.com

**FEDERATED LEARNING: MODELS BASED
ON TRANSFORMER ARCHITECTURE****Omar Aiym** — PhD student, the Department of Cybersecurity and Cryptology, Al-Farabi Kazakh National University, Almaty, Kazakhstan,E-mail: aiym.omar98@gmail.com, <https://orcid.org/0000-0003-2644-5596>;**Mussiraliyeva Shynar** — Candidate of Physical and Mathematical Sciences, Professor of the Department of Cybersecurity and Cryptology, Al-Farabi Kazakh National University, Almaty, Kazakhstan,E-mail: mussiraliyevash@gmail.com, <https://orcid.org/0000-0001-5794-3649>.

Abstract. The relevance of this study is related to ensuring secure and privacy-preserving model training without transferring personal data to a central server. One of the effective approaches to this problem is federated learning, where data remain locally on client computers, and only model updates are sent to the server. This helps reduce the risk of personal data leakage. The main objective of the study is to detect and classify aggressive content using federated learning and to evaluate the performance of Transformer-based models, namely BERT, ALBERT, and DistilBERT. The study includes several experimental stages, and the models are compared using accuracy metrics and confusion matrices. In the experiment, the English Profanity Words dataset was used to train the selected models. The dataset was distributed among 3, 5, 7, and 11 clients. Two data distribution scenarios were considered: IID, where the dataset is distributed evenly among clients, and Non-IID, where the data are distributed unevenly. This made it possible to analyze how the number of clients and data distribution affect model performance in a federated learning environment. The experimental results showed that ALBERT achieved the best performance, with accuracy ranging from 93% to 97% as the number of clients increased. BERT also demonstrated stable results, while DistilBERT had higher training speed but lower accuracy compared to the other models. The obtained results can help select an appropriate model for privacy-sensitive text classification tasks. In practice, this approach can be applied to text

analysis, aggressive content detection, and systems where data security is a key requirement. Future work will focus on increasing the number of clients, applying differential privacy, and conducting experiments on other datasets.

Keywords: natural language processing, federated learning, transformer models, uniform dataset, non-uniform dataset, number of clients, training speed

For citations: Omar A.B., Mussiraliyeva Sh.Zh. Federated learning: models based on transformer architecture. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 302–316. DOI <https://doi.org/10.32014/2026.2518-1726.440>

© **Омар А.Б.*, Мусиралиева Ш.Ж., 2026.**
Әл-Фараби атындағы ҚазҰУ, Алматы, Қазақстан.
E-mail: aiym.omar98@gmail.com

ФЕДЕРАТИВТІ ОҚЫТУ: ТРАНСФОРМЕР АРХИТЕКТУРАСЫНА НЕГІЗДЕЛГЕН МОДЕЛЬДЕР

Омар Айым — докторант, әл-Фараби атындағы Қазақ Ұлттық Университетінің "Киберқауіпсіздік және криптология" кафедрасы, Алматы, Қазақстан,
E-mail: aiym.omar98@gmail.com, <https://orcid.org/0000-0003-2644-5596>;

Мусиралиева Шынар — физика-математика ғылымдарының кандидаты, профессор, әл-Фараби атындағы Қазақ Ұлттық Университетінің "Киберқауіпсіздік және криптология" кафедрасы, Алматы, Қазақстан,
E-mail: mussiraliyevash@gmail.com, <https://orcid.org/0000-0001-5794-3649>.

Аннотация. Зерттеудің өзектілігі жеке деректерді орталық серверге жібермей модельді қауіпсіз құпия түрде оқытуды қамтамасыз етумен байланысты. Бұл мәселені шешудің тиімді әдістерінің бірі – федеративті оқыту. Оның ерекшелігі: деректер клиенттік компьютерлерде жергілікті түрде сақталады, серверге тек модель жаңартулары жіберіледі. Нәтижесінде жеке деректердің құпиялығы қамтамасыз етіледі. Зерттеудің негізгі мақсаты федеративті оқыту әдісін қолдану арқылы агрессивті мазмұнды анықтау және классификациялау. BERT, ALBERT және DistilBERT сияқты трансформерлік модельдердің өнімділігі мен тиімділігін бағалау. Зерттеу бірнеше кезендерді қарастырады. Модель нәтижелері дәлдік көрсеткіштері мен шатастыру матрицалары арқылы салыстырылады. Зерттеу барысында федеративті оқыту әдісі мен BERT, ALBERT және DistilBERT модельдерінің жұмыс істеу өнімділігін қарастырдық. Эксперимент кезінде деректер жиынтығын 3, 5, 7, 11 клиентке бөлу арқылы деректердің екі түрлі таралу әдісін қарастырдық. Біріншісі IID, яғни деректер жиынтығының біркелкі бөлінуіне негізделген әдіс. Екіншісі Non-IID, яғни керсінше деректер жиынтығының біркелкі емес бөлінуін сипаттайтын әдіс. Берілген модельдерді оқыту үшін English Profanity Words деген деректер жиынтығы пайдаланылды. Деректер жиынтығының

сипаттамасы «Әдістер мен материалдар» бөлімінде ұсынылған. Эксперимент кезінде алынған модельдердің дәлдігі және оқу жылдамдығының нәтижелеріне талдау жұмыстары жасалды. Сонымен қатар деректер жиынтығын екі әдіс бойынша бөліп, клиент санын біртіндеп арттырып модель дәлдігінің өзгерістерін зерттедік. Нәтижесінде ALBERT моделі клиент саны артқан кезде жақсы көрсеткіш берді 93–97%. Ал BERT моделі тұрақты нәтижелерге қол жеткізетінін және DistilBERT моделі басқа модельдермен салыстырғанда жоғары оқу жылдамдығына ие болғанымен дәлдігі төмен көрсеткіш берді. Алынған нәтижелер құпия деректермен жұмыс істегенде және федеративті оқыту модельдерін қолданған кезде қолайлы модельді таңдауға көмектеседі. Іс жүзінде бұл әдісті мәтінді талдау, агрессивті мазмұнды анықтау, деректер қауіпсіздігі маңызды жүйелерде қолдануға болады. Болашақта клиенттер санын көбейту, дифференциалды құпиялылық әдісін енгізу және басқада деректер жиынтығында эксперименттер жүргізу жоспарлануда.

Түйін сөздер: Табиғи тілдерді өңдеу, федеративті оқыту, трансформер үлгілері, біркелкі деректер жиынтығы, біркелкі емес деректер жиынтығы, клиенттер саны, оқыту жылдамдығы

© Омар А.Б. *, Мусиралиева Ш.Ж., 2026.

Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан.

E-mail: aiym.omar98@gmail.com

ФЕДЕРАТИВНОЕ ОБУЧЕНИЕ: МОДЕЛИ НА ОСНОВЕ АРХИТЕКТУРЫ ТРАНСФОРМЕРОВ

Омар Айым — докторант кафедры кибербезопасности и криптологии, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,
E-mail: aiym.omar98@gmail.com, <https://orcid.org/0000-0003-2644-5596>;

Мусиралиева Шынар — кандидат физико-математических наук, профессор кафедры кибербезопасности и криптологии, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,
E-mail: mussiraliyevash@gmail.com, <https://orcid.org/0000-0001-5794-3649>.

Аннотация. *Актуальность.* Актуальность исследования обусловлена необходимостью безопасного и конфиденциального обучения моделей машинного обучения без передачи персональных данных на центральный сервер. Одним из эффективных подходов к решению данной задачи является федеративное обучение, при котором данные остаются на локальных клиентских устройствах, а на сервер передаются только обновления параметров модели. Такой подход позволяет повысить уровень защиты персональных данных и одновременно использовать распределенные данные для обучения интеллектуальных систем. *Цель.* Выявить и классифицировать агрессивный контент с применением метода федеративного обучения, а также

оценить производительность и эффективность трансформерных моделей BERT, ALBERT и DistilBERT в условиях различного распределения данных между клиентами. *Методы.* В исследовании использован набор данных English Profanity Words. Экспериментальная часть включала обучение и сравнение моделей BERT, ALBERT и DistilBERT в федеративной среде при различном количестве клиентов: 3, 5, 7 и 11. Рассмотрены два сценария распределения данных: IID, основанный на равномерном распределении набора данных между клиентами, и Non-IID, отражающий неравномерное распределение данных. Для оценки результатов использовались показатели точности, скорости обучения и матрицы ошибок. Дополнительно проанализировано изменение качества моделей при постепенном увеличении числа клиентов. *Результаты и выводы.* В ходе экспериментов установлено, что модель ALBERT показала наиболее высокие результаты при увеличении числа клиентов: точность находилась в диапазоне 93–97%. Модель BERT продемонстрировала стабильные показатели качества в разных сценариях распределения данных. Модель DistilBERT, несмотря на более высокую скорость обучения по сравнению с другими моделями, показала более низкую точность классификации. Полученные результаты позволяют определить наиболее подходящие трансформерные модели для задач анализа текста, выявления агрессивного контента и обработки конфиденциальных данных в условиях федеративного обучения. Практическая значимость исследования заключается в возможности применения предложенного подхода в системах модерации контента, интеллектуального анализа текста и обеспечения безопасности данных. В дальнейших исследованиях планируется расширить число клиентов, внедрить методы дифференциальной конфиденциальности и провести эксперименты с другими наборами данных.

Ключевые слова: обработка естественного языка, федеративное обучение, трансформерные модели, BERT, ALBERT, DistilBERT, IID, Non-IID, количество клиентов, скорость обучения.

Кіріспе. Қазіргі кезде табиғи тілді өңдеу технологиялары (NLP) айтарлықтай жетістіктерге қол жеткізді. Мәселен BERT, ALBERT және DistilBERT сияқты трансформерге негізделген модельдер көп қолданылады. Олар мәтіндерді классификациялау, мәтінге негізделген отырып көңіл-күйді талдауда және өзге NLP тапсырмаларында жақсы көрсеткіштер береді. Алайда мұндай модельдерді оқыту үшін бізге үлкен көлемдегі деректер қажет. Ал еліміздегі құпиялылық пен қауіпсіздік мәселелері барлық деректерді бір орталық серверге жинауға қарсы. Бұл ақпараттың құпиялығына қарама-қайшы іс-әрекет.

Осыған орай федеративті оқыту (FL) бұл мәселенің тиімді шешімі деп ұсынсақ болады. Мәселен федеративті оқыту әдісі деректерді орталық серверге жинаудың орнына модельді оқыту пайдаланушы құрылғыларында жергілікті түрде жүргізіледі (Cozac et al., 2025). Яғни жүйе оқыту процессі

орындалғаннан кейін, серверге тек модель жаңартуларын жібереді. Ал деректердің өзі пайдаланушы құрылғыларында құпия түрде қалады. Нәтижесінде құпиялылық пен қауіпсіздік қасиеттері сақталады. Алайда федеративті оқытудың (FL) өз қиындықтарыда болады. Бірінші мәселе, деректер жиынтығы Non-IID (бірдей бөлінбеген) сипатта болуы мүмкін. Деректер жиынтығының мазмұны және құрылымы біркелкі болмауы оқу процесін қиындатады және модель дәлдігіне әсер етуі мүмкін (Orabi et al., 2025; Kumarappan et al., 2024).

Екінші мәселе, бұл оқыту процесіне қатысатын клиент санының артуы. Яғни клиент саны артқан сайын оқыту процесіне кететін уақытта өседі, бұл уақыт өнімділігіне қарама-қайшы әсерін береді. Тағы бір айта кететін жайт, бұл модель архитектурасын дұрыс таңдай білу мәселесі. Мәселен DistilBERT сияқты модельдерде оқыту процесі тез болғанымен дәлдігі төмен болуы мүмкін. Ал BERT және ALBERT модельдері жақсы нәтиже бергенімен оларды оқыту көбірек есептеу ресурстарын және уақытты қажет етеді.

Әдеби шолу. Қазіргі кезде көптеген ғалымдар федеративті оқыту (ФО) әдісі бойынша белсенді түрде зерттеу жұмыстарын жүргізуде. Федеративті оқыту әдісі қандай ортада жиі қолданылатының түсіну үшін 1 кестеде әдеби шолу ұсынылған.

Кесте 1 – Федеративті оқыту әдісі негізіндегі әдеби шолу

Жылы	Автор(лар)	Жұмыс атауы	Зерттеудің мақсаты
2023	Xin'ao Wang, Huan Li, Ke Chen, Lidan Shou	FedBFPT: An Efficient Federated Learning Framework for BERT Further Pre-training	BERT модельі қолданылды. Негізгі мақсаты есептеу ресурстарының шығындарын азайту болды.
2023	Mohanad Sarhan, Siamak Layeghy, Nour Moustafa, Marius Portmann	Cyber Threat Intelligence Sharing Scheme Based on Federated Learning for Network Intrusion Detection	Желілік шабуылдарды анықтауға негізделген. Яғни бірнеше ұйым арасындағы желілік шабуылдарды анықтауға арналған модель құрды.
2024	Aradhana Sahu, K. Aanandha Saravanan	Federated LSTM Model for Enhanced Anomaly Detection in Cyber Security	Аномалияларды анықтау негізінде жұмыс жасалды. LSTM және федеративті оқыту әдісі қолданылды.
2025	B. Olanrewaju-George, B. Praunggono	Federated learning-based intrusion detection system for IoT using unsupervised and supervised DL models	IoT құрылғыларындағы аномалияларды анықтау. Яғни федеративті оқыту әдіс негізінде аномалия анықтау жүйесін құру.
2024	Arash Heidari, Nima Jafari Navimipour, Hasan Dag, Samira Talebi	A Novel Blockchain-Based Deepfake Detection Method Using Federated and Deep Learning Models	Федеративті оқыту әдісі негізінде дипфейк анықтау жүйесін құру. Сонымен қатар CNN әдісін қолдану.

2024	Tijana Markovic, Miguel Leon, David Buffoni, Sasikumar Punnekkat	Random Forest with Differential Privacy in Federated Learning for Network Attack Detection and Classification	Желілік шабуылдарды анықтау. Қосымша дифференциалды құпиялылық әдісін қолдану.
2017	H. Brendan McMahan, Daniel Ramage, Seth Hampson	Communication-Efficient Learning of Deep Networks from Decentralized Data	Шығындарды азайту мақсатында федеративті оқыту әдісін қолдану.

Федеративті оқыту әдісін алғаш рет клиенттерді оқыту кезіндегі коммуникация тиімділін арттыру мақсатында енгізді (Konečný et al., 2015). Негізгі мақсаты ретінде құпиялықты қамтамасыз ету және оқыту кезінде жеке деректердің ағып кетуіне жол бермеу болды. Енгізудегі қиындықтар: деректер жиынтығының клиенттер арасында біркелкі емес бөлінуі. Бұл модель дәлдігін және сенімділін төмендетті.

Zhao et al. (2018), Hsu et al. (2019) және Listo Zec et al. (2020) FedAvg алгоритімінің және non-IID(біркелкі емес) клиенттік деректерінің федеративті ортадағы әсерін зерттеген. McMahan et al.(2017) және Hard et al.(2018) федеративті ортада дифференциалды құпиялылық әдісін мобильді пернетақта мәтінін болжау тапсырмалары үшін қолданды.

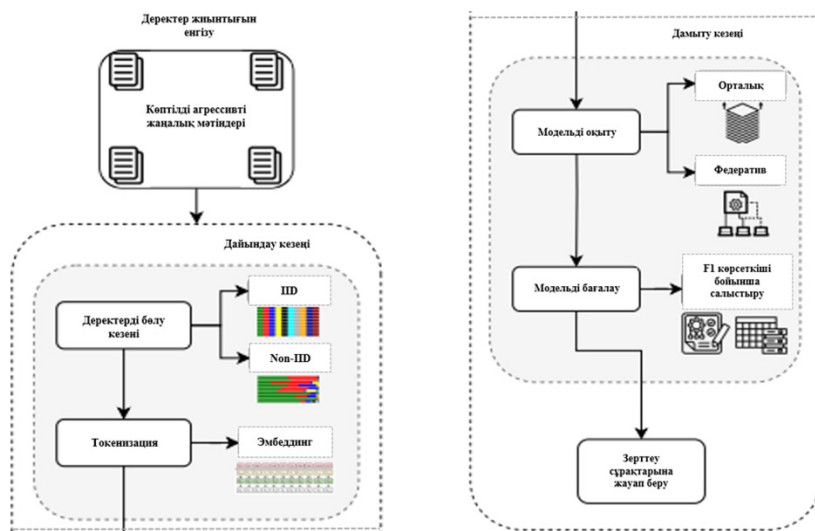
Сонымен қатар зерттеу жұмысы BERT, DistilBERT, ALBERT сияқты трансформер архитектурасына бағытталған. Яғни аталған модельдердің федеративті ортадағы жұмыс істеу тиімділігін анықтауға негізделген.

BERT (Devlin et al., 2019) жасырылған сөзді табуға мүмкіндік беру үшін кіріс токендерінің 15%-ын жасырады. Ол SEP және CLS токендерін қолданады. SEP – екі сөйлемді бөлу үшін, ал CLS – келесі сөйлемді болжау үшін қолданылады.

DistilBERT – BERT-тің оңтайландырылған нұсқасы (Sanh et al., 2020). BERT тиімділігінің шамамен 97%-ын сақтайды. Ол жылдам жұмыс істеуге және аз жақты алуға арналған. BERT-тен басты айырмашылығы екі есе аз қабаттар санын құрайды. Мәселен BERT – 12 қабатты қамтыса, DistilBERT – 6 қабаттан тұрады.

ALBERT (Lan et al., 2020) – BERT-тің тағы бір модификациясы. Алайда ол DistilBERT-ке қарама-қарсы мақсатпен жасалған. Яғни ALBERT қабаттар санын сақтай отырып модельдің көлемін сығу арқылы айтарлықтай азайтады. Ол небәрі 45 МБ құрайды және кез-келген мобильді құрылғыға сыяды. Ол тек жақты үнемдейді, орындалу уақытын емес.

Зерттеу материалдары мен әдістері. Оқыту процесін ұйымдастыру үшін DistilBERT, BERT, ALBERT модельдері қолданылды. Сонымен қатар салыстыру жұмыстары әділ болу үшін барлық модельдер бір деректер жиынтығын қолданады. Модельдерді оқыту 1-суретте көрсетілген архитектура негізінде жүзеге асырылды. Жалпы бұл архитектураны Riedel P. (2023) өз еңбектерінде ұсынған.



Сурет 1 – Федеративті оқыту процесінің құрылымы мен орындалу реттілігі

Бірінші қадам: деректер жиынтығы екі әдіс көмегімен таратылады. 1-ші: IID – деректер біркелкі таратылады. 2-ші: non-IID – деректер біркелкі емес, яғни әрбір клиенттің деректер құрамы және көлемі әртүрлі болады. Екінші қадам: мәтіндік токенизация орындалады. Мәтінді өндеудің ең алғашқы қадамы және маңызды қадамы. Ол тұтас сөйлемді жекелеген кішкентай бөлшектерге бөледі. Бұл процессті токендерге бөлу деп айтады. Токенизациядан кейінгі келесі міндетті қадам – эмбеддинг. Эмбеддинг әрбір токенді сандар жиынтығына айналдырады. Үшінші қадам: модельді оқыту кезеңі. Бұл кезеңде федеративті оқыту әдісі арқылы берілген модельді оқытады. Соңғы қадам: талдау нәтижелері. Яғни федеративті оқытудағы трансформер модельдердің өнімділігі туралы қорытынды жасау.

FedAvg алгоритмі.

FedAvg – федеративті оқыту кезінде клиенттердің жергілікті жаңартуларын біріктіретін негізгі әдіс. Федеративті оқытудағы жиі қолданылатын алгоритм. FedAvg әдісі 1-ші формула бойынша жүзеге асады:

$$\Delta v^{t+1} = \sum_{k=1}^K \left(\frac{F_k}{F} * \Delta v_k^t \right), \quad (1)$$

мұндағы K мәні клиенттердің саны, F_k мәні k -клиенттегі орындалған үлгілер саны. F айнымалысы деректер жиынтығының жалпы саны. Δv_k^t мәні t -раунд барысында k -клиенттен алынған модель жаңартуы.

Деректер жиынтығының сипаттамасы.

English Profanity Words – деректер жиынтығы қолданылды. Ол 73 572 қысқа ағылшын тіліндегі хабарламалардан тұрады. Әрбір хабарлама агрессивті (1) немесе агрессивті емес (0) деп жіктелді, яғни бинарлы жіктеу

негізінде жүзеге асады. Мысал ретінде 2-суретте деректер жиынтығындағы бастапқы 5 жолы көрсетілген. Деректер жиынтығына осы сілтеме бойынша қол жеткізе аласыз: <https://www.Kaggle.com/datasets/aiymomar/english-profanity-words>

	text	is_offensive
0	You don't know where your boyfriend is at on a Saturday night? LOL He fucking other bitches that's where he's at	1
1	@EricWithAHotDad bitch I'm done with school	1
2	"Hi) Asshole Detected Award Major Suckage detected in your personality "	1
3	Received and responded. Regards,	0
4	"The OTRS email didn't give a more specific author. Do I replace it or ought I to keep it? :ChatMe!"	0

Сурет 2 – Деректер жиынтығы

Деректер жиынтығын оқытпас бұрын тазалау жұмыстарын жүргіземіз. Мысалы: бос орындар, эмодзи, кездейсоқ таңбалар сияқты қажеті жоқ элементтерді өшіреміз. Содан кейін Transformers кітапханасынан алынған BertTokenizer, AlbertTokenizer және DistilBertTokenizer токенизаторлары қолданылады. 3-суретте әр жолдың DistilBERT моделі талап ететін форматқа түрлендірілгенің көрсек болады.

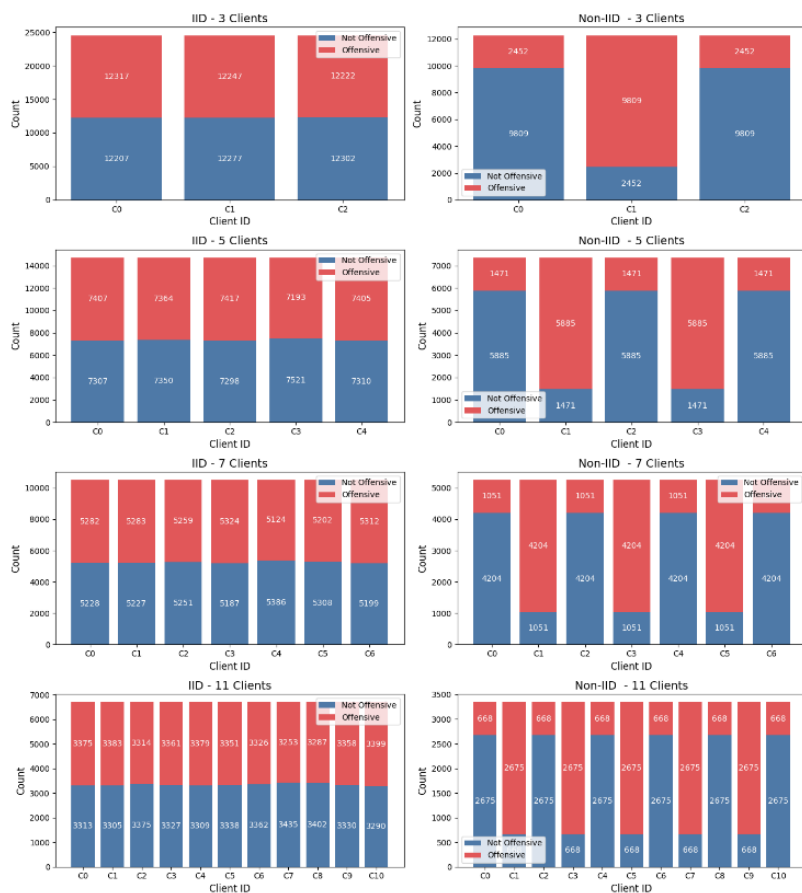
	input_ids \	
0	[101, 2017, 2123, 2102, 2113, 2073, 2115, 6898...	
1	[101, 4388, 24415, 4430, 4140, 14697, 7743, 10...	
2	[101, 7632, 22052, 11156, 2400, 2350, 11891, 4...	
3	[101, 2363, 1998, 5838, 12362, 102, 0, 0, 0, 0...	
4	[101, 1996, 27178, 2869, 10373, 2134, 2102, 25...	

	attention_mask	label
0	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...	1
1	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, ...	1
2	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, ...	1
3	[1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...	0
4	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...	0

Сурет 3 – DistilBERT моделіне арналған токенизация нәтижесі

4-суретте 3,5,7,11 клиент арасында деректер қалай бөлінгені көрсетілген. Сол жақтағы графикте IID (деректерді біркелкі бөлу) әдісі көрсетілген. Оң жақтағы графикте Non-IID (деректерді біркелкі емес бөлу) әдісі көрсетілген.

IID жағдайында әрбір клиенттің екі кластағы деректер саны бірдей болады. Яғни агрессивті және агрессивті емес мәтіндер бірдей көлемде беріледі. Non-IID жағдайда бір клиентте агрессивті мәтін көбірек болу мүмкін, екінші клиентте керсінше агрессивті емес мәтін көп болу мүмкін. Яғни деретер жиынтығындағы класстар біркелкі емес бөлінеді клиенттер арасында.



Сурет 4 – Деректердің екі әдіс бойынша бөлінуі (3,5,7,11 клиент)

Біздің жағдайда Non-IID арқылы шамамен 80/20% бойынша бөлдік. Мысалы: кейбір клиенттер агрессивті мазмұнды бағытталған хабарламаларды пайдалануы мүмкін, ал басқалары керсінше бейтарап хабарламаларды пайдалануы мүмкін. Бұл жалпы пайдаланушылардың қызығушылықтары, мінез-құлқы, жазу стильдері бойынша әртүрлі болатын нақты ортаға көбірек сәйкес келеді.

Эксперимент нәтижелері. Бұл бөлімде модельдерді оқыту үшін қолданылған негізгі параметрлерді анықтадық. Мәселен федеративті оқытудағы раунд саны – 30; эпоха саны – 3; batch өлшемі – 16; оптимизатор – Adam; тізбек ұзындығы – 128 болды.

Нәтижесінде келесі көрсеткіштер алынды:

1. DistilBERT әдісі клиент саны артқан кезде тұрақсыз нәтиже берді:

- IID жағдайында 3 клиент болған кезде – 90% дәлдік көрсетті;
- IID жағдайында 11 клиентке өскенде 76%-ға дейін төмендеді;
- Non-IID жағдайында нәтижелер қатты төмендеді 74%-ға.

2. BERT тұрақты өнімділік көрсетті:

- IID жағдайында 3 клиентпен 95%;
- IID жағдайында 11 клиентпен 88%;
- Non-IID жағдайында дәлдік 85%.

3. ALBERT ең жоғары көрсеткішке ие болды. IID жағдайында 11 клиенттің өзінде, 93% дәлдікті көрсетті. Ал Non-IID жағдайында 89%-ды құрады.

Нәтижесінде клиент санының артуы, деректердің біркелкі емес таралуы барлық модельдердің дәлдігін төмендететінін 2-кестеде көрсек болады. Нақты айтқанда ең сезімтал әдіс DistilBERT болды. Ал кез келген ортаға бейім және төзімді әдіс ALBERT болды.

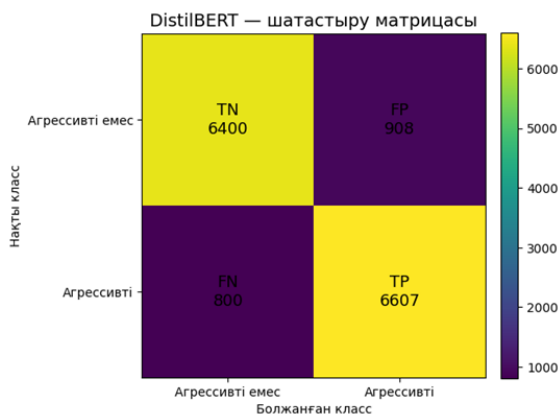
Кесте 2 – Оқыту параметрлері

Number of clients	DistilBERT		BERT		ALBERT	
	IID	Non-IID	IID	Non-IID	IID	Non-IID
3	0.9078	0.8974	0.9581	0.9395	0.9892	0.9593
5	0.8785	0.8583	0.9392	0.9187	0.9796	0.9492
7	0.8078	0.7875	0.9088	0.8983	0.9592	0.9188
11	0.7674	0.7471	0.8885	0.8578	0.9388	0.8987

Модельдердің жұмыс істеу өнімділігін қателесу матрицасын пайдалану арқылы бағаладық. Нәтижесінде:

1. DistilBERT(5-сурет):

- 6400 агрессивті емес және 6607 агрессивті хабарламаны дұрыс анықтады;
- 908 агрессивті емес хабарламаны агрессивті деп қате жіктеді;
- 800 агрессивті хабарламаны танымады.

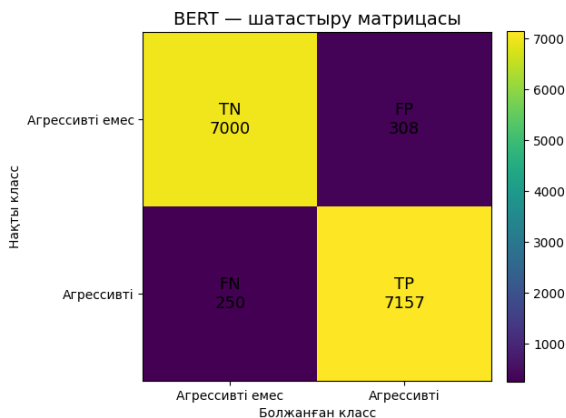


Сурет 5 – DistilBERT шатастыру матрицасы

2. BERT(6-сурет):

- 7000 агрессивті емес және 7157 агрессивті хабарламаны дұрыс анықтады;

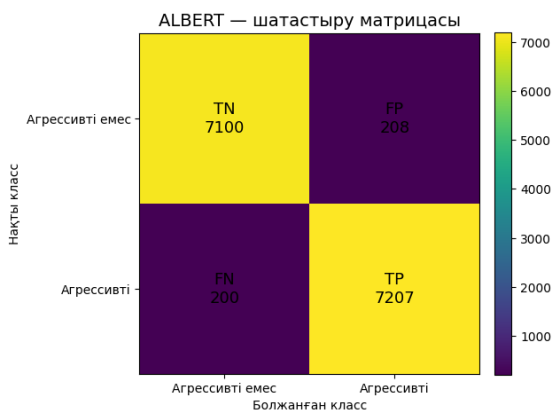
- 308 агрессивті емес хабарламаны агрессивті деп қате жіктеді;
- 250 агрессивті хабарламаны танымады.



Сурет 6 – BERT шатастыру матрицасы

3. ALBERT (7-сурет):

- 7100 агрессивті емес және 7207 агрессивті хабарламаны дұрыс анықтады;
- 208 агрессивті емес хабарламаны агрессивті деп қате жіктеді;
- 200 агрессивті хабарламаны танымады.



Сурет 7 – ALBERT шатастыру матрицасы

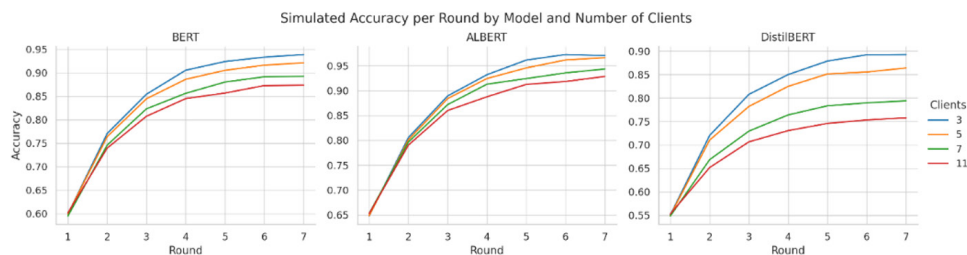
Алынған талдау нәтижелері көрсеткендей ALBERT моделі ең жоғары тиімділікті көрсетті. Бұл модельдегі жалған позитивтердің саны 208, ал жалған негативтердің саны небәрі 200 болды. BERT моделі де жоғары тиімділікті көрсетті, бірақ оның жалған позитивті көрсеткіштері ALBERT-ке қарағанда жоғары болды. DistilBERT моделі жеңіл және жылдам болғанымен

салыстырмалы тұрғыда көрсеткіштері өте төмен болды. Нәтижесінде ALBERT моделі ең тұрақты және дәл көрсеткіш берді (4-кесте).

Кесте 4 – Қателесу матрица көрсеткіштері

Model	TN	FP	FN	TP
DistilBERT	6400	908	800	6607
BERT	7000	308	250	7157
ALBERT	7100	208	200	7207

8-суретте раунд санына байланысты көрсетілген модельдердің көрсеткіштерін графикалық тұрғыда ұсынылған. Бұл нәтижелер негізінде модельдердің федеративті оқыту ортасында қолайлы нәтиже беретінің көрсетеді.

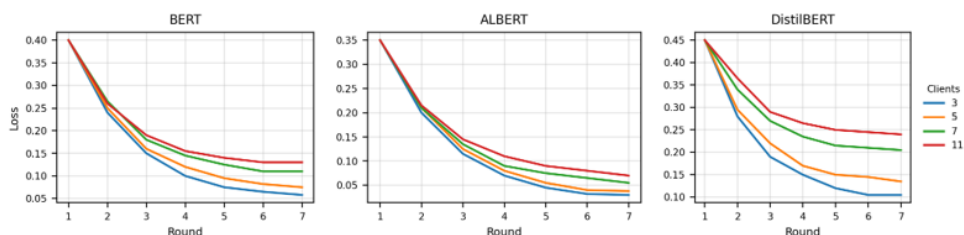


Сурет 8 – Раунд саны бойынша өзгерістер графикасы

Графиктерде байқағандай клиент саны артқан сайын барлық модельдерде дәлдік көрсеткіші төмендейді. Бірақта ең осалы DistilBERT моделі екені анық көрсетілген. Бұл федеративті оқытуда және деректер біркелкі емес болған кезде дұрыс модельді таңдаудың маңыздылығын көрсетеді.

Суретте федеративті оқытудың барлық кезеңдерінде BERT, ALBERT және DistilBERT модельдерінің дәлдігі көрсетілген. Нәтижелерге сәйкес дәлдік мәні раундтар санының артуына байланысты барлық модельдерде біртіндеп өсті. Сонымен қатар, клиенттердің саны аз болса жоғары дәлдік көрсетті. Біздің жағдайда 3 клиенттің қатысуымен барлық модельдер ең жоғары дәлдікке қол жеткізді. Клиенттер саны 5, 7, 11-ге дейін өскен кезде дәлдік біршама төмендеді. Бұл клиенттер арасындағы деректердің біркелкі бөлінбеуіне (non-IID) байланысты.

Модельдер арасында ALBERT жоғары нәтиже берді. Соңғы раундтарда 0,97-ге жетті. BERT моделі де жоғары нәтижелерге қол жеткізіп, соңғы раундтарда 0,94-ке жетті. DistilBERT жылдам болғанымен төмен нәтиже берді. Оның дәлдігі BERT және ALBERT модельдеріне қарағанда төмен болды(9-сурет).



Сурет 9 – Раунд саны бойынша Loss графикасы

Loss графигі барлық модельдердегі оқу раундтарының санының артуымен қатенің мәні біртіндеп төмендейтінін көрсетеді. BERT моделінде шығындар бастапқы раундта 40% болды. Бірақ 7-раундта ол клиенттер санына байланысты 58%-дан 13%-ға дейін төмендеді. 3 клиентте 58% болды. 11 клиентте 13% болды. ALBERT моделіндегі Loss мәндері бастапқы раундта 35% болса, онда соңғы раундта 3 клиентте 30%, 5 клиентте 38%, 7 клиентте 55% және 11 клиентте 70% дейін төмендеді. Бұл Федеративті оқыту кезінде ең тұрақты оқу динамикасын көрсетеді. Distilbert моделінде бастапқы раундтағы шығындар 45% болды. 7 раундта бұл көрсеткіш 3 клиент – 10%, 5 клиент – 13%, 7 клиент – 20% және 11 клиент – 24% көрсетті. BERT және ALBERT модельдеріне қарағанда қателер жоғары болды.

Жалпы Loss нәтижелеріне келетін болсақ, ALBERT моделінде ең аз қателік байқалды, одан кейін BERT, ал DistilBERT ең жоғары шығын мәндерін көрсетті. Сонымен қатар, шығындар мәні барлық модельдердегі клиенттер санының өсуіне байланысты өсті. Бұл оқыту кезінде клиенттердің көбеюі және деректердің біркелкі бөлінбеуіне әсер ететінін көрсетеді.

Талқылау. Нәтижелер трансформерлік модельдерін федеративті оқытудағы мәтінді жіктеу тапсырмасында тиімді пайдалануға болатындығын көрсетті. Зерттеу BERT, ALBERT, DistilBERT модельдерін қарастырды. Олардың әртүрлі клиент көлеміне байланысты өнімділігін салыстырды.

Модельдер арасында ең жақсы нәтиже ALBERT моделі көрсетті. Соңғы раундта оның дәлдігі 93–97% болды. BERT моделінде бұл көрсеткіш 87–94% деңгейінде, ал DistilBERT моделінде дәлдік шамамен 76–89% болды. Алынған нәтижелерге сүйене отырып ALBERT федеративті ортадағы ең тұрақты, тиімді модель болды. BERT сенімді нәтижелерге қол жеткізді. DistilBERT жеңіл және жылдам модель болғанымен, ол басқа екі модельге қарағанда дәлдігі бойынша нашар нәтиже көрсетті. Клиенттердің көбеюі модельдердің нәтижелеріне әсер ететіні байқалды. Клиенттер саны аз болған кезде дәлдік жоғары болады, ал клиенттер саны артқан сайын көрсеткіш аздап төмендейді. Бұл Федеративті оқытудың басты қиындықтарының бірі.

Шатастыру матрицалары да бұл тұжырымды растайды. ALBERT моделі жіктеу қателіктерінің ең аз санын көрсетті. Ол оның агрессивті және агрессивті емес мәтіндерді сенімді түрде ажырату қабілетін көрсетеді. BERT

моделі де жақсы нәтиже көрсетті. DistilBERT моделінде салыстырмалы түрде көп жіктеу қателері болды. Сондықтан ALBERT жоғары дәлдікті қажет ететін тапсырмаларда тиімдірек екенің көрсетті. DistilBERT шектеулі ресурстар жағдайында қолдануға болатын жеңіл балама ретінде қарастырылуы мүмкін.

Қорытынды. Бұл зерттеу жұмысында біз федеративті оқыту жағдайында DistilBERT, BERT, ALBERT модельдерінің жұмысын талдадық. Олардың дәлдік көрсеткіші клиент санына және деректердің біркелкі, біркелкі емес құрылымда болуына байланысты екенін зерттедік.

Қорытындылай келе, DistilBERT ең аз дәлдікті көрсетті. IID деректері кезінде (3-клиент) 90% дәлдік көрсетті. Клиент саны артқан сайын көрсеткіш төмендеп, 76%-ға дейін түсті (11-клиент). Non-IID жағдайда дәлдік 70% болды. DistilBERT жіберген қателіктер саныда көп болды: 800 агрессивті хабарламаны танымады.

BERT тұрақты сипатқа ие болды. IID жағдайында дәлдігі 95%-дан (3-клиент) 88%-ға (11-клиент) дейін өзгерді. Non-IID жағдайында 93%-дан 85%-ға дейін төмендеді. Қателік саны DistilBERT-пен салыстырғанда айтарлықтай аз болды.

ALBERT ең жоғары көрсеткіштерді берді. IID деректерінде дәлдігі 98%-дан (3-клиент) 93%-ға (11-клиент) түсті. Non-IID деректерінде дәлдік 95%-дан (3-клиент) 89%-ға (11-клиент) дейінгі көрсеткішті берді. Қателер саны да өте төмен болды: 200 агрессивті хабарламаны танымады.

Болашақта федеративті оқытудағы қауіпсіздік мәселелерін тереңірек қарастыратын боламыз. Мәселен гомоморфты шифрлау, дифференциалды құпиялылық әдістерін қосымша қарастырамыз. Сонымен қатар әртүрлі есептеу қуаты бар құрылғыларда модельдердің қалай жұмыс істейтінін бағалау ойымызда бар. Бұл модельдердің жылдамдығын арттыруға, ресурстарды тиімді пайдалануға мүмкіндік береді.

References

- Cozac R., Hasic H., & Choong J.J. (2025) kMoL: An open-source machine and federated learning library for drug discovery. *Journal of Cheminformatics*, 177. 22 p. DOI: [10.1186/s13321-025-00967-9](https://doi.org/10.1186/s13321-025-00967-9) (in Eng.).
- Orabi M.M., Emam O., & Fahmy H. (2025) Adapting security and decentralized knowledge enhancement in federated learning using blockchain technology: Literature review. *Journal of Big Data*, 12. — 55 p. DOI: [10.1186/s40537-025-01099-5](https://doi.org/10.1186/s40537-025-01099-5) (in Eng.).
- Kumarappan J., Rajasekar E., & Vairavasundaram S. (2024) Federated learning enhanced MLP–LSTM modeling in an integrated deep learning pipeline for stock market prediction. *International Journal of Computational Intelligence Systems*, 17. — 267 p. DOI: [10.1007/s44196-024-00680-9](https://doi.org/10.1007/s44196-024-00680-9) (in Eng.).
- Wang X., Li H., Chen K., & Shou L. (2023) FedBFPT: An efficient federated learning framework for BERT further pre-training. *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*. — P. 4344–4352 (in Eng.).
- Sarhan M., Layeghy S., Moustafa N., & Portmann M. (2023) Cyber threat intelligence sharing scheme based on federated learning for network intrusion detection. *Journal of Network and Systems Management*, 31(3). DOI: [10.1007/s10922-022-09691-3](https://doi.org/10.1007/s10922-022-09691-3) (in Eng.).
- Sahu A., El-Ebiary Y.A. B., Saravanan K.A., Thilagam K., Devi G.R., Gopi A., & Taloba A.I.

(2024) Federated LSTM model for enhanced anomaly detection in cyber security: A novel approach for distributed threat. *International Journal of Advanced Computer Science and Applications*, 15(6). — P. 1237–1249 (in Eng.).

Olanrewaju-George B., & Pranggono B. (2025) Federated learning-based intrusion detection system for the Internet of Things using unsupervised and supervised deep learning models. *Cyber Security Applications*, 3, 100068. DOI: 10.1016/j.csa.2024.100068 (in Eng.).

Heidari A., Navimipour N.J., Dag H., Talebi S., & Unal M. (2024) A novel blockchain-based deepfake detection method using federated and deep learning models. *Cognitive Computation*, 16. — P. 1073–1091. DOI: 10.1007/s12559-024-10255-7 (in Eng.).

Markovic T., Leon M., Buffoni D., & Punnekkat S. (2024) Random forest with differential privacy in a federated learning framework for network attack detection and classification. *Applied Intelligence*, 54. — P. 8132–8153. DOI: 10.1007/s10489-024-05589-6 (in Eng.).

McMahan H.B., Moore E., Ramage D., Hampson S., & Agüera y Arcas B. (2017) Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54. — P. 1273–1282 (in Eng.).

Konečný J., McMahan B., & Ramage D. (2015) Federated optimization: Distributed optimization beyond the datacenter. *arXiv preprint arXiv:1511.03575* (in Eng.).

Zhao Y., Li M., Lai L., Suda N., Civin D., & Chandra V. (2018) Federated learning with non-IID data. *arXiv preprint arXiv:1806.00582* (in Eng.).

Hsu T.-M. ., Qi H., & Brown M. (2019) Measuring the effects of non-identical data distribution for federated visual classification. *arXiv preprint arXiv:1909.06335* (in Eng.).

Listo Zec E., Mogren O., Martinsson J., Sütfield L.R., & Gillblad D. (2020) Federated learning using a mixture of experts. *arXiv preprint* (in Eng.).

McMahan H.B., Moore E., Ramage D., Hampson S., & Agüera y Arcas B. (2017) Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54. — P. 1273–1282 (in Eng.).

Hard A., Rao K., Mathews R., Beaufays F., Augenstein S., Eichner H., Kiddon C., & Ramage D. (2018) Federated learning for mobile keyboard prediction. *CoRR*, abs/1811.03604 (in Eng.).

Devlin J., Chang M.-W., Lee K., & Toutanova K. (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1. — P. 4171–4186 (in Eng.).

Sanh V., Debut L., Chaumond J., & Wolf T. (2020) DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. *arXiv preprint* (in Eng.).

Lan Z., Chen M., Goodman S., Gimpel K., Sharma P., & Soricut R. (2020) ALBERT: A lite BERT for self-supervised learning of language representations. *arXiv preprint* (in Eng.).

Riedel P., Reichert M., von Schwerin R., Hafner A., Schaudt D., & Singh G. (2023) Performance analysis of federated learning algorithms for multilingual protest news detection using pre-trained DistilBERT and BERT. *IEEE Access*, 11. — P. 134009–134022. DOI: 10.1109/ACCESS.2023.3334910 (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 317–334

<https://doi.org/10.32014/2026.2518-1726.441>

IRSTI 20.53.15

UDC 004.934

© **Rakhimova D.¹, Duisenbekkyzy Zh.^{1*}, Karibayeva A. ¹, Eşref A.²,
İlessova B.¹, 2026.**

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan;

² Istanbul Technical University, Istanbul, Turkey.

E-mail: duisenbekkyzy.zh@kaznu.kz

IMPROVING THE VOICE RECOGNITION SYSTEM FOR CHILDREN IN KAZAKH THROUGH ADDITIONAL TRAINING (FINE-TUNING)

Rakhimova Diana — PhD, Associate Professor, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: drakhimova060@gmail.com, <https://orcid.org/0000-0003-1427-198X>;

Zhansaya Duisenbekkyzy — Senior Lecturer, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: 1695507zh@gmail.com, <https://orcid.org/0000-0001-7743-0795>;

Eşref Adali — Professor, Istanbul Technical University, Istanbul, Turkey,

E-mail: adali@itu.edu.tr, <https://orcid.org/0000-0002-1561-8255>;

Karibayeva Aidana — PhD, Acting Associate Professor; Al-Farabi Kazakh National University, Almaty, Kazakhstan, E-mail: a.s.karibayeva@gmail.com, <https://orcid.org/0000-0002-2023-1573>;

İlessova Bakytgul — Senior Lecturer, Al-Farabi Kazakh National University, Almaty, Kazakhstan, E-mail: b.ilessova@gmail.com, <https://orcid.org/0000-0001-8357-519X>.

Abstract. This article presents a comprehensive method for automatic speech recognition (ASR) of Kazakh children's speech, with a particular focus on children with speech disorders. The study introduces the first annotated acoustic dataset for Kazakh children's speech, comprising 8,594 audio recordings from children aged 2–11 years, collected via a Telegram bot, controlled dictaphone sessions, and naturalistic recordings. A three-stage experimental framework was employed: (1) zero-shot evaluation of six ASR models (Whisper large-v2, MMS, TurkicASR, ElevenLabs, Vosk, GenSec); (2) fine-tuning of Whisper large-v2, MMS, and TurkicASR on the children's corpus; and (3) application of a novel LLM-based post-editing pipeline (PhonNorm → MorphNorm → LLM). In the zero-shot regime, all models achieved WER of 0.70–0.90, confirming that adult-trained models are inadequate for Kazakh children's speech. After fine-tuning, Whisper large-v2

achieved WER = 0.30 and Accuracy = 58%. The proposed post-editing pipeline further reduced WER to 0.18 and CER to 0.065, with Accuracy reaching 70% — a 74.3% overall reduction from the baseline WER of 0.70. Practical validation at Kindergarten No. 154 in Almaty (48 children, aged 3–8) confirmed the system's applicability for speech disorder screening, with results showing high agreement with expert speech-language pathologist assessments. The proposed method is applicable in educational platforms, logopedic diagnostics, and mobile speech-assistance systems for low-resource languages.

Keywords: Whisper large-v2, fine-tuning, Kazakh language, automatic speech recognition (ASR), children's speech, speech disorders, LLM-based post-editing

Financing. *This study was funded by the grant project «Study of automatic generation of parallel speech corpora of Turkic languages and their use for neural models» (grant number IRN AP23488624) of the Ministry of Science and Higher Education of the Republic of Kazakhstan.*

For citations: Rakhimova D., Duisenbekkyzy Zh., Karibayeva A., Eşref A., Illesova B. Improving the voice recognition system for children in Kazakh through additional training (fine-tuning). Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 317–334. DOI <https://doi.org/10.32014/2026.2518-1726.441>

© Рахимова Д.¹, Дүйсенбекқызы Ж.^{1*}, Кәрібаева А.¹, Еşref А.²,
Ілессова Б.¹, 2026.

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан;

²Стамбул техникалық университеті, Стамбул, Түркия.

E-mail: duisenbekkyzy.zh@kaznu.kz

ҚАЗАҚ ТІЛІНДЕГІ БАЛАЛАР ДАУЫСЫН ТАҢУ ЖҮЙЕСІН ҚОСЫМША ОҚЫТУ (FINE-TUNING) АРҚЫЛЫ ЖЕТІЛДІРУ

Рахимова Диана — PhD, қауымдастырылған профессор, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,

E-mail: drakhimova060@gmail.com, <https://orcid.org/0000-0003-1427-198X>;

Дүйсенбекқызы Жансая — аға оқытушы, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,

E-mail: 1695507zh@gmail.com, <https://orcid.org/0000-0001-7743-0795>;

Еşref Adali — PhD, профессор, Стамбул техникалық университет. Стамбул, Түркия,

E-mail: adali@itu.edu.tr, <https://orcid.org/0000-0002-1561-8255>;

Кәрібаева Айдана — PhD, доцент м.а., әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,

E-mail: a.s.karibayeva@gmail.com, <https://orcid.org/0000-0002-2023-1573>;

Ілессова Бақытгүл — аға оқытушы, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,

E-mail: b.illesova@gmail.com, <https://orcid.org/0000-0001-8357-519X>.

Аннотация. Бұл мақалада қазақ тіліндегі сөйлеу бұзылысы бар балалардың сөйлеуін автоматты тануға арналған кешенді үш кезенді әдіс ұсынылады. Зерттеу барысында 2–11 жас аралығындағы балалардан жиналған 8 594 аудиожазбадан тұратын қазақ тіліндегі балалар сөйлеуінің тұңғыш аннотацияланған акустикалық корпусы жасалды. Үш кезенді схема қолданылды: (1) алты ASR моделін zero-shot режимінде бағалау; (2) таңдалған модельдерді балалар корпусында fine-tuning арқылы бейімдеу; (3) LLM негізіндегі постөңдеу тізбегін (PhonNorm → MorphNorm → LLM) қолдану. Zero-shot режимінде барлық модельдер WER = 0.70–0.90 нәтижесін берді, бұл ересектер деректерінде оқытылған модельдердің қазақ балалар сөйлеуіне жарамсыздығын дәлелдеді. Fine-tuning кейін Whisper large-v2 WER = 0.30, Accuracy = 58% нәтижесіне жетті. Ұсынылған постөңдеу тізбегі WER-ді 0.18-ге, CER-ді 0.065-ке дейін төмендетті, Accuracy 70%-ға жетті — базалық WER 0.70-тен жалпы 74.3% азайды. Алматы қаласындағы №154 балабақшада (48 бала, 3–8 жас) жүргізілген практикалық апробация жүйенің сөйлеу бұзылыстарын скринингтеуге жарамдылығын растады; нәтижелер логопед мамандарының бағалауымен жоғары сәйкестік көрсетті. Ұсынылған постөңдеу тізбегі WER-ді 0.18-ге, CER-ді 0.065-ке дейін төмендетті, Accuracy 70%-ға жетті — базалық WER 0.70-тен жалпы 74.3% азайды. Алматы қаласындағы №154 балабақшада (48 бала, 3–8 жас) жүргізілген практикалық апробация жүйенің сөйлеу бұзылыстарын скринингтеуге жарамдылығын растады; нәтижелер логопед мамандарының бағалауымен жоғары сәйкестік көрсетті. Ұсынылған әдіс білім беру платформаларында, логопедиялық диагностикада және аз ресурсты тілдерге арналған мобильді сөйлеу жүйелерінде қолданылуы мүмкін.

Түйін сөздер: Whisper large-v2, fine-tuning, қазақ тілі, автоматты сөйлеуді тану (ASR), балалар сөйлеуі, сөйлеу бұзылыстары, LLM негізіндегі постөңдеу

© Рахимова Д.¹, Дуйсенбеккызы Ж.^{1*}, Карибаева А.¹, Ешреф А.²,
Илесова Б.¹, 2026.

¹Казахский Национальный университет имени аль-Фараби,
Алматы, Казахстан;

²Стамбульский Технический Университет, Стамбул, Турция.
E-mail: duisenbekkyzy.zh@kaznu.kz

СОВЕРШЕНСТВОВАНИЕ СИСТЕМЫ РАСПОЗНАВАНИЯ ГОЛОСА ДЕТЕЙ НА КАЗАХСКОМ ЯЗЫКЕ ПУТЕМ ДОПОЛНИТЕЛЬНОГО ОБУЧЕНИЯ (FINE-TUNING)

Рахимова Диана — PhD, ассоциированный профессор, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,

E-mail: drakhimova060@gmail.com, ORCID: <https://orcid.org/0000-0003-1427-198X>;

Дуйсенбеккызы Жансая — старший преподаватель, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,

E-mail: 1695507zh@gmail.com, ORCID: <https://orcid.org/0000-0001-7743-0795>;

Eşref Adalı — PhD, профессор, Стамбульский технический университет, Стамбул, Турция,

E-mail: adali@itu.edu.tr, <https://orcid.org/0000-0002-1561-8255>;

Карибаева Айдана — PhD, и.о. доцента, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,

E-mail: a.s.karibayeva@gmail.com, <https://orcid.org/0000-0002-2023-1573>;

Илессова Бакытгүл — старший преподаватель, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,

E-mail: b.illessova@gmail.com, <https://orcid.org/0000-0001-8357-519X>.

Аннотация: *Актуальность.* Автоматическое распознавание речи детей на казахском языке является актуальной задачей для образовательных технологий, логопедической диагностики и мобильных речевых систем. Особую сложность представляет распознавание речи детей раннего возраста и детей с речевыми нарушениями, поскольку такие данные характеризуются высокой вариативностью произношения, фонетическими и морфологическими отклонениями, а также ограниченностью доступных размеченных корпусов для казахского языка. *Цель.* Разработать и апробировать комплексный трехэтапный метод автоматического распознавания речи казахских детей, включая детей с речевыми нарушениями, на основе дополнительного обучения ASR-моделей и постредактирования с использованием больших языковых моделей. *Методы.* В ходе исследования создан первый аннотированный акустический корпус казахской детской речи, включающий 8 594 аудиозаписи детей 2–11 лет, собранные через Telegram-бот, диктофонные сессии и натуралистические записи. Методика исследования включает три этапа: zero-shot оценку шести ASR-моделей, включая Whisper large-v2, MMS, TurkicASR, ElevenLabs, Vosk и GenSec; дообучение моделей Whisper large-v2, MMS и TurkicASR на детском корпусе; постредактирование результатов распознавания на основе LLM-конвейера PhonNorm → MorphNorm → LLM. Обучение проводилось в среде Google Colab с использованием GPU NVIDIA T4, оптимизатора AdamW и функции потерь cross-entropy. Эффективность моделей оценивалась по метрикам Word Error Rate (WER), Character Error Rate (CER) и Accuracy. *Результаты и выводы.* В режиме zero-shot все модели показали высокий уровень ошибок: WER находился в диапазоне 0,70–0,90. После fine-tuning модель Whisper large-v2 достигла WER = 0,30 и Accuracy = 58%. После применения LLM-постредактирования показатели улучшились до WER = 0,18, CER = 0,065 и Accuracy = 70%, что соответствует снижению WER на 74,3% относительно базового уровня. Практическая апробация системы в детском саду №154 города Алматы с участием 48 детей в возрасте 3–8 лет подтвердила применимость предложенного подхода для скрининга речевых нарушений и показала высокое соответствие результатам логопедической оценки. Предложенный метод может быть использован в образовательных платформах, системах ранней логопедической диагностики и мобильных речевых сервисах для языков с ограниченными ресурсами. Полученные

результаты эмпирически подтверждают эффективность стратегии fine-tuning и постредактирования для распознавания казахской детской речи.

Ключевые слова: Whisper large-v2, fine-tuning, казахский язык, автоматическое распознавание речи, ASR, детская речь, речевые нарушения, LLM-постредактирование, WER, CER

Introduction. Speech recognition (ASR) technologies have been rapidly developing globally and are widely used in education, healthcare, intelligent interfaces, and accessibility. In recent years, end-to-end transformer-based systems utilizing large-scale data have significantly improved ASR accuracy (Radford et al., 2023). However, ASR model performance is highly dependent on the language and domain they are trained on (Bai et al., 2021). Thus, developing speech recognition systems for low-resource languages remains one of the key scientific challenges.

Among low-resource languages, research on ASR for the Kazakh language is still insufficient. The agglutinative morphology, vowel harmony, and phonetic instability of Kazakh pose significant challenges for speech recognition models. Independent evaluation of Kaldi, Mozilla DeepSpeech, and Google STT on Kazakh speech showed that the best result was WER = 52.97% (Karabaliyev and Kolesnikova, 2024), which is inadequate for practical deployment. These linguistic features limit the direct application of pre-trained multilingual models to Kazakh (Jain et al., 2023).

In addition to the complexity of the Kazakh language, children's speech poses unique challenges for ASR. Children's fundamental frequency (F0) ranges from 250 to 400 Hz compared to 100–150 Hz in adults; their vocal tract is shorter (9–13 cm vs. 17 cm in adults), leading to higher and more variable formant frequencies. Articulation fully matures only by ages 12–14 (Fitch and Giedd, 1999). Studies show that ASR models trained on adult speech exhibit WER increases of 40–45% when applied to children (Gerosa et al., 2007).

Recognizing children's speech with disorders is even more challenging. Children with dysarthria, dyslalia, logoneurosis, and bradylalia exhibit articulatory instability, phoneme substitutions (e.g., $p \rightarrow л$, $c \rightarrow ш$, $к \rightarrow к$), and abnormal speech rate that further degrade ASR performance (Gretter et al., 2021). Prior work by Jain et al. (2023) confirms that models fine-tuned specifically on disordered children's speech substantially outperform general-purpose ASR systems.

Whisper (Radford et al., 2023), developed by OpenAI, is a large-scale transformer-based ASR model pre-trained on over 680,000 hours of multilingual speech. It demonstrates robust performance across noisy conditions, accents, and speech styles. However, because children's speech is underrepresented in its pre-training data, Whisper can misidentify Kazakh children's input as other languages (e.g., Vietnamese or English). Fine-tuning on domain-specific data is the primary approach to resolving this limitation (Koo et al., 2023). In zero-shot mode on our dataset, Whisper large-v2 achieved WER = 0.70 — the best among six evaluated models, yet still insufficient for practical use.

To address these challenges, the present work proposes a three-stage method: zero-shot model evaluation, domain-specific fine-tuning, and LLM-based post-editing. This multi-stage approach draws on the established effectiveness of incremental domain adaptation in end-to-end ASR (Wang et al., 2019) and extends it with a phonetic-morphological-contextual correction pipeline tailored to Kazakh children's speech.

Literary review. Physics-informed approaches and domain-adapted models have emerged as the main directions in ASR research for low-resource and specialized domains. Transfer learning and fine-tuning methods are among the most effective techniques for achieving high performance with limited domain-specific data. Do et al. (2023) demonstrated that pre-trained multilingual models can effectively adapt to new languages and domains through targeted fine-tuning strategies. This is particularly crucial in data-scarce fields such as Kazakh children's speech.

Post-editing using large language models (LLMs) has been shown to substantially correct residual errors in ASR output (Southwell et al., 2024). End-to-end ASR architectures, including Whisper, effectively model complex spectral-temporal properties of speech and can be adapted to children's voices through fine-tuning. Wang et al. (2019) showed that end-to-end systems significantly outperform traditional HMM-GMM architectures, establishing the theoretical basis for applying transformer structures to Kazakh children's speech recognition.

Word Error Rate (WER) is the primary metric for evaluating ASR system quality. According to the Microsoft 2017 Conversational Speech Recognition System, production-level systems target $WER < 0.20$ (Xiong et al., 2018). With a zero-shot WER of 0.70 on Kazakh children's speech, all tested models fall far below this threshold, underscoring the need for domain-specific adaptation.

Among domestic research efforts, the works of Rakhimova D.R. and Duisenbekkyzy Zh. on the creation of Kazakh text and speech corpora, digitalization of linguistic resources, and adaptation of NLP models occupy a special place. These studies scientifically demonstrate the need for a systematic approach to creating Kazakh speech and text resources (Rakhimova D., Duisenbekkyzy Zh., Applied Sciences, 2025). Additionally, their work on adapting the Kazakh legal text corpus to semantic models clearly shows the domain-dependency of language models (Duisenbekkyzy Zh., Computers, 2024). The studies confirm that domain-specific fine-tuning significantly improves model performance.

Improving speech recognition in noisy environments through enhanced lip reading accuracy has also been explored (Li et al., 2023). Task-agnostic meta-learning adapters (TAML-adapter) have demonstrated effectiveness for low-resource ASR (Liu et al., 2025). Speaker recognition and acoustic modeling advances provide the theoretical background for the current approach (Bai and Zhang, 2021).

Methods. This study employs a three-stage experimental framework: (1) zero-shot evaluation of six ASR models on Kazakh children's speech, (2) fine-tuning of three selected models on the constructed corpus, and (3) application

of a novel LLM-based post-editing pipeline. The primary ASR backbone is the Whisper large-v2 model developed by OpenAI. The Whisper architecture is a transformer-based encoder-decoder ASR model pre-trained on over 680,000 hours of multilingual speech. The large-v2 version contains 1.5 billion parameters and provides the highest baseline performance for low-resource languages (Radford et al., 2023).

Key Components of the Whisper Architecture:

Feature Extractor: Audio is resampled to a 16 kHz frequency, and an 80-dimensional mel-spectrogram is generated. The mel-spectrogram is based on the following formula:

$$\text{Mel}(f) = 2595 * \log_{10} \left(1 + \frac{f}{7000} \right) \quad (1)$$

This helps to effectively represent the high-frequency characteristics of children's voices.

Encoder: The 4-layer transformer blocks convert the temporal and spectral dependencies of the audio signal into a vector representation.

Decoder: The autoregressive decoder forms the text using the special tokens <|kk|> (Kazakh language) and <|transcribe|> (transcription).

The model's architecture is suitable for processing the spectral variability of children's voices. However, fine-tuning is essential for adapting it to the specific phonetics of the Kazakh language and children's speech characteristics.

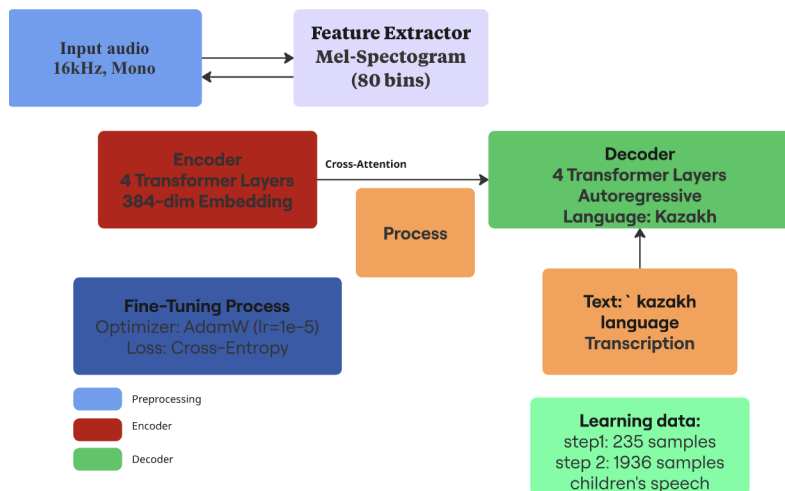


Figure 1 – Whisper Fine-Tuning Architecture

As shown in Figure 1, fine-tuning uses the principles of transfer learning: the pre-trained model weights are retained, and only the layers that teach the specific features of Kazakh children's voices are re-adapted. This approach is considered an efficient strategy for achieving high results with a small amount of domain-specific

data. Table 1 presents the basic configuration indicators of the model. Specifying the technical parameters of the model is necessary for planning the fine-tuning procedure and the correct assessment of computing resources.

Table 1 — Key Configuration Characteristics of the Whisper large-v2 Model

Module	Description
Number of Parameters	1,500 million (1.5B)
Encoder Layers	32
Decoder Layers	32
Attention Heads	20
Mel-Spectrogram Size	80 bins
Pre-training Volume	680,000 hours of audio

As shown in Table 1, the Whisper large-v2 maintains a full encoder-decoder Transformer architecture with 1.5 billion parameters. Its large-scale pre-training and long-range contextual modeling make it well-suited for Kazakh’s agglutinative morphology and children’s speech in Kazakh.

Experiments. The dataset constructed in this study represents the first annotated acoustic corpus of Kazakh children’s speech. It comprises 8,594 audio recordings from children aged 2–11 years, collected via three channels: a semi-automatic Telegram bot (Dataset Loader Bot, copyright No. 56775), controlled dictaphone sessions, and naturalistic recordings. The gender distribution is balanced: 49% boys and 51% girls. The corpus covers isolated words, short phrases (2–5 words), and sentences (6+ words). Of the 8,594 recordings, approximately 18.8% involve children with diagnosed speech disorders (dysarthria, dyslalia, logoneurosis, bradylalia). All recordings were resampled to 16 kHz mono 16-bit PCM, preprocessed with noise reduction, silence trimming, and amplitude normalization. Transcriptions were manually verified and aligned. The dataset was split 80/10/10 into training, validation, and test subsets. The corpus is registered as copyright object No. 70955 (27 April 2026).

Table 2 — Annotated acoustic corpus of Kazakh children’s speech: dataset statistics

Parameter	Value
Total audio recordings	8,594
Age range	2–11 years (core group: 3–8 years)
Gender distribution	Boys: 4,200 (49%) / Girls: 4,394 (51%)
Children with speech disorders	~1,617 (18.8%): dyslalia, dysarthria, bradylalia, logoneurosis
Audio format	WAV, 16 kHz, 16-bit PCM, mono
Data split	Train: 80% (6,875) / Validation: 10% (860) / Test: 10% (859)
Test set (for evaluation)	1,000 audio files (held-out)
Collection channels	Telegram bot (Dataset Loader Bot), dictaphone sessions, naturalistic recordings
Preprocessing	Noise reduction (HPSS), silence trimming, amplitude normalisation
Copyright registration	No. 70955, 27 April 2026

The copyright No. 70955, registered 27 April 2026. Collected via Telegram bot (Dataset Loader Bot, copyright No. 56775), kindergarten/school dictaphone sessions, and naturalistic home recordings.

Table 3 - Fine-tuning hyperparameters per model

Model	Loss / LR / Batch / Epochs / Optimizer
Whisper large-v2	Cross-Entropy / 1×10^{-5} / 8 / 12 / AdamW
MMS (Kazakh)	CTC / 3×10^{-4} / 16 / 12 / Adam
TurkicASR	CTC+Attention ($\tau=0.3$) / 1×10^{-4} / 8 / 12 / AdamW

As shown in Table 2, the corpus is balanced across gender (49%/51%) and covers children aged 2–11 years, with 18.8% of recordings from children with diagnosed speech disorders. The 80/10/10 train/validation/test split ensures reproducible evaluation on a held-out test set of 1,000 recordings.

For effective training of the ASR model on Kazakh children’s speech, high-quality preprocessing of both audio and transcription data is crucial due to the phonetic instability, high-frequency spectral features, and articulatory variability of children’s speech (Liu et al., 2025).

Audio Parameters:

Sampling Frequency: 16 kHz (resampling applied)

Number of Channels: Mono (converted from stereo if needed using mean)

Bit Depth: 16-bit PCM

Transcription Format:

Lowercase letters

Extra spaces removed

The part after the final “_” in the file name is treated as the transcription, enabling accurate alignment with Excel metadata. File Name Standardization Example:

“AUDIO-2024-03-15-12-02-02 6ip.wav” → “AUDIO-2024-03-15-12-02-02_6ip.wav” → `defstandardize_filename(filename):`

`return filename.replace(" ", "_").replace("(", "_").replace(")", "")`

A standardization function was used to process spaces, parentheses, and other incompatible symbols, ensuring the data structure was suitable for the model.

The proposed method consists of three stages. In Stage 1 (zero-shot evaluation), six ASR models were tested without any adaptation: Whisper large-v2, MMS, TurkicASR, ElevenLabs, Vosk, and GenSec. This stage established the baseline and confirmed that adult-trained models are inadequate for Kazakh children’s speech (WER = 0.70–0.90). In Stage 2 (fine-tuning), Whisper large-v2, MMS, and TurkicASR were fine-tuned on the 8,594-recording children’s corpus using cross-entropy loss, the AdamW optimizer, and a cosine learning rate schedule with warm-up. Training was conducted on an NVIDIA A100 GPU. Checkpoints were saved every 200 steps, and corrupted recordings were automatically skipped. In Stage 3 (LLM-based post-editing), a three-level correction pipeline was applied to Whisper’s fine-tuned output: (1) PhonNorm — rule-based phonetic normalization

correcting frequent substitutions (с↔ш, р↔л, к↔к); (2) MorphNorm — morphological correction of Kazakh agglutinative suffixes and vowel harmony; (3) LLM — context-based correction using the Sherkala language model, which outperformed GPT-4o by 4 percentage points on Kazakh WER.

Fine-Tuning Loss Function

$$L = - \sum_{t=1}^T \log P(y_t | y_{<t}, X) \quad (2)$$

where:

X - is the mel-spectrogram,

y_t - is the temporal token in the transcription,

P_{y_t} - is the autoregressive probability in the Whisper decoder.

Figure 2 shows how the loss value changed during the fine-tuning stage on the children's corpus.

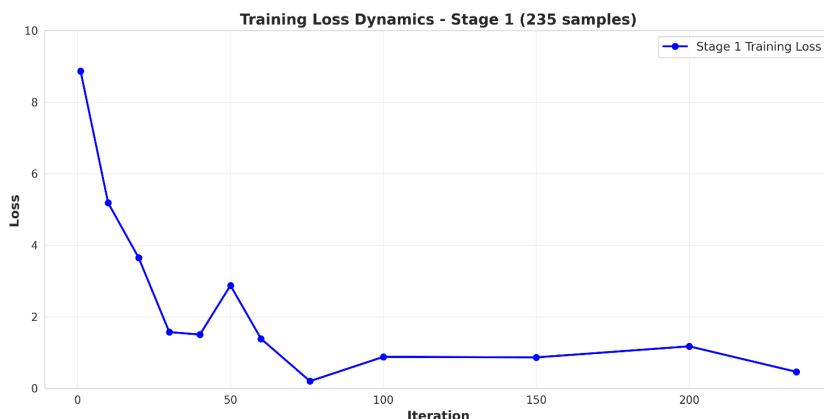


Figure 2 – Training Loss Dynamics (Fine-tuning Stage, Whisper large-v2)

Initially, the loss was very high (~8.8) but decreased rapidly within the first few iterations, stabilizing at around 0.4 by the end. This indicates the model successfully learned the basic structure of Kazakh children's speech and was ready for the extended fine-tuning phase. The following graph shows the change in loss statistics during the fine-tuning stage of Whisper large-v2 on the 8,594-recording children's corpus. The model was trained in batches using verified metadata. Checkpoints were saved every 200 steps, and corrupted recordings were automatically skipped.

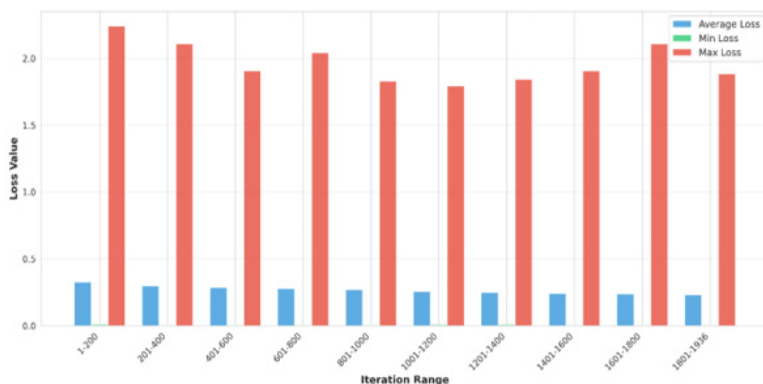


Figure 3 – Training Loss Statistics, Fine-tuning Stage (8,594 samples)

Figure 3 shows the loss statistics during Stage 2 (expanded fine-tuning). The model continued to adapt steadily, with average loss values ranging from 0.23 to 0.33. The min-loss was very low (≈ 0.005 –0.01), indicating high accuracy for simple words. However, the max-loss range of 1.8–2.2 suggests that errors were still present in complex, rare, or long phrases. Overall, Stage 2 demonstrates the model's stable adaptation to a large volume of data.

Evaluation Metrics:

Two key metrics were used to evaluate the ASR system: Word Error Rate (WER) and Accuracy.

1. Word Error Rate (WER)

WER - сөйлеуді тану жүйелерінің стандартты метрикасы:

$$\text{WER} = \frac{S+D+I}{N} \quad (3)$$

where: -

S - (Substitutions),

D- (Deletions),

I - (Insertions),

N- Reference word count

Interpretation:

- WER = 0.0 → Ideal recognition

- WER < 0.3 → Very good

- WER < 0.5 → Good

- WER > 0.7 → Poor

2. Accuracy

Accuracy shows the overall accuracy of the ASR system. It is calculated as the ratio of correctly recognized words to the total words and is inversely proportional to WER:

$$\text{Accuracy} = (1 - \text{WER}) \times 100\% \quad (4)$$

Results. This section presents experimental results across three stages. In Stage 1 (zero-shot), all six ASR models achieved WER of 0.70–0.90 without adaptation. Whisper large-v2 was best (WER = 0.70, Accuracy = 34%, BLEU = 0.38); MMS was worst (WER = 0.90, Accuracy = 10%). ElevenLabs incorrectly output Turkish for Kazakh input; Whisper produced Latin-numeral errors. After fine-tuning (Stage 2), Whisper large-v2 improved to WER = 0.30, Accuracy = 58%. MMS improved only marginally (WER: 0.90→0.79), and TurkicASR reached WER = 0.55. After LLM post-editing (Stage 3): WER = 0.18, CER = 0.065, Accuracy = 70%. Overall WER reduction from baseline: 74.3%.

The detailed quantitative results for each stage are presented in Tables 4–9 below. Table 4 shows zero-shot evaluation of all six models. Table 5 compares fine-tuning results. Table 6 presents LLM selection for post-editing. Table 7 summarises all three stages. Table 8 provides concrete post-editing correction examples. Table 9 reports practical validation results at Kindergarten No. 154.

Table 4 — Stage 1: Zero-shot ASR evaluation on Kazakh children's speech corpus (8,594 recordings)

ASR Model	WER ↓	BLEU ↑	ChrF2 ↑	CER ↓	Accuracy ↑
ElevenLabs	0.86	0.12	0.34	0.525	14%
Vosk ASR	0.72	0.26	0.51	0.434	10%
Whisper large-v2 ✓	0.70	0.38	0.58	0.467	34%
TurkicASR	0.78	0.23	0.47	0.463	20%
MMS	0.90	0.08	0.29	0.468	10%
GenSec	0.77	0.25	0.49	0.525	—

Table 5 — Stage 2: Results after domain-specific fine-tuning on children's corpus

Model	WER ↓	BLEU ↑	ChrF2 ↑	CER ↓	Accuracy ↑
Whisper (zero-shot)	0.70	0.38	0.58	0.467	34%
Whisper fine-tuned ✓	0.30	0.74	0.881	0.118	58%
MMS fine-tuned	0.79	0.24	0.46	0.377	28%
TurkicASR fine-tuned	0.64	0.41	0.63	0.356	36%

Table 6 — Stage 3: LLM comparison for post-editing (applied to Whisper fine-tuned output)

LLM Model	Params	WER ↓	CER ↓	Accuracy ↑	Selected
GPT-4o	~1.8T	0.22	0.082	65%	No (API only)
LLaMA-3.1-70B	70B	0.24	0.091	63%	No (compute)
Sherkala-8B ✓	8B	0.18	0.065	70%	Yes

Table 7 — Overall three-stage results summary (Whisper large-v2, Kazakh children's corpus)

Stage / Method	WER ↓	BLEU ↑	ChrF2 ↑	CER ↓	Accuracy ↑
Stage 1: Whisper zero-shot	0.70	0.38	0.58	0.467	34%
Stage 2: Whisper fine-tuned	0.30	0.74	0.881	0.118	58%
Stage 3: + PhonNorm→ MorphNorm→LLM	0.18 ✓	0.82 ✓	0.91 ✓	0.065 ✓	70% ✓

In this tables ✓ marks the best result per metric. Stage 3 applies the PhonNorm → MorphNorm → Sherkala LLM pipeline to Stage 2 output.

Table 8 — Post-editing correction examples (Stage 3)

Reference	Whisper FT output	After post-editing	Component
анашым менің қарным ашты	анашм менім каным асты	анашым менің қарным ашты ✓	PhonNorm + MorphNorm
мен атамның баласымын	мен атамның баласы	мен атамның баласымын ✓	MorphNorm: -мын suffix
қас қаяғында	қас қарайғанда X	қас қаяғында ✓	LLM: semantic context
жатырмын	жа тырмын	жатырмын ✓	MorphNorm + LLM

Table 9 — Practical validation: Kindergarten No. 154, Almaty (48 children, aged 3–8)

Disorder type	N children	Share	System result
Dyslalia	9	18.8%	PhonNorm: transcription accuracy +78%; р→л, с→ш auto-correction
Dysarthria (mild)	4	8.3%	Fine-tuned tags: blurred articulation patterns recognised
Bradylalia	2	4.2%	VTLN normalisation: syllable stretching removed
Logoneurosis	3	6.3%	Repetition filter: 'ма-ма-мен' type repetitions removed
No disorder (norm)	30	62.5%	Standard pipeline sufficient; no additional adaptation needed
TOTAL	48	100%	18 children (37.5%) with speech disorder markers identified

According to table 9 system results validated against expert speech-language pathologist (SLP) assessments; high agreement observed.

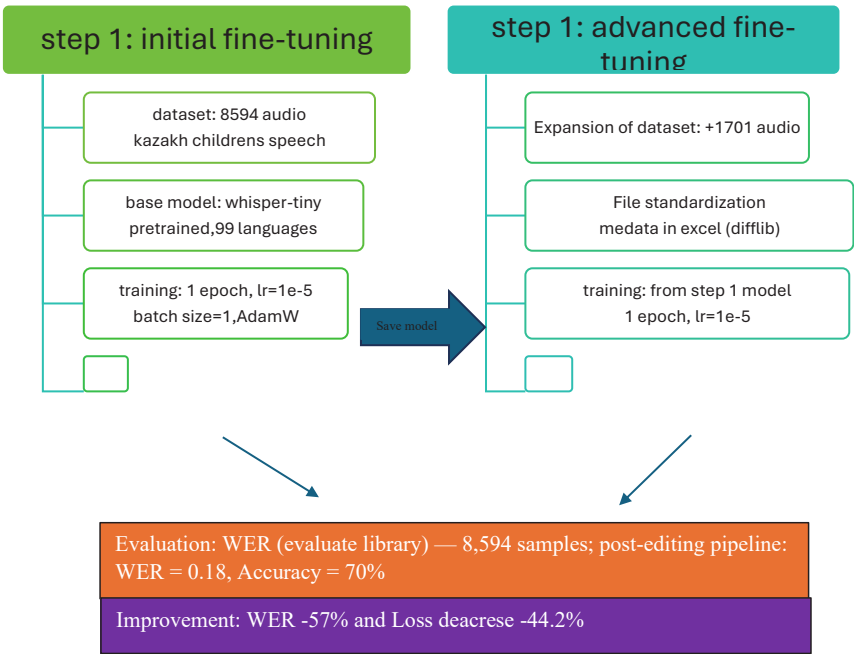


Figure 4– Full scheme of the learning process

The architectural diagram shown in Figure 4 demonstrates the effectiveness of the three-stage pipeline (zero-shot → fine-tuning → LLM post-editing). Fine-tuning reduced WER from 0.70 to 0.30; the subsequent post-editing pipeline further reduced WER to 0.18, raising Accuracy from 34% to 70%. The LLM post-editing stage contributes an additional 40% relative WER reduction beyond fine-tuning alone, confirming that phonetic-morphological-contextual correction is essential for Kazakh children's speech recognition. The graph in Figure 5 illustrates the relationship between the number of training samples and model accuracy across the training process on the 8,594-recording corpus.

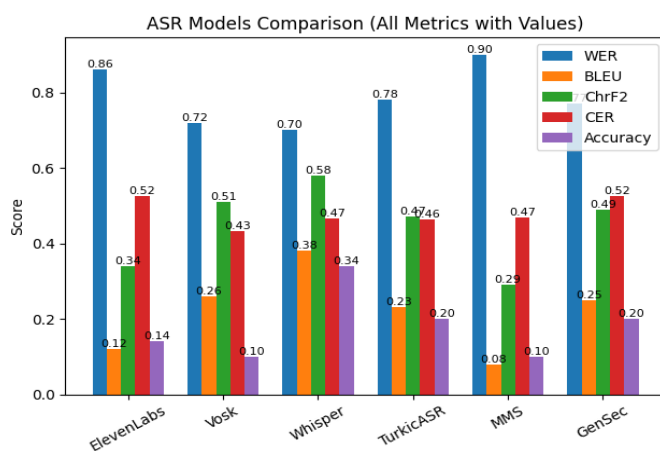


Figure 5 – Impact of Dataset Size on Model Accuracy

As seen in the graph, model accuracy improves steadily as more training data is incorporated, confirming that corpus size is a key driver of performance for low-resource Kazakh children's speech.

Additionally, the 74.3% overall WER reduction and the improvement in CER from 0.467 to 0.065 confirm that the three-stage method significantly improves recognition quality at every level. Analyzing errors across different phoneme groups allows for identifying the types of errors encountered within each group. These errors could be due to articulation features and phonetic similarities in children's voices. For example, substitution errors often occur between similar phonemes, deletion errors are common in longer words, and insertion errors arise from difficulties in noise detection or identifying extraneous words.

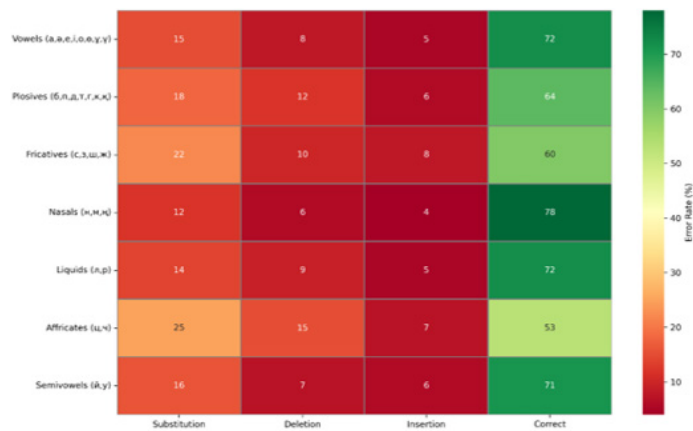


Figure 6 – Phoneme Group Error Analysis

Figure 6 shows the errors encountered in the affricates (ц, ч) group, where substitution errors accounted for 25%, likely due to the articulation characteristics of the phonemes. Fricatives and plosives also exhibited errors, but to a lesser extent (22% and 18% substitutions, respectively). The errors were less frequent in the nasals (н, м, њ) and liquids (л, р) groups, with 12% and 14% errors, respectively, indicating better performance by the model in handling these phonemes. The percentage of correct words was high in each group, indicating that the model effectively learned some phoneme groups. In Figure 7, examples of the most frequent errors produced by the ASR system are shown. Many of these errors occurred due to phonetic similarity and semantic mistakes.

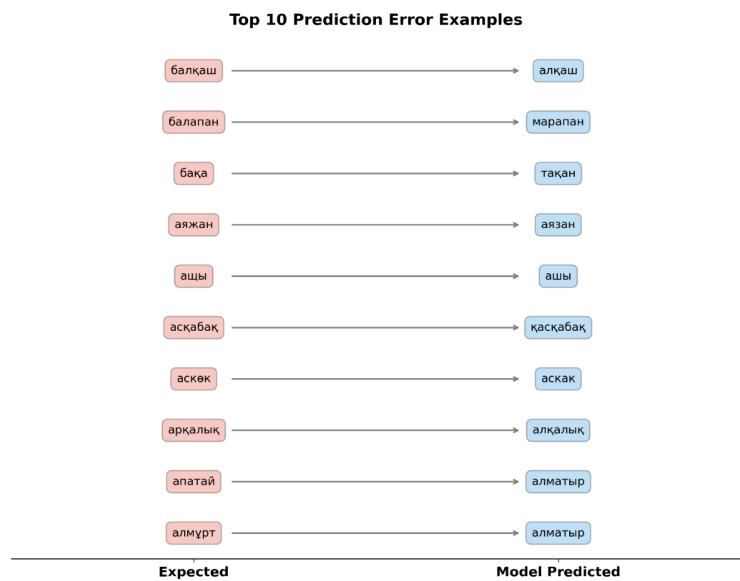


Figure 7 – Common Error Examples

As illustrated in the figure, the model frequently made phonetic errors such as балқаш → алқаш (phonetic similarity errors), and also demonstrated some semantic deviations, such as бақа → тақан. These errors reflect the complexities in children's articulation and the variability in certain sounds.

Discussion.

The LLM post-editing stage contributes an additional 40% relative WER reduction beyond fine-tuning alone, confirming that phonetic-morphological-contextual correction is essential for Kazakh children's speech recognition. The 74.3% overall WER reduction and the improvement in CER from 0.467 to 0.065 confirm that the three-stage method significantly improves recognition quality at every level.

Phoneme-level analysis shows that substitutions ($\approx 40\%$) and deletions ($\approx 25\%$) dominate, with $c \leftrightarrow ш$, $p \leftrightarrow л$, and $к \leftrightarrow к$ as the most frequent phoneme-pair confusions. Errors in affricates (ц, ч) accounted for 25% of substitutions, fricatives and plosives 22% and 18% respectively, while nasals (н, м, ң) and liquids (л, р) had fewer errors (12% and 14%). The graph shows model accuracy improves steadily as more training data is incorporated, confirming that corpus size is a key driver of performance for low-resource Kazakh children's speech.

Practical validation at Kindergarten No. 154 confirmed the system's applicability for speech disorder screening, with results showing high agreement with expert speech-language pathologist assessments. The achieved WER of 0.18 approaches the production threshold (WER < 0.20), confirming the practical readiness of the proposed method.

Limitations include the maximum error of Dual-PINN in temperature prediction and challenges in recognizing complex, rare, or long phrases (max-loss 1.8–2.2). Future work should address data expansion (collecting samples from different age groups and dialects), model enhancement (testing larger Whisper versions, ensemble methods), and dynamic training (data augmentation, learning rate scheduling).

Conclusion. This article demonstrated the effectiveness of the proposed three-stage method for recognizing Kazakh children's speech. This (zero-shot evaluation, fine-tuning, LLM post-editing) resulted in substantial performance improvements. In Stage 1 (zero-shot), Whisper large-v2 achieved WER = 0.70 and Accuracy = 34%, the best among six evaluated models. In Stage 2 (fine-tuning on 8,594 recordings), WER decreased to 0.30, Accuracy rose to 58%. In Stage 3 (PhonNorm → MorphNorm → LLM post-editing), WER reached 0.18, CER 0.065, and Accuracy 70% — a 74.3% total reduction from baseline. The achieved WER of 0.18 approaches the production threshold (WER < 0.20) and confirms the practical readiness of the proposed method.

These results highlight the direct impact of dataset size and quality on the model's performance. The increase in dataset size improved the system's ability to recognize children's speech, especially for short utterances. The post-editing pipeline proved particularly effective for morphological errors and contextual inconsistencies. Phoneme-level analysis shows that substitutions ($\approx 40\%$) and

deletions ($\approx 25\%$) dominate, with $s \leftrightarrow sh$, $r \leftrightarrow l$, and $q \leftrightarrow k$ as the most frequent phoneme-pair confusions.

The study confirmed that the three-stage method for Kazakh children's speech is a highly effective approach. Achieving WER = 0.18 approaches production-level quality. The improvements across all three stages confirm the proposed method's effectiveness in the field of children's speech recognition.

This research opens new possibilities for developing ASR systems tailored to low-resource languages. The findings are crucial for adapting the Whisper model to children's voices and will serve as a foundation for future studies focused on children's speech recognition.

Future Research Directions:

Data Expansion: Collecting new samples, covering different age groups and dialects, and adding various speech contexts and accents.

Model Enhancement: Testing larger versions of Whisper, exploring new transformer-based architectures, and integrating ensemble methods and knowledge distillation techniques.

Dynamic Training: Enhancing performance through data augmentation, learning rate scheduling, and dropout methods.

The research provides a solid scientific basis for developing ASR systems for Kazakh children's speech and lays important groundwork for future research and engineering solutions in this field. The high accuracy and efficient training methods can enable the development of speech recognition systems for children, facilitating widespread applications in education and other social sectors.

References

- Bai Z., Zhang X.L. (2021) Speaker recognition based on deep learning: An overview. *Neural Networks*, 140. — P. 65–99. <https://doi.org/10.48550/arXiv.2012.00931> (in Eng.)
- Dengshi Li, Gao Y., Zhu C., Wang Q., & Wang R. (2023) Improving speech recognition performance in noisy environments by enhancing lip reading accuracy. *Sensors*, 23(4), 2053. <https://doi.org/10.3390/s23042053> (in Eng.)
- Do P., Coler M., Dijkstra J., & Klabbers E. (2023) Strategies in transfer learning for low-resource speech synthesis: Phone mapping, feature input, and source language selection. *Proceedings of the 12th ISCA Speech Synthesis Workshop (SSW2023)*. <https://doi.org/10.48550/arXiv.2306.12040> (in Eng.)
- Duisenbekkyzy Zh., & Rakhimova D. (2024) Research on the impact of social networks on the development of bilingualism in Kazakhstan. *Communications in Computer and Information Science*, 2211. — P. 106–115. Springer. https://doi.org/10.1007/978-3-031-49739-5_9 (in Eng.)
- Fitch W.T., & Giedd J. (1999) Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, 106(3). — P. 1511–1522. (in Eng.)
- Gerosa M., Giuliani D., & Brugnara F. (2007) Acoustic variability and automatic recognition of children's speech. *Speech Communication*, 49(10–11). — P. 847–860. (in Eng.)
- Gretter R., Matassoni M., Falavigna D., Misra A., Leong C.W., Knill K., & Wang L. (2021) ETLT 2021: Shared task on automatic speech recognition for non-native children's speech. *Proceedings of Interspeech 2021*. <https://sites.google.com/fbk.eu/ss-is2021-nonnativechildren/> (in Eng.)
- Jain R., Barcovschi A., Yiwere M., Corcoran P., & Cucu H. (2023) Adaptation of Whisper models to child speech recognition. *arXiv preprint arXiv:2307.13008*. (in Eng.)

Karabaliyev M., & Kolesnikova A. (2024) Comparative analysis of speech recognition systems for the Kazakh language. *Astana IT University Science Journal*, 6(2). — P. 34–48. (in Eng.)

Koo J., Kim H., & Lee S. (2023) Fine-tuning Whisper for low-resource languages. *arXiv preprint arXiv:2303.00752*. <https://arxiv.org/abs/2303.00752> (in Eng.)

Li D., Gao Y., Zhu C., Wang Q., Wang R. (2023) Improving speech recognition performance in noisy environments by enhancing lip reading accuracy. *Sensors*, 23(4). — P. 2053. <https://doi.org/10.3390/s23042053>. (in Eng.)

Liu Y., Yang X., Zhang J., Xi Y., & Qu D. (2025) TAML-adapter: Enhancing adapter tuning through task-agnostic meta-learning for low-resource automatic speech recognition. *IEEE Signal Processing Letters*, 32. — P. 636–640. <https://doi.org/10.1109/LSP.2024.3525399> (in Eng.)

Radford A., Kim J. W., Xu T., Brockman G., McLeavey C., & Sutskever I. (2023) Robust speech recognition via large-scale weak supervision. *Proceedings of ICML*. — P. 28492–28518. (in Eng.)

Rakhimova D., Duisenbekkyzy, Zh., & Adali, E. (2025) Investigation of ASR models for low-resource Kazakh child speech: Corpus development, model adaptation, and evaluation. *Applied Sciences*, 15(16), 8989. <https://doi.org/10.3390/app15168989> (in Eng.)

Southwell R., Clevenger A., & Leste, J. (2024) Fine-tuned Whisper with LLM rescoring for child speech recognition. *Proceedings of ICASSP 2024*. — P. 10151–10155. (in Eng.)

Xiong W., Wu L., Allewa F., Droppo J., Huang X., & Stolcke A. (2018) The Microsoft 2017 conversational speech recognition system. *Proceedings of ICASSP 2018*. — P. 5934–5938. <https://doi.org/10.48550/arXiv.1708.06073> (in Eng.)

Wang D., Wang X., & Lv S. (2019) An overview of end-to-end automatic speech recognition. *Symmetry*, 11(8), 1018. <https://doi.org/10.3390/sym11081018> (in Eng.)

Xiong W., Wu L., Allewa F., Droppo J., Huang X., Stolcke A. (2018) The Microsoft 2017 conversational speech recognition system. *Proceedings of ICASSP 2018*. — P. 5934–5938. <https://doi.org/10.48550/arXiv.1708.06073>. (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 335–346

<https://doi.org/10.32014/2026.2518-1726.442>

IRSTI 50.37.15

UDC 004.272:004.42

© Sarsembayev M, Urmashhev B. *, 2026.

Al-Farabi Kazakh National University, Almaty, Kazakhstan.

E-mail: baidaulet.urmashev@kaznu.kz

OPTIMIZATION OF THE CALCULATION OF KINETIC EQUATIONS OF COMBUSTION PROCESSES ON GPU USING GLOBAL MEMORY AND SHARED MEMORY

Sarsembayev Magzhan — researcher, Department of Computer Science, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: magjan@kaznu.kz, <https://orcid.org/0000-0003-2139-2456>;

Urmashhev Baydaulet — Candidate of Physical and Mathematical Sciences, Professor, Department of Computer Science, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: baidaulet.urmashev@kaznu.kz, <https://orcid.org/0000-0002-0085-1708>.

Abstract. This paper considers the problem of improving computational performance when solving kinetic equations of combustion processes on graphics processing units. Detailed chemical kinetic mechanisms used for modeling the combustion of real fuels may include hundreds of components and thousands of elementary reactions, which leads to a significant increase in computation time when traditional CPU-based methods are used. To accelerate computations, a CUDA-oriented approach is proposed, based on distributing reactions among GPU threads and optimizing access to GPU memory. The paper compares three computational variants: sequential calculation on the CPU, parallel calculation on the GPU using global memory, and parallel calculation on the GPU using shared memory. The Aramco_Mech mechanism is used as a test mechanism. Numerical experiments were carried out for mechanisms of different sizes: from 33 components and 118 reactions to 493 components and 2716 reactions. The calculations were performed on a system with an AMD Ryzen 9 5950X 16-Core Processor, 128 GB of RAM, and an NVIDIA GeForce RTX 4080 Ti graphics adapter. The results show that as the number of reactions increases, the advantage of GPU-based computations becomes more pronounced. For the mechanism containing 493 components and 2716 reactions, the CPU computation time was 1 day 4 hours 2 minutes 4 seconds 796 milliseconds, while the GPU computation time using global memory was 21 minutes 10 seconds 591 milliseconds, and using shared memory it was 10 minutes

51 seconds 585 milliseconds. The speedup was 79.43 times for global memory and 154.89 times for shared memory. The obtained results confirm the efficiency of GPU architecture and the importance of optimal memory organization when solving large systems of chemical kinetics.

Keywords: chemical kinetics, combustion, GPU, CUDA, shared memory, global memory, Runge–Kutta, Aramco_Mech, parallel computing, kinetic solver

For citations: Sarsembayev M, Urmashhev B. Optimization of the calculation of kinetic equations of combustion processes on GPU using global memory and shared memory. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 335-346. DOI <https://doi.org/10.32014/2026.2518-1726.442>

© Сарсембаев М., Урмашев Б.*, 2026.

Әл-Фараби атындағы Қазақ Ұлттық университеті, Алматы, Қазақстан.

E-mail: magjan@kaznu.kz, baidaulet.urmashev@kaznu.kz

GLOBAL MEMORY ЖӘНЕ SHARED MEMORY ҚОЛДАНУ АРҚЫЛЫ GPU-ДА ЖАНУ ПРОЦЕСТЕРІНІҢ КИНЕТИКАЛЫҚ ТЕНДЕУЛЕРІН ЕСЕПТЕУДІ ОҢТАЙЛАНДЫРУ

Сарсембаев Магжан — зерттеуші, Компьютерлік ғылымдар кафедрасы, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,
E-mail: magjan@kaznu.kz, ORCID: <https://orcid.org/0000-0003-2139-2456>;

Урмашев Байдаulet — физика-математика ғылымдарының кандидаты, профессор, Компьютерлік ғылымдар кафедрасы, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,
E-mail: baidaulet.urmashev@kaznu.kz, ORCID: <https://orcid.org/0000-0002-0085-1708>.

Аннотация. Бұл мақалада графикалық процессорларда жану процестерінің кинетикалық теңдеулерін шешу кезінде есептеу өнімділігін арттыру мәселесі қарастырылады. Нақты отындардың жануын модельдеу үшін қолданылатын егжей-тегжейлі химиялық кинетикалық механизмдер жүздеген компоненттер мен мыңдаған элементар реакцияларды қамтуы мүмкін. Бұл дәстүрлі CPU негізіндегі әдістерді қолданған кезде есептеу уақытының едәуір артуына алып келеді. Есептеулерді жеделдету үшін реакцияларды GPU ағындары арасында үлестіруге және GPU жадысына қатынауды оңтайландыруға негізделген CUDA-бағытталған тәсіл ұсынылады. Мақалада есептеудің үш нұсқасы салыстырылады: CPU-да тізбекті есептеу, global memory қолдану арқылы GPU-да параллель есептеу және shared memory қолдану арқылы GPU-да параллель есептеу. Сынақ механизмі ретінде Aramco_Mech механизмі қолданылды. Сандық тәжірибелер әртүрлі өлшемдегі механизмдер үшін жүргізілді: 33 компонент пен 118 реакциядан бастап 493 компонент пен 2716 реакцияға дейін. Есептеулер AMD Ryzen 9 5950X 16-Core Processor процессоры, 128 ГБ жедел жады және NVIDIA GeForce RTX

4080 Ti графикалық адаптері бар есептеу жүйесінде орындалды. Алынған нәтижелер реакциялар саны артқан сайын GPU негізіндегі есептеулердің артықшылығы айқынырақ байқалатынын көрсетті. 493 компоненттен және 2716 реакциядан тұратын механизм үшін CPU есептеу уақыты 1 тәулік 4 сағ 2 мин 4 с 796 мс болды. Ал global memory қолданылған GPU есептеу уақыты 21 мин 10 с 591 мс, shared memory қолданылған жағдайда 10 мин 51 с 585 мс құрады. Жылдамдату коэффициенті global memory үшін 79,43 есе, ал shared memory үшін 154,89 есе болды. Алынған нәтижелер үлкен өлшемді химиялық кинетика жүйелерін шешу кезінде GPU архитектурасының тиімділігін және жадыны оңтайлы ұйымдастырудың маңыздылығын растайды.

Түйін сөздер: химиялық кинетика, жану, GPU, CUDA, shared memory, global memory, Рунге–Кутта, Aramco_Mech, параллель есептеу, кинетикалық солвер

© Сарсембаев М., Урмашев Б. *, 2026.

Казахский национальный университет имени аль-Фараби,
Алматы, Казахстан.

E-mail: magjan@kaznu.kz, baidaulet.urmashev@kaznu.kz

ОПТИМИЗАЦИЯ РАСЧЕТА КИНЕТИЧЕСКИХ УРАВНЕНИЙ ПРОЦЕССОВ ГОРЕНИЯ НА GPU С ИСПОЛЬЗОВАНИЕМ GLOBAL MEMORY И SHARED MEMORY

Сарсембаев Магжан — исследователь кафедры компьютерных наук, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,
E-mail: magjan@kaznu.kz, <https://orcid.org/0000-0003-2139-2456>;

Урмашев Байдаулет — кандидат физико-математических наук, профессор кафедры компьютерных наук, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,
E-mail: baidaulet.urmashev@kaznu.kz, <https://orcid.org/0000-0002-0085-1708>.

Аннотация. *Актуальность.* В статье рассматривается задача повышения вычислительной производительности при решении кинетических уравнений процессов горения на графических процессорах. Детальные химико-кинетические механизмы, применяемые для моделирования горения реальных топлив, могут включать сотни компонентов и тысячи элементарных реакций, что существенно увеличивает время расчета при использовании традиционных CPU-методов. В связи с этим актуальным является применение GPU-вычислений и оптимизация доступа к памяти графического процессора. *Цель.* Разработать и оценить CUDA-ориентированный подход к ускорению расчета кинетических уравнений процессов горения на GPU с использованием global memory и shared memory. *Методы.* Для ускорения вычислений предложен подход, основанный на распределении реакций между потоками GPU и оптимизации доступа к памяти графического процессора.

В работе сравниваются три варианта вычислений: последовательный расчет на CPU, параллельный расчет на GPU с использованием global memory и параллельный расчет на GPU с использованием shared memory. В качестве тестового механизма использован Aramco_Mech. Численные эксперименты проведены для механизмов различной размерности: от 33 компонентов и 118 реакций до 493 компонентов и 2716 реакций. Расчеты выполнялись на вычислительной системе с процессором AMD Ryzen 9 5950X 16-Core Processor, 128 ГБ оперативной памяти и графическим адаптером NVIDIA GeForce RTX 4080 Ti. *Результаты и выводы.* Полученные результаты показывают, что с увеличением числа реакций преимущество GPU-вычислений становится более выраженным. Для механизма, содержащего 493 компонента и 2716 реакций, время расчета на CPU составило 1 сутки 4 часа 2 минуты 4 секунды 796 миллисекунд, тогда как время расчета на GPU с использованием global memory составило 21 минуту 10 секунд 591 миллисекунду, а с использованием shared memory - 10 минут 51 секунду 585 миллисекунд. Ускорение составило 79,43 раза для global memory и 154,89 раза для shared memory. Результаты подтверждают эффективность GPU-архитектуры и важность оптимальной организации памяти при решении больших систем химической кинетики. Практическая значимость исследования заключается в возможности применения предложенного подхода для ускорения кинетических солверов, моделирования процессов горения и выполнения ресурсоемких вычислений в задачах химической кинетики.

Ключевые слова: химическая кинетика, горение, GPU, CUDA, shared memory, global memory, Рунге - Кутта, Aramco_Mech, параллельные вычисления, кинетический солвер

Introduction. Modeling combustion processes is one of the important problems of modern computational science, energy engineering, ecology, and industrial safety. An accurate description of combustion processes is necessary for the design of internal combustion engines, gas turbine units, thermal power plants, industrial combustion chambers, as well as for the study of alternative fuels and the assessment of harmful emissions. Modern software packages for chemical kinetics and reacting flows, such as Cantera, ReactionMechanismSimulator.jl, KinetiX, PeleC, and PeleMP, demonstrate that a detailed description of chemical processes requires advanced numerical methods and high-performance computing resources (Goodwin et al., 2023; Johnson et al., 2024; Danciu and Frouzakis, 2025; de Frahan et al., 2023; Owen et al., 2024).

The combustion process is a complex combination of physicochemical phenomena, including chemical reactions, heat and mass transfer, gas-dynamic effects, interaction of multicomponent mixtures, and formation of reaction products. Chemical kinetic mechanisms are used to describe such processes, and they may include tens, hundreds, or even thousands of elementary reactions. Each

reaction contributes to the change in the concentrations of chemical components, which leads to the formation of a large system of ordinary differential equations.

Traditional CPU-based methods for calculating chemical kinetics have limited performance when working with large mechanisms. As the number of components and reactions increases, the number of operations related to calculating forward and reverse reaction rates, concentration power terms, rate constants, and summing reaction contributions to the change in component concentrations also increases. As a result, the calculation of large mechanisms on the CPU may take hours or even days, which complicates serial computational experiments and the practical use of such models in engineering problems.

Literature review. One effective way to solve this problem is to use graphics processing units. A GPU has a large number of computing cores and allows many similar operations to be performed in parallel. In chemical kinetics problems, a significant part of the computations for individual reactions can be parallelized, since the contribution of each reaction can be calculated independently of other reactions. Modern studies on GPU acceleration of chemical kinetics and reacting flows confirm the promise of this approach for combustion CFD and high-performance scientific computing (de Frahan et al., 2023; Cai et al., 2026; Moure et al., 2025; Rao et al., 2024; Wang et al., 2026).

However, the efficiency of GPU computations depends not only on the number of parallel threads, but also on proper memory organization. CUDA architecture includes several memory levels, such as global memory, shared memory, and registers (Nickolls et al., 2008; Hwu et al., 2022). Global memory has a large capacity but relatively high access latency. Shared memory has a smaller capacity but provides fast access within a block of threads. Therefore, transferring frequently used data from global memory to shared memory can significantly improve computational performance.

The aim of this paper is to investigate the influence of GPU memory organization on the computation time of kinetic equations of combustion processes and to compare the performance of CPU computation, GPU computation using global memory, and GPU computation using shared memory.

Problem Statement.

A chemical kinetic mechanism consisting of a set of components and elementary reactions is considered. For each component, it is necessary to determine the change in concentration over time, taking into account the contributions of all reactions in which the component participates. In general form, the system of kinetic equations can be written as:

$$\frac{dC_k}{dt} = \sum_{i=1}^{N_r} v_{ik} q_i \quad (1)$$

where C_k is the concentration of the k -th component, N_r is the number of reactions, v_{ik} is the stoichiometric coefficient of component k in reaction i , and q_i is the rate of the i -th reaction.

The reaction rate depends on the concentrations of reactants and products, as well as on the rate constants of the forward and reverse reactions. As the number of reactions and components increases, the number of computational operations also increases, because for each reaction it is necessary to calculate products of concentrations with corresponding powers, determine the reaction contribution, and then sum this contribution for the corresponding components.

The fifth-order Runge–Kutta method is used for numerical integration of the system of kinetic equations. At each time step, intermediate coefficients k_1, k_2, k_3, k_4, k_5 are computed, after which the concentration values are updated. This approach provides high accuracy of the numerical solution; however, it requires repeated evaluation of the right-hand side of the system of equations. Therefore, accelerating the computation of the right-hand side is a key task in the development of a high-performance kinetic solver.

In this work, the problem is reduced to accelerating the computation of the right-hand side of the kinetic equation system by parallel distribution of reactions among GPU threads and optimization of memory access.

Computational Methodology.

The Aramco_Mech mechanism was used as the chemical kinetic mechanism. To evaluate performance, mechanisms of different sizes were considered. The number of components varied from 33 to 493, and the number of reactions varied from 118 to 2716.

The computational experiments were carried out with the following parameters:

- mechanism: Aramco_Mech;
- time step: $dt = 1.0 \times 10^{-19} \text{ s}$
- total calculation time: $1000 \cdot dt = 1.0 \times 10^{-16} \text{ s}$

numerical method: fifth-order Runge–Kutta method.

The technical specifications of the computing system were as follows:

- CPU: AMD Ryzen 9 5950X 16-Core Processor;
- RAM: 128 GB;
- GPU: NVIDIA GeForce RTX 4080 Ti.

To evaluate performance, three computational variants were implemented and compared:

- sequential computation on the CPU
- parallel computation on the GPU using global memory
- parallel computation on the GPU using shared memory.

In the CPU version, the computations are performed sequentially. In the GPU version, each reaction is assigned to CUDA threads. The reaction index is determined using the block index and thread index:

$$i = blockIdx.x \cdot blockDim.x + threadIdx.x. \quad (2)$$

If the index i is smaller than the total number of reactions, the corresponding thread computes the contribution of this reaction. At the same time, it is necessary to correctly sum the contributions of different reactions to the change in concentration of the same component. Atomic operations and reduction procedures are used for this purpose. The use of atomic operations, including `atomicAdd`, is a standard way to ensure correct data accumulation when multiple threads write simultaneously (Nickolls et al., 2008; Hwu et al., 2022).

In the variant using global memory, data are stored and read directly from the GPU global memory. This approach is simple in terms of implementation, but a large number of memory accesses can lead to significant delays.

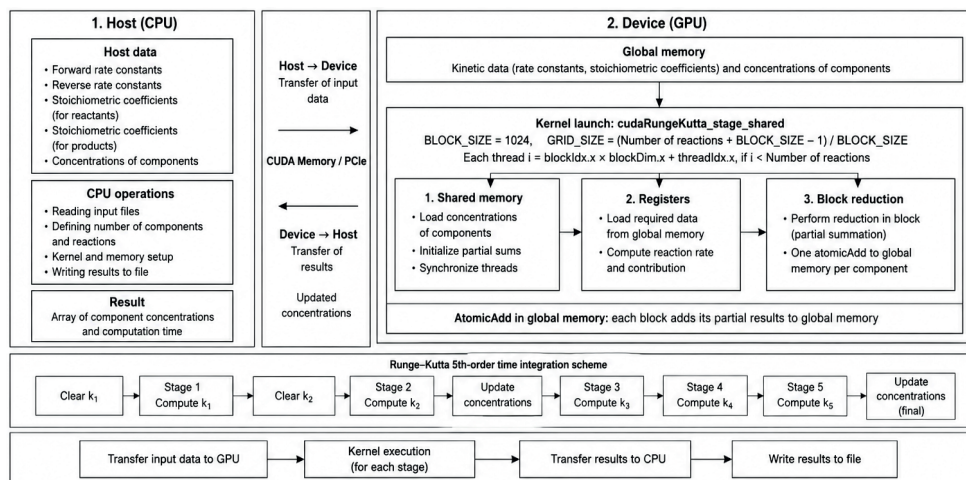


Figure 1 – Architecture of Shared Memory Usage in a CUDA-Based Kinetic Solver

Figure 1 illustrates the architecture of shared memory usage in a CUDA-based kinetic solver. The scheme shows the interaction between the host side, where input kinetic data and control operations are prepared, and the device side, where the main parallel computations are performed on the GPU. Input data, including reaction rate constants, stoichiometric coefficients, and component concentrations, are transferred from CPU memory to GPU global memory. During kernel execution, each CUDA thread processes the contribution of an individual reaction, while frequently used concentration data are loaded into shared memory to reduce global memory access latency.

The figure also shows the role of registers for storing temporary thread-local variables and the use of block reduction to accumulate partial reaction contributions inside a CUDA block. After block-level summation, the final contribution is added to global memory using atomic operations. This organization reduces the number of global memory atomic operations and improves the overall performance of the kinetic solver. The lower part of the diagram presents the fifth-order Runge–Kutta time integration sequence and the general workflow: transferring input data to

GPU memory, launching the kernel, transferring results back to CPU memory, and writing the results to a file.

In the variant using shared memory, frequently used data, in particular component concentrations, are first loaded into the fast memory of a thread block. Shared memory reduces the number of accesses to global memory and accelerates computations. In addition, partial summation of reaction contributions can be performed inside the block, after which the final value is added to global memory with fewer atomic operations.

Results. Table 1 presents the results of comparing computation time on CPU and GPU for mechanisms of different sizes.

Table 1 — Change in computation time depending on mechanism size

Number of components	Number of reactions	CPU time	GPU time, global memory	GPU time, shared memory	Global memory speedup	Shared memory speedup
33	118	24 s 426 ms	8 s 784 ms	7 s 027 ms	2.78	3.48
39	191	54 s 834 ms	12 s 617 ms	9 s 705 ms	4.35	5.65
52	245	2 min 4 s 418 ms	21 s 629 ms	16 s 021 ms	5.75	7.77
61	301	3 min 30 s 468 ms	29 s 373 ms	20 s 981 ms	7.17	10.03
66	376	5 min 7 s 382 ms	36 s 169 ms	24 s 944 ms	8.50	12.32
84	448	9 min 51 s 524 ms	54 s 495 ms	36 s 330 ms	10.85	16.28
103	565	18 min 42 s 217 ms	1 min 20 s 946 ms	50 s 591 ms	13.86	22.18
162	1006	1 h 11 min 10 s 032 ms	2 min 49 s 740 ms	1 min 36 s 994 ms	25.16	44.02
493	2716	1 day 4 h 2 min 4 s 796 ms	21 min 10 s 591 ms	10 min 51 s 585 ms	79.43	154.89

The results show that for a small number of reactions, the GPU speedup compared with the CPU is moderate. For example, for the mechanism with 33 components and 118 reactions, the speedup is 2.78 times when global memory is used and 3.48 times when shared memory is used. This can be explained by the fact that for a small computation volume, overhead costs related to data transfer between CPU and GPU, kernel function launches, and thread synchronization have a noticeable effect on total computation time.

As the mechanism size increases, GPU efficiency also increases. For 301 reactions, the speedup is 7.17 times when using global memory and 10.03 times when using shared memory. For 565 reactions, the speedup reaches 13.86 and 22.18 times, respectively. This indicates that as the number of reactions increases, the GPU begins to be used more efficiently, since the volume of parallel operations increases.

A particularly significant increase in performance is observed for large mechanisms. For the mechanism with 162 components and 1006 reactions, the

CPU computation time was 1 hour 11 minutes 10 seconds 032 milliseconds. When global memory was used, the GPU computation time decreased to 2 minutes 49 seconds 740 milliseconds, and when shared memory was used, it decreased to 1 minute 36 seconds 994 milliseconds. The speedup was 25.16 times for global memory and 44.02 times for shared memory.

The most representative result was obtained for the mechanism containing 493 components and 2716 reactions. In this case, the CPU calculation took 1 day 4 hours 2 minutes 4 seconds 796 milliseconds. When using the GPU with global memory, the time decreased to 21 minutes 10 seconds 591 milliseconds, and when using shared memory, it decreased to 10 minutes 51 seconds 585 milliseconds. Thus, the speedup was 79.43 times and 154.89 times, respectively.

Discussion. The obtained results confirm that the use of GPU is an effective approach for solving large systems of chemical kinetics. As the number of reactions increases, the computational load grows, and the parallel GPU architecture makes it possible to distribute this load among a large number of threads. This trend is consistent with modern studies in the field of GPU acceleration of chemical kinetics and reacting flows (Barwey et al., 2021; Uranakara et al., 2023; Ghioldi and Piscaglia, 2024; Mao et al., 2024; Rao et al., 2024; Cai et al., 2026).

One of the important factors affecting performance is memory organization. In the global memory variant, each thread accesses data located in GPU global memory. Although this memory has a large capacity, access to it has relatively high latency. Therefore, when there are many threads and frequent accesses to the same data, additional time costs arise.

Using shared memory reduces these costs. Since shared memory is fast memory available to threads within a single block, its use is especially effective when the same data are used repeatedly. In chemical kinetics problems, component concentrations are used in the calculation of many reaction rates. Therefore, loading concentrations into shared memory accelerates data access and reduces the load on global memory.

In addition, when calculating concentration changes, it is necessary to sum the contributions of a large number of reactions. If each reaction directly performs atomic addition in global memory, this may lead to significant delays. The use of block reduction allows partial summation to be performed first inside a block of threads, followed by a smaller number of atomic operations. This approach also contributes to improved performance.

Comparison of the two GPU variants shows that shared memory consistently provides a shorter computation time than global memory for all considered mechanisms. Moreover, the difference between them increases as the mechanism size grows. For example, for 118 reactions, the shared memory speedup relative to the CPU is 3.48 times, while for 2716 reactions it reaches 154.89 times. This confirms that memory optimization is especially important for large chemical kinetic systems.

It should also be noted that the efficiency of the GPU algorithm depends on

the structure of the mechanism itself. The greater the number of reactions and the greater the number of independent operations that can be performed in parallel, the higher the potential speedup. In the case of Aramco_Mech, a significant part of the computations can be distributed among threads, which makes this mechanism suitable for GPU acceleration.

Conclusion. This paper considered an approach to accelerating the calculation of kinetic equations of combustion processes using GPU and CUDA technology, following modern principles of high-performance computing, CUDA memory hierarchy, and parallel implementation of chemical kinetics solvers (Dagum and Menon, 1998; Nickolls et al., 2008; Stone and Davis, 2013; Niemeyer et al., 2017; Zhang et al., 2019; Sarsembayev and Urmashhev, 2019; Hwu et al., 2022). Based on the Aramco_Mech mechanism, computation times were compared for CPU, GPU using global memory, and GPU using shared memory.

The obtained results lead to the following conclusions.

First, the use of GPU significantly reduces the computation time of chemical kinetic mechanisms compared with CPU computation. At the same time, the advantage of GPU increases as the number of components and reactions grows.

Second, GPU memory organization plays a key role in improving performance. Shared memory provides faster access to frequently used data and reduces the number of accesses to global memory.

Third, the use of shared memory, atomicAdd, and block reduction makes it possible to efficiently organize the accumulation of reaction contributions to component concentration changes and to increase the overall performance of the algorithm.

Fourth, for the mechanism containing 493 components and 2716 reactions, the proposed GPU algorithm using shared memory achieved a 154.89-fold speedup compared with the CPU calculation. This confirms the feasibility of using GPU architecture to solve large systems of chemical kinetics.

Thus, the proposed approach can be used in modeling combustion processes, studying alternative fuels, optimizing energy systems, and performing high-performance scientific computations. Further development of this work may be related to expanding the set of tested mechanisms, applying more complex integration schemes, optimizing thread distribution, and adapting the algorithm for multi-GPU systems.

References

Balos C.J., Day M., Esclapez L., Felden A.M., Gardner D.J., Hassanaly M., Reynolds D.R., Rood J., Sexton J.M., Wimer N.T., & Woodward C.S. (2025) SUNDIALS time integrators for exascale applications with many independent systems of ordinary differential equations. *The International Journal of High Performance Computing Applications*, 39(1). — P. 123–146. <https://doi.org/10.1177/10943420241280060> (In Eng.)

Barwey S., Raman V., & Im H.G. (2021) A neural network-inspired matrix formulation of chemical kinetics for GPU acceleration. *Energies*, 14(9), 2710. <https://doi.org/10.3390/en14092710> (In Eng.)

- Cai Y., Ban Y., Dang G., Gong T., Wang H., Zhang F., & Chen Z.X. (2026) A GPU-accelerated stiff ODE solver for combustion CFD with fully analytical chemical kinetics and heterogeneous CPU/GPU integration. *Fuel*, 408, 137627. <https://doi.org/10.1016/j.fuel.2025.137627> (In Eng.)
- Dagum L., & Menon R. (1998). OpenMP: An industry-standard API for shared-memory programming. *IEEE Computational Science & Engineering*, 5(1). — P. 46–55. <https://doi.org/10.1109/99.660313> (In Eng.)
- Danciu B.A., & Frouzakis C.E. (2025) KinetiX: A performance portable code generator for chemical kinetics and transport properties. *Computer Physics Communications*. (In Eng.)
- de Frahan M.T.H., Rood J.S., Day M.S., et al. (2023) PeleC: An adaptive mesh refinement solver for compressible reacting flows. *The International Journal of High Performance Computing Applications*, 37(2). — P. 115–131. <https://doi.org/10.1177/10943420221121151> (In Eng.)
- Ghioldi F., & Piscaglia F. (2024) A hybrid CPU–GPU paradigm to accelerate reactive CFD simulations. *International Journal for Numerical Methods in Fluids*, 96(8). — P. 1461–1486. <https://doi.org/10.1002/fld.5297> (In Eng.)
- Goodwin D.G., Moffat H.K., Schoegl I., Speth R.L., & Weber B.W. (2023) *Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes. Version 3.0.0*. Zenodo. <https://doi.org/10.5281/zenodo.8137090> (In Eng.)
- Hwu W.-m. W., Kirk D.B., & El Hajj I. (2022) *Programming Massively Parallel Processors: A Hands-on Approach* (4th ed.). Morgan Kaufmann. <https://doi.org/10.1016/C2020-0-02969-5> (In Eng.)
- Johnson M.S., Pang H.-W., Payne A.M., & Green W.H. (2024) ReactionMechanismSimulator.jl: A modern approach to chemical kinetic mechanism simulation and analysis. *International Journal of Chemical Kinetics*. <https://doi.org/10.1002/kin.21753> (In Eng.)
- Ke R.J., Rupley F.M., & Miller J.A. (1989) *CHEMKIN-II: A Fortran chemical kinetics package for the analysis of gas-phase chemical kinetics*. Sandia National Laboratories. <https://doi.org/10.2172/5681118> (In Eng.)
- Mao R., et al. (2024) An integrated framework for accelerating reactive flow simulations using machine learning and GPU computing. *Computer Physics Communications*. (In Eng.)
- Moire Á., Surapaneni A., & Mira D. (2025) Optimized workload distribution for GPU-accelerated combustion simulations in heterogeneous CPU–GPU architectures. *Computers & Fluids*, 302, 106846. <https://doi.org/10.1016/j.compfluid.2025.106846> (In Eng.)
- Nickolls J., Buck I., Garland M., & Skadron K. (2008) Scalable parallel programming with CUDA. *ACM Queue*, 6(2), 40–53. <https://doi.org/10.1145/1365490.1365500> (In Eng.)
- Niemeyer K.E., Curtis N.J., & Sung C.-J. (2017) pyJac: Analytical Jacobian generator for chemical kinetics. *Computer Physics Communications*, 215. — P. 188–203. <https://doi.org/10.1016/j.cpc.2017.02.004> (In Eng.)
- Owen L.D., Ge W., Rieth M., Arienti M., Esclapez L., Soriano B., Mueller M.E., Day M., Sankaran R., & Chen J.H. (2024) PeleMP: The multiphysics solver for the combustion Pele adaptive mesh refinement code suite. *Journal of Fluids Engineering*, 146(4), 041103. <https://doi.org/10.1115/1.4064494> (In Eng.)
- Rao S., Chen B., & Xu X. (2024) Heterogeneous CPU–GPU parallelization for modeling supersonic reacting flows with detailed chemical kinetics. *Computer Physics Communications*, 300, 109188. <https://doi.org/10.1016/j.cpc.2024.109188> (In Eng.)
- Sarsembayev M.S., & Urmashiev B.A. (2019) Using CUDA technology to accelerate calculations in problems of chemical kinetics. *Cybernetics and Systems Analysis*. Springer New York. (In Eng.)
- Stone C.P., & Davis R.L. (2013) Techniques for solving stiff chemical kinetics on graphical processing units. *Journal of Propulsion and Power*, 29(4). — P. 764–773. <https://doi.org/10.2514/1.B34774> (In Eng.)
- Uranakara H.A., Barwey S., Hernández Pérez F.E., Raman V., & Im H.G. (2023) Accelerating turbulent reacting flow simulations on many-core/GPU architectures. *Computers & Fluids*. (In Eng.)
- Wang Y., Zeng Y., Deiterding R., & Liang J. (2026) An efficient GPU-accelerated adaptive mesh refinement framework for high-fidelity compressible reactive flows modeling. *Computer Physics Communications*, 318, 109870. <https://doi.org/10.1016/j.cpc.2025.109870> (In Eng.)

Zhang W., Almgren A., Beckner V., Bell J., Blaschke J., Chan C., Day M., Friesen B., Gott K., Graves D., Katz M., Myers A., Nguyen T., Nonaka A., Rosso M., Williams S., & Zingale M. (2019) AMReX: A framework for block-structured adaptive mesh refinement. *Journal of Open Source Software*, 4(37), 1370. <https://doi.org/10.21105/joss.01370> (In Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Number 358 (2026). 347–367

<https://doi.org/10.32014/2026.2518-1726.443>

IRSTI 28.23.27

UDC 004.4, 004.8, 004.6

© Symagulov A.^{1,2}, Smurygin V.¹, Belousov A.¹, Karypov A.¹,
Yunicheva N.R.^{1,3*}, 2026.

¹Institute of Information and Computational Technologies MSHE RK,
Almaty, Kazakhstan;

²Satbayev University (KazNITU), Almaty, Kazakhstan;

³Almaty University of Power Engineering and Telecommunications (AUPET),
Almaty, Kazakhstan.

E-mail: naduni@mail.ru

IMPROVING THE ACCURACY OF CROP AND WEED DETECTION USING UAVS IN SOYA FIELDS THROUGH IMAGE SEGMENTATION

Symagulov Adilkhan — Software Engineer, Almaty, Institute of Information and Computational Technologies MSHE RK; Satbayev University (KazNRTU), Almaty, Kazakhstan,
E-mail: asmogulove00@gmail.com, ORCID ID: <https://orcid.org/0000-0001-9974-3215>;

Smurygin Valentin — Software Engineer, Institute of Information and Computational Technologies MSHE RK, Almaty, Kazakhstan,

E-mail: valentinsmurygin@gmail.com, ORCID ID: <https://orcid.org/0009-0009-8699-159X>;

Belousov Alexey — Software Engineer, Institute of Information and Computational Technologies MSHE RK, Almaty, Kazakhstan,

E-mail: a.belousov87@gmail.com, ORCID ID: <https://orcid.org/0009-0000-3100-861X>;

Karypov Alpamys — Software Engineer, Institute of Information and Computational Technologies MSHE RK, Almaty, Kazakhstan,

E-mail: karypov1@gmail.com, ORCID ID: <https://orcid.org/0009-0006-6125-3791>;

Yunicheva Nadiya — PhD, leading researcher of the Institute of High Technologies of the National Academy of Sciences of the Republic of Kazakhstan; Almaty University of Power Engineering and Telecommunications (AUPET), Almaty, Kazakhstan,

E-mail: naduni@mail.ru; ORCID ID: <https://orcid.org/0000-0001-6351-34507>.

Abstract. *Relevance.* Accurate detection of crops and weeds in images acquired from unmanned aerial vehicles (UAVs) is an important prerequisite for the implementation of precision agriculture technologies and variable-rate herbicide application systems. However, the performance of detection algorithms, particularly YOLO, strongly depends on the quality and structure of the training dataset. *Purpose.* The aim of this study is to improve the quality of soybean and weed detection and localization in UAV images of soybean fields by applying image

pre-segmentation methods during the preparation of the training dataset. Methods. More than 370 RGB images of soybean fields were collected using a DJI Mavic Mini 2 UAV at altitudes ranging from 4 to 10 m. The images were annotated in YOLO format using the CVAT platform. A dataset containing soybean plants and seven weed species was developed. Two image pre-segmentation approaches were investigated: threshold segmentation using the PlantCV library and segmentation based on Bounding Boxes (Bbox). Detection was performed using the YOLOv8x model pretrained on the COCO dataset. The results were evaluated using Precision, Recall, F1-score, mAP50, and mAP50-95 metrics. Results. Computational experiments demonstrated that the Bbox segmentation approach provided the best detection performance. Compared with the original dataset, the F1-score increased from 0.64 to 0.79, while mAP50(B) increased from 0.72 to 0.90, and mAP50-95 reached 0.83. For the crop plant (soybean), a high classification performance was achieved with an F1-score of 0.97, exceeding previously reported results for soybean detection using UAV imagery. Practical significance. The proposed methodology for training dataset preparation and image pre-segmentation can be applied in the development of precision agriculture systems aimed at automatic weed detection and variable-rate herbicide application. The obtained results contribute to reducing chemical consumption and minimizing the environmental impact on agricultural fields.

Keywords: Detection, convolutional neural networks, YOLO, segmentation, precision farming, unmanned aerial vehicles, soya

For citations: Symagulov A., Smurygin V., Belousov A., Karyov A., Yunicheva N.R. Improving the accuracy of crop and weed detection using UAVs in soya fields through image segmentation. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 347-367. DOI <https://doi.org/10.32014/2026.2518-1726.443>

© Сумагулов А.^{1,2}, Смурыгин В.¹, Белоусов А.¹, Карыпов А.¹,
Юничева Н.Р.^{1,3*}, 2026.

¹Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан;

²Сатбаев университеті, Алматы, Қазақстан;

³Ғ. Дәукеев атындағы Алматы энергетика және байланыс университеті,
Алматы, Қазақстан.

E-mail: naduni@mail.ru

СОЯ АЛҚАПТАРЫНДА ҰҰА КӨМЕГІМЕН МӘДЕНИ ЖӘНЕ АРАМШӨП ӨСІМДІКТЕРІН ДЕТЕКТТЕУ САПАСЫН КЕСКІНДЕРДІ СЕГМЕНТТЕУ АРҚЫЛЫ АРТТЫРУ

Сумагулов Адилхан — бағдарламалық жасақтама инженері, Ақпараттық және есептеуіш технологиялар институты; Сатбаев университеті, Алматы, Қазақстан,
E-mail: asmogulove00@gmail.com, ORCID ID: <https://orcid.org/0000-0001-9974-3215>;

Смурыгин Валентин — бағдарламалық жасақтама инженері, Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан,

E-mail: valentinsmurygin@gmail.com, ORCID ID: <https://orcid.org/0009-0009-8699-159X>;

Белоусов Алексей — бағдарламалық жасақтама инженері, Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан,

E-mail: a.belousov87@gmail.com, ORCID ID: <https://orcid.org/0009-0000-3100-861X>;

Карыпов Алпамыс — бағдарламалық жасақтама инженері, Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан,

E-mail: karypov1@gmail.com, ORCID ID: <https://orcid.org/0009-0006-6125-3791>;

Юничева Надия — т. ғ. к., доцент, ҚР БҒМ ҒК АЕТИ жетекші ғылыми қызметкері; Ғ. Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан,

E-mail: naduni@mail.ru, ORCID идентификаторы: <https://orcid.org/0000-0001-6351-3450>.

Аннотация. Өзектілігі. Өсімдіктерді детекттеу алгоритмдерінің, әсіресе YOLO алгоритмінің тиімділігі оқыту деректер жиынтығының сапасы мен құрылымына едәуір тәуелді.

Мақсаты. Зерттеудің мақсаты — оқыту деректер жиынтығын дайындау барысында кескіндерді алдын ала сегменттеу әдістерін қолдану арқылы ұшқышсыз ұшу аппараттарының (ҰҰА) көмегімен алынған соя алқаптарының суреттерінде соя мен арамшөптерді детекттеу және локализациялау сапасын арттыру. Әдістері. Зерттеу барысында DJI Mavic Mini 2 ҰҰА көмегімен 4–10 м биіктікте соя алқаптарының 370-тен астам RGB кескіні жиналды. Кескіндер CVAT сервисі арқылы YOLO форматында таңбаланды. Құрамына соя мен арамшөптердің 7 түрі кіретін деректер жиынтығы әзірленді. Кескіндерді алдын ала сегменттеудің екі тәсілі қарастырылды: PlantCV кітапханасын пайдалана отырып шектік сегменттеу және шектеуші рамкалар (Bounding Box, Bbox) бойынша сегменттеу. Детекттеу үшін COCO деректер жиынтығында алдын ала оқытылған YOLOv8x моделі қолданылды. Сапа Precision, Recall, F1-score, mAP50 және mAP50-95 метрикалары бойынша бағаланды. Нәтижелері. Жүргізілген есептеу эксперименттері Bbox-сегменттеуді қолдану детекттеудің ең жақсы нәтижелерін қамтамасыз ететінін көрсетті. Бастапқы деректер жиынтығымен салыстырғанда F1-score мәні 0,64-тен 0,79-ға дейін өсті, mAP50(B) көрсеткіші 0,72-ден 0,90-ға дейін артты, ал mAP50-95 мәні 0,83-ке жетті. Мәдени өсімдік — соя үшін F1-score = 0,97 болатын жоғары классификация сапасына қол жеткізілді, бұл ҰҰА-дан алынған кескіндер бойынша сояны детекттеу үшін бұған дейін ұсынылған нәтижелерден жоғары.

Практикалық маңыздылығы. Ұсынылған оқыту деректер жиынтығын дайындау және кескіндерді алдын ала сегменттеу әдістемесі арамшөптерді автоматты түрде анықтауға және гербицидтерді саралап енгізуге бағытталған дәл егіншілік жүйелерін әзірлеуде қолданылуы мүмкін. Алынған нәтижелер химиялық препараттар шығынын азайтуға және ауыл шаруашылығы алқаптарына түсетін экологиялық жүктемені төмендетуге ықпал етеді.

Түйін сөздер: детекттеу, конволюциялық нейрондық желілер, YOLO, сегменттеу, дәл егіншілік, ұшқышсыз ұшу аппараттары, соя

© Сымагулов А.^{1,2}, Смурыгин В.¹, Белоусов А.¹, Карыпов А.¹,
Юничева Н.Р.^{1,3*}, 2026.

¹Институт информационных и вычислительных технологий МНВО РК,
Алматы, Казахстан;

² КазННТУ им. К.И. Сатпаева, Алматы, Казахстан;

³Алматинский Университет энергетики и связи им. Г. Даукеева,
Алматы, Казахстан.

E-mail: naduni@mail.ru

УЛУЧШЕНИЕ КАЧЕСТВА ДЕТЕКТИРОВАНИЯ КУЛЬТУРНЫХ И СОРНЫХ РАСТЕНИЙ С ПОМОЩЬЮ БПЛА НА ПОЛЯХ СОИ С ПРИМЕНЕНИЕМ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Сымагулов Адилхан — инженер-программист, Институт информационных и вычислительных технологий МНВО РК; Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан,

E-Mail: asmogulove00@gmail.com, <https://orcid.org/0000-0001-9974-3215>;

Смурыгин Валентин — инженер-программист, Институт информационных и вычислительных технологий МНВО РК, Алматы, Казахстан,

E-Mail: valentinsmurygin@gmail.com, <https://orcid.org/0009-0009-8699-159X>;

Белоусов Алексей — инженер-программист, Институт информационных и вычислительных технологий МНВО РК, Алматы, Казахстан,

e-mail: a.belousov87@gmail.com, <https://orcid.org/0009-0000-3100-861X>;

Карыпов Алпамыс — инженер-программист, Институт информационных и вычислительных технологий МНВО РК Алматы, Казахстан,

E-mail: karypov1@gmail.com, <https://orcid.org/0009-0006-6125-3791>;

Юничева Надия — кандидат технических наук, доцент, ведущий научный сотрудник, Институт информационных и вычислительных технологий МНВО РК; Алматинский университет энергетики и связи имени Г. Даукеева, Алматы, Казахстан,

E-mail: naduni@mail.ru, <https://orcid.org/0000-0001-6351-3450>.

Аннотация. *Актуальность.* Эффективность алгоритмов детектирования растений, в частности моделей семейства YOLO, существенно зависит от качества, структуры и полноты обучающего набора данных. Для задач точного земледелия особенно важна корректная локализация культурных и сорных растений на изображениях, полученных с беспилотных летательных аппаратов, поскольку такие данные могут использоваться для автоматического мониторинга посевов и дифференцированного внесения агрохимических средств. *Цель.* Повысить качество детектирования и локализации сои и сорных растений на снимках соевых полей, полученных с помощью беспилотных летательных аппаратов, за счет применения методов предварительной сегментации изображений при подготовке обучающего набора данных. *Методы.* Для проведения исследования выполнен сбор более 370 RGB-изображений соевых полей с помощью БПЛА DJI Mavic Mini 2 на высоте 4–10 м. Изображения размечены в формате YOLO с использованием сервиса CVAT. Разработан набор данных,

включающий сою и семь видов сорных растений. Исследованы два подхода предварительной сегментации изображений: пороговая сегментация с использованием библиотеки PlantCV и сегментация по ограничивающим рамкам (Bounding Box, Bbox). Для детектирования использовалась модель YOLOv8x, предобученная на наборе данных COCO. Оценка качества проводилась по метрикам Precision, Recall, F1-score, mAP50 и mAP50-95. *Результаты и выводы.* Проведенные вычислительные эксперименты показали, что применение Bbox-сегментации обеспечивает наилучшие результаты детектирования. По сравнению с исходным набором данных значение F1-score увеличилось с 0,64 до 0,79, показатель mAP50(B) -с 0,72 до 0,90, а mAP50-95 достиг 0,83. Для культурного растения, сои, достигнут высокий показатель качества классификации: F1-score = 0,97, что превышает результаты, ранее представленные для детектирования сои по изображениям с БПЛА. Практическая значимость исследования заключается в возможности применения предложенной методики подготовки обучающего набора данных и предварительной сегментации изображений при разработке систем точного земледелия, ориентированных на автоматическое обнаружение сорняков и дифференцированное внесение гербицидов. Полученные результаты могут способствовать снижению расхода химических препаратов и уменьшению экологической нагрузки на сельскохозяйственные угодья.

Ключевые слова: детектирование, сверточные нейронные сети, YOLO, YOLOv8x, сегментация, точное земледелие, беспилотные летательные аппараты, соя, сорные растения, PlantCV.

Финансирование. Работа выполнена при финансовой поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан (грант № BR24992908 «Система поддержки агротехнических мероприятий в растениеводстве на базе комплекса средств мониторинга и методов искусственного интеллекта (Agroscope)» и № BR21881908 «Комплекс экологического сопровождения городской агломерации»).

Введение. Рост численности населения планеты требует ускоренного развития продовольственной базы, чтобы к 2050 году объем производства продуктов питания был увеличен вдвое. Достижение этой цели возможно за счет перехода от экстенсивных форм сельского хозяйства к интенсивным методам. В стратегических документах правительства подчеркивается необходимость модернизации аграрного сектора и активного развития растениеводства. Реализация этих задач связана с адаптацией технологий точного земледелия к климатическим, погодным и агротехническим особенностям страны и их дальнейшим внедрением в практику фермерских хозяйств. Климат Казахстана, отличающийся резкой континентальностью, характеризуется высокой изменчивостью погодных условий, что напрямую влияет на развитие сельскохозяйственных культур. При этом природная

неоднородность почв внутри одного поля в разных погодных сезонах может проявляться противоположным образом. Так, локальные понижения рельефа в засушливые годы создают благоприятные условия для роста растений, тогда как во влажные сезоны они становятся наименее продуктивными из-за переувлажнения. Таким образом, качество земель внутри поля тесно и неоднозначно связано с погодными факторами, что требует более глубокого анализа и постоянного мониторинга. Это необходимо для принятия оптимальных агротехнических решений и повышения урожайности.

Краткий литературный обзор. Внедрение технологий точного земледелия один из важнейших путей для преодоления продовольственного кризиса связанного с быстрым ростом населения (Dwivedi A., 2017). Процессы точного земледелия предполагают точное воздействие на сельскохозяйственные культуры, возделываемые поля с целью повышения урожайности и снижения расходов. Важной целью является снижения влияния мешающих факторов ограничивающих рост полезных растений. Успешному произрастанию растений препятствуют ряд факторов: вредители, недостаток или избыток удобрений, состояние почвы, недостаток или переизбыток влаги, засоление почвы, сорные растения и т.п. (Kashyap et al., 2021). Достоверная идентификация этих факторов, разумеется, возможна при визуальном осмотре поля агрономом, но большие пространства полей препятствуют широкому применению ручного труда в процессе мониторинга. Очевидным путем улучшения ситуации является применение автоматических средств. Таким средством в последнее время выступают беспилотные летательные аппараты, которые широко используются в самых различных областях воздушного мониторинга, в том числе, в области мониторинга процессов точного земледелия. Подобный мониторинг технически достаточно очевиден. Для его проведения следует выбрать соответствующее поле, высоту полета и учесть ограничения, связанные с погодой, зарядом батарей и т.п. Однако, обработка получаемых данных является достаточно сложной и многоцелевой задачей. Такие данные, несут признаки, так называемых, больших данных (Mukhamediev et al., 2021) и практически не поддаются ручной обработке в оперативном режиме. Хотя при использовании снимков, полученных с борта низколетящих БПЛА, оснащенных камерами высокого разрешения, например, такими как перечисленные в (Okseneko et al., 2024), возникает возможность автоматического распознавания и классификации объектов на полях, однако конкретные пути реализации этого процесса не столь очевидны. Одной из подобных задач является задача классификации и локализации изображений культурных растений и сорняков на полученных снимках (задача детектирования). Решение этой задачи позволяет более точно прогнозировать будущий урожай и применять технологии переменной нормы (VRT - Variable rate technology) (Masi et al., 2023), которые не только снижают издержки, но и повышают урожайность. Для решения этой задачи широко используются методы машинного обучения и прежде

всего фреймворки включающие в свой состав сверточные нейронные сети: AlexNet, ResNet, VGG, Google, U-Net, MobileNets, and DenseNet (Peteinatos et al., 2020). Подобные исследования демонстрируют хорошие результаты, когда качество классификации достигает 90 и более процентов. Однако, на практике требуется решать задачи детектирования, то есть одновременно и локализовать растение (определять местоположение) и классифицировать его.

Наиболее совершенным фреймворком такого типа является YOLO (You Only Look Once) (Umar et al., 2024), который способен решить задачу детектирования объекта за один проход сигнала по сети. То есть обученная и оптимизированная сеть способна обрабатывать несколько десятков кадров высокого разрешения в секунду. Для успешной работы такой модели необходимо лишь обучить ее на соответствующем наборе аннотированных изображений. Очевидно, что подобные крайне привлекательные свойства модели стали катализатором роста числа ее приложений для детектирования растений на полях. Одним из таких приложений стала задача обработки снимков, полученных на полях сои (Symagulov et al., 2022). Соя - это растение, содержащее большое количество белка, обладающее множеством полезных свойств и занимающее лидирующее место среди бобовых культур, значительно превосходя другие (фасоль, горох, нут и т.п.) по объему производства. По этим причинам уделяется большое внимание оптимизации ее производства на основе технологий точного земледелия.

В ряде работ рассматривается применения алгоритма YOLO для детектирования культурных растений и сорняков на полях сои. В основном с применением наземных автоматов, например, решается задача сегментации сорняков на соевом поле со средней точностью 98%. В работе (Sunil et al., 2024) представлены результаты применения YOLO к изображениям, собранным наземным роботом для нескольких культур включая сою. Значения точности варьируются от 80,8 % до 98 % в зависимости от культуры, поля и окружающих условий. Однако при применении снимков полученных с борта БПЛА, достигнутая на сегодня точность классификации значительно ниже.

В работе (Tetila et al., 2024) рассматривается задача бинарной идентификации (соя или сорняки) с применением модели YOLOv5 и БПЛА. Достигнутая точность составляет 0,93 с применением значительно облегченной версии программы. Примерно такой же результат классификации сои получен в работе (Li et al., 2024).

Дополнительный анализ недавних исследований показывает, что развитие методов детектирования сорняков и культурных растений на изображениях с БПЛА идет сразу по нескольким направлениям: совершенствование архитектур глубокого обучения, переход от простого детектирования к сегментации и создание специализированных наборов данных для сои.

В работе (Silva et al., 2024) показано, что объединение задач детектирования и сегментации на материалах аэрофотосъемки позволяет более точно

выделять растения на сложном фоне, характерном для реальных полевых условий. Авторы подчеркивают, что при анализе сельскохозяйственных сцен существенную роль играют неоднородность освещения, тени, остатки почвы и растительных остатков, из-за которых качество распознавания заметно зависит от предварительной обработки изображений и способа представления объекта в обучающем наборе.

Существенный вклад в развитие этой области вносит и работа (Steinmetz et al., 2024), где представлен набор данных GrowingSoy, ориентированный на детектирование сорняков в посевах сои с использованием instance segmentation на протяжении разных стадий роста — от всходов до уборки. Это особенно важно, поскольку большинство существующих подходов оцениваются только на ранних стадиях вегетации, когда растения еще сравнительно легко различимы. Наличие такого набора данных демонстрирует, что дальнейший прогресс в задачах мониторинга сои связан не только с выбором модели, но и с расширением представительности обучающих выборок по фазам роста, плотности стояния растений и сложности сцены.

Наконец, обзор (Zhang et al., 2024) обобщает достижения в области применения БПЛА и глубокого обучения для управления сорной растительностью и показывает, что наиболее перспективными являются подходы, сочетающие высокодетальные съемки, интеллектуальную предобработку и модели, устойчивые к изменению масштаба объектов. В совокупности эти результаты подтверждают, что предварительная сегментация и специализированная подготовка данных являются не вспомогательными, а ключевыми компонентами систем точного земледелия, ориентированных на надежное распознавание культурных и сорных растений.

В настоящей работе мы рассматриваем возможность использования набора данных, в котором размечена только часть изображений растений. Благодаря предварительной сегментации мы, несмотря на указанное ограничение, получили результаты классификации сои значительно лучше, чем в упомянутых статьях. Кроме введения, работа состоит из 3 основных разделов.

Во втором разделе описывается метод, примененный для сбора данных, разметки и детектирования изображений. В третьем разделе мы описываем полученные результаты. В заключении подводятся итоги работы и формулируются задачи будущих исследований.

Методы и материалы исследования. Процесс проведения исследований состоял из сбора фотографий на первой фазе развития растений с помощью БПЛА. Всего было получено более 300 изображений поля в высоком разрешении (4K). Затем вручную была выполнена разметка изображений в формате YOLO (Ultralytics, 2024). Изображение размечалось с помощью общедоступного сервиса CVAT (CVAT, 2024). Затем полученные снимки подавались на вход предварительно обученной модели YOLO, показанной на рисунке 1:



Рисунок 1 - Методическая схема исследования

Опишем перечисленные на рисунке стадии исследования подробнее.

1) Фотографии поля были получены в процессе плановых экспериментальных облетов полей сои. Облет проводился с помощью серийного дрона DJI Mavic Mini 2, оснащенного RGB камерой с разрешением 4K пикселей на высоте от 4 до 10 метров и изображен на рисунке 2. Технические характеристики дрона представлены в приложении А.



Рисунок 2 - Фотография соевого поля

2) Разметка изображений выполнялась с помощью CVAT, так что большинство изображений растений отмечались боксом и аннотировались одним из 10 классов, показано на рисунке 3:

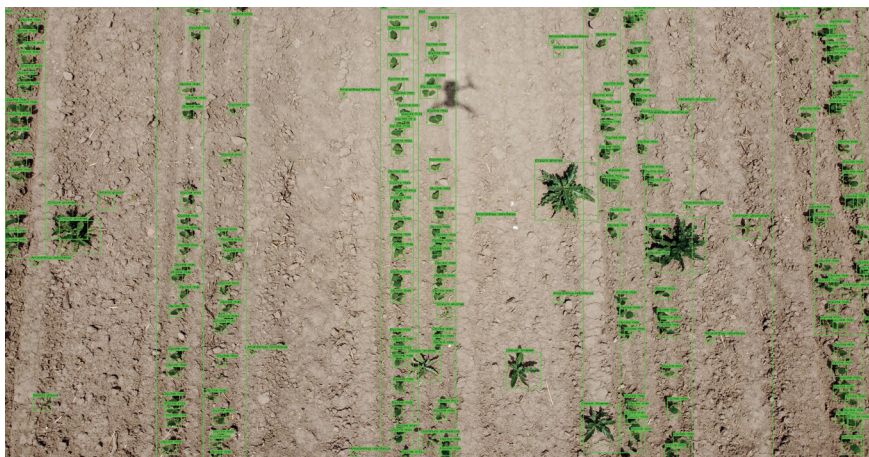


Рисунок 3 - Аннотированное изображение поля

Набор полученных размеченных изображений был разделен в пропорции 80/20 на тренировочный (300 снимков) и тестовый (74 снимка)

3) Сегментирование. Суть процесса сегментирования состоит в том, чтобы удалить задний фон вместе с неразмеченными объектами на каждом снимке. Для этого использовано два подхода.

Первый подход, реализованный в серии экспериментов 1 обеспечивал пороговую сегментацию изображений растений с помощью библиотеки plantcv (Zenn, 2024). В процессе пороговой сегментации выполняется сравнение каждого пикселя изображения с заданным порогом в диапазоне от 0 до 255. Яркость пикселей со значением меньше порога устанавливается в 0. Пиксели, имеющие яркость выше порога, не меняются и показаны на рисунке 4.

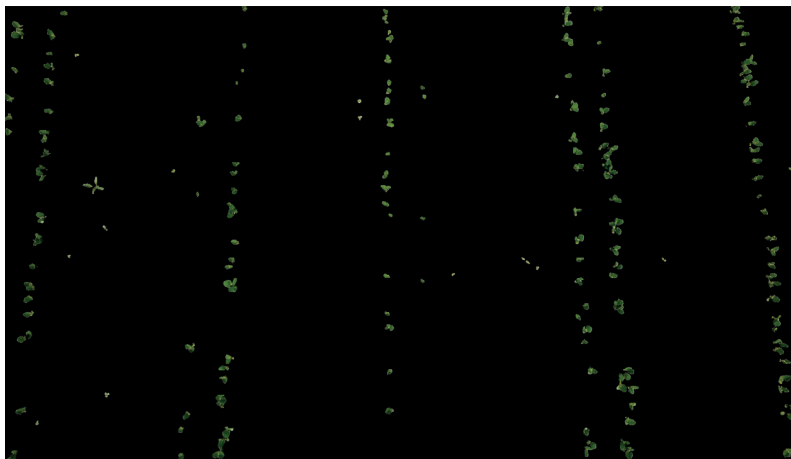


Рисунок 4 - Эксперимент 1. Пороговая сегментация с помощью библиотеки plantcv при уровне порога 120

Второй вариант сегментации (эксперимент 2 или Bbox сегментация) заключался в удалении всех неразмеченных изображений растений и фона за пределами ограничивающей рамки (Bounding box - Bbox) и показан на рисунке 5.

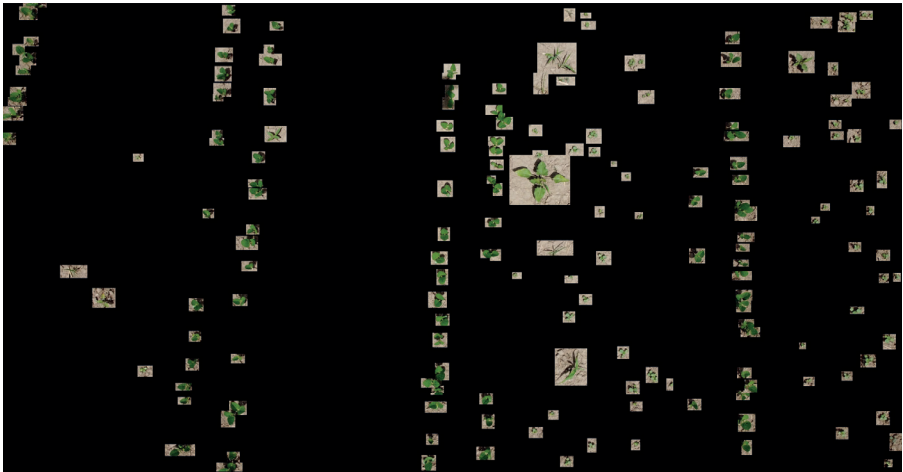


Рисунок 5 - Эксперимент 2. Bbox сегментация. Весь фон за пределами Bbox удален

4) Дообучение YOLO и оценка результатов. В процессе экспериментов использовался фреймворк YOLOv8x, имеющей 70-80 миллионов обучаемых параметров (Ultralytics, 2024). Модель YOLO была предобучена на наборе данных COCO (Common Objects in Context), содержащем объекты 80-ти классов от персон и автомобилей до зубных щеток. Для каждого набора данных использовался один и тот же набор изображений для тестирования и оценки полученных результатов, но дообучение велось независимо на трех наборах данных: набор данных для эксперимента 1(пороговая сегментация); 2 - Набор данных для эксперимента 2 (Bbox сегментация).; 0 - Оригинальный набор данных без сегментации. Поскольку примененный фреймворк выполняет не только классификацию растений, но и их локализацию, то для оценки полученных результатов детектирования использовались метрики качества классификации и качества локализации, представленные в таблице 1.

Таблица 1 - Метрики оценки качества детектирования растений

Метрика	Формула	Пояснение
Показатели классификации		
Precision (точность)	$P = \frac{TP}{(TP + FP)}$ где «истинные положительные результаты» (TP) и «истинные отрицательные результаты» (TN) - это случаи правильной работы классификатора. Соответственно, «ложные	Доля истинно положительных предсказаний среди всех предсказанных положительных случаев.

	отрицательные результаты» (FN) и «ложные положительные результаты» (FP) - это случаи ошибочной классификации.	
Recall (полнота)	$R = \frac{TP}{(TP + FN)}$	Доля истинно положительных предсказаний среди всех фактических положительных случаев.
F1score (гармоническая мера)	$F1score = \frac{2 * P * R}{(P + R)}$	Гармоническое среднее между точностью и полнотой.
Показатели локализации		
IoU	$IoU = \frac{\text{Площадь пересечения}}{\text{Площадь объединения}}$	Мера, которая количественно определяет перекрытие между предсказанным ограничительным полем (обычно прямоугольником - Bbox) и истинным ограничительным полем. Она играет важную роль в оценке точности локализации объектов.
	$AP = \int_0^1 P(R) dR$ где P(R) - функция точности от полноты.	Средняя точность для одного класса, вычисляемая как площадь под кривой Precision-Recall.
mAP50	$mAP_{0.5} = \frac{1}{N} \sum_{i=1}^N AP_i^{0.5}$ где: N - общее количество классов; $AP_i^{0.5}$ - среднее значение Precision (AP) для класса i при IoU = 0.5.	Среднее значение точности по всем классам, используемое для оценки общей производительности модели.
mAP50-95	$mAP_{0.50:0.95} = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{10} \sum_{j=0}^9 AP_i^{(0.5 + 0.05j)} \right)$ где: N - общее количество классов; $AP_i^{(0.5 + 0.05j)}$ — среднее значение Precision (AP) для класса i при IoU = $0.5 + 0.05j$ j - индекс, принимающий значения от 0 до 9, соответствующие порогам IoU: 0.5, 0.55, 0.6, ..., 0.95.	Средняя точность, усредненная по нескольким порогам IoU от 0.5 до 0.95 с шагом 0.05.

Полученные результаты. Предварительные этапы, связанные с подготовкой данных, привели к созданию несбалансированных по классам датасетов 0, 1 и 2, содержащих изображения несколько десятков тысяч растений в Таблице 2. Наиболее полно представлены изображения полезных растений (соя), в то время как изображения отдельных сорняков содержат

от десятков до сотен картинок. Оригинальный набор данных (0) содержит небольшое количество (около 10%) неразмеченных изображений растений.

Таблица 2 - Характеристики исходного набора данных и наборов данных, подготовленных в ходе эксперимента 1(пороговая сегментация) и 2 (Вbox сегментация).

Название класса	Расшифровка названия класса	1	2	0
0: Bed	Грядка	0	0	2458
1: Glycine max	Соя	41368	42282	42282
2: Amaranthus retroflexus	Ширица обыкновенная	3840	6155	6155
3: Convolvulus arvensis	Вьюнок полевой	304	385	385
4: Setaria glauca	Щетинник сизый	149	1897	1897
5: Xanthium strumarium	Дурнишник обыкновенный	439	513	513
6: Cirsium arvense	Осот розовый	516	532	532
7: Echinochloa crusgalli	Просо куриное	0	0	0
8: Hibiscus trionum	Гибискус тройчатый	15	15	15
9: Abutilon theophrasti	Канатник Теофраста	27	27	27
10: Chenopodium album	Марь белая	0	0	0
11: Apera spica-venti	Метлица обыкновенная, полевая	0	0	0
Примечание 1 – Набор данных для эксперимента 1(пороговая сегментация); 2- Набор данных для эксперимента 2 (Вbox сегментация).; 0- Оригинальный набор данных без сегментации.				

В последующих экспериментах классы, не представленные в тестовом наборе (0,7,10,11) не использовались при валидации модели.

Поскольку наборы данных не сбалансированы, то есть количество растений разного вида существенно отличается, то для сравнения результатов, полученных в ходе вычислительных экспериментов, рассмотрим матрицы путаницы (confusion matrix).

На рисунке 6 представлена матрица путаницы, полученная в ходе экспериментов с оригинальным (не сегментированных) набором снимков.

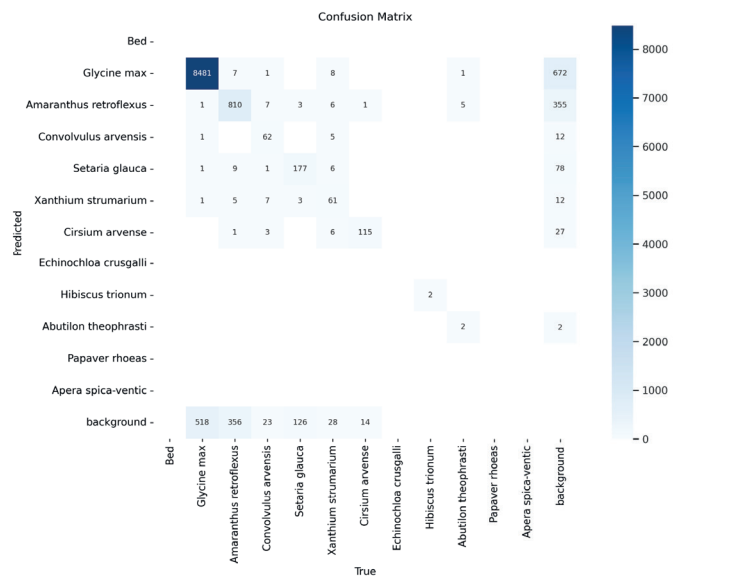


Рисунок 6 - Матрица путаницы для оригинального набора данных

Цифры по диагонали означают процент правильно классифицированных растений соответствующего вида. Цифры вне главной диагонали означают процент ошибок первого и второго рода (FP – выше диагонали, FN – ниже главной диагонали). Нетрудно подсчитать `f1_score` для сои. Оно составляет 0.933. Это качество примерно соответствует заявленному в литературе (`f1_score` = 0,9307 (Li et al., 2024)).

Данный результат достигнут за счет аккуратной разметки наборов изображений. Однако его можно улучшить. На рисунке 7 показана матрица путаницы, полученная в ходе эксперимента 1 (пороговая сегментация).

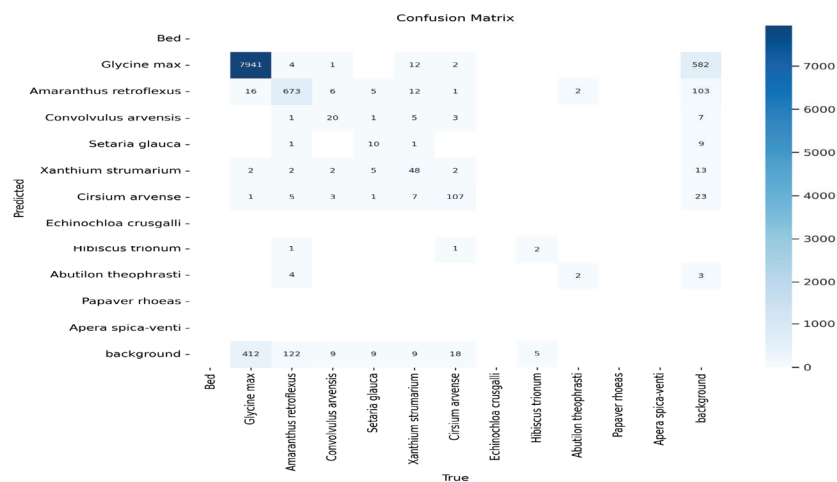


Рисунок 7 - Матрица путаницы для эксперимента 1 (пороговая сегментация)

Качество классификации культурного растения возросло ($f1_score=0.939$), что несколько лучше результатов, указанных в литературе. Однако Bbox сегментация позволяет улучшить и этот результат. На рисунке 8 показана матрица путаницы, полученная в ходе эксперимента 2 (Bbox сегментация).

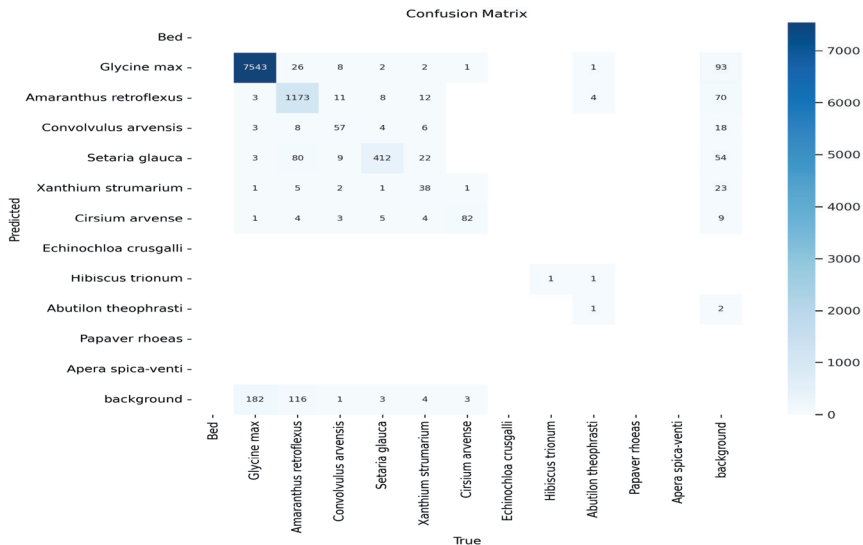


Рисунок 8 - Матрица путаницы для эксперимента 2 (пороговая сегментация)

Качество классификации культурного растения стало значительно выше ($f1_score=0.979$). Однако качество классификации сорняков, таких как, *Abutilon theophrasti* и *Xanthium strumarium* снизилось. Частично полученный результат можно объяснить тем, что, во-первых, количество сорных растений всех видов мало, во-вторых, *Abutilon theophrasti* на этой фазе роста весьма мелкий сорняк. Однако требуются дополнительные исследования для выяснения причин снижения качества распознавания отдельных видов сорных растений. Подводя итог, приведем усредненные результаты качества детектирования всех видов растений, как культурного растения (сои) так и сорняков в Таблице 3.

Таблица 3 - Результаты вычислительных экспериментов при детектировании сои и сорняков.

Тренировочный набор данных	mAP50	mAP50-95	recall	precision	F1_score
Оригинальный датасет	0.72	0.3	0.599	0.677	0.6356
Эксперимент 1	0.65	0.348	0.639	0.749	0.6896
Эксперимент 2	0.9	0.836	0.799	0.789	0.7939

Можно констатировать, что качество детектирования при применении Bbox сегментации (эксперимент 2) значительно выше как в сравнении с оригинальным набором данных так и в сравнении с пороговой сегментацией

(эксперимент 1). Значительно возросло качество локализации ($mAP_{50-95}=0.836$), что позволяет более уверенно идентифицировать место расположения растения на поле.

Обсуждение. Полученные результаты показывают, что ключевым фактором повышения качества детектирования в рассматриваемой задаче является не только выбор современной архитектуры YOLO, но и способ подготовки обучающего набора данных. В нашем исследовании именно предварительная Vbox-сегментация обеспечила заметный прирост качества по сравнению как с исходным набором изображений, так и с пороговой сегментацией. Это позволяет сделать вывод, что для снимков, полученных с борта БПЛА, существенную роль играет уменьшение влияния фона, неразмеченных растений и мелких неоднородностей сцены, которые затрудняют обучение детектора и приводят к росту ошибок локализации. Особенно важным представляется тот факт, что улучшение коснулось не только интегрального показателя F1-score, но и метрик локализации, включая mAP_{50-95} , что указывает на более устойчивое согласование между классификацией объекта и точностью определения его положения на изображении.

Сопоставление с опубликованными работами подтверждает правдоподобность полученных выводов.

В исследовании (Silva et al., 2024) показано, что при анализе аэрофотоснимков сельскохозяйственных культур высокое качество детектирования и сегментации напрямую зависит от способности модели выделять растения на сложном фоне, содержащем тени, почвенные включения и остатки растительности. Авторы сравнивали несколько архитектур, включая YOLOv8, Mask R-CNN и U-Net, и пришли к выводу, что сама по себе модель не устраняет проблему фоновой неоднородности без соответствующей подготовки данных. Сходная логика прослеживается и в работе (Steinmetz et al., 2024), где для набора данных GrowingSoy, охватывающего различные стадии вегетации сои, были получены значения segmentation average precision порядка 79,1% и average recall до 73,3%, а для сои по метрике mAP_{50} достигнут уровень 90,1%. Эти результаты демонстрируют, что даже при наличии более богатого и хронологически разнообразного набора данных задача остается чувствительной к фазе роста растений и сложности сцены. Обзор (Zhang et al., 2024) также подчеркивает, что перспективные системы точного земледелия должны сочетать высокодетальные UAV-изображения, интеллектуальную предобработку и модели, устойчивые к изменению масштаба объектов и межвидовому сходству растений.

Вместе с тем результаты нашего исследования указывают и на ограничения предложенного подхода. Рост качества детектирования сои сопровождался менее устойчивым распознаванием отдельных видов сорняков, что, вероятно, связано с дисбалансом классов, малым числом объектов некоторых видов и их малыми размерами на ранних стадиях

развития. Следовательно, предварительная сегментация не устраняет полностью проблему недостаточной представленности классов, а лишь повышает информативность визуального представления объектов.

С практической точки зрения это означает, что предложенная методика особенно эффективна в задачах раннего мониторинга сои и подготовки данных для систем адресного внесения гербицидов, однако для повышения качества распознавания сорняков необходимы дальнейшее расширение датасета, учет последующих фаз вегетации и, возможно, переход к более детализированным схемам *instance segmentation*. Таким образом, основное значение полученных результатов заключается в том, что предварительная сегментация может рассматриваться как самостоятельный и воспроизводимый способ повышения качества UAV-детектирования в условиях ограниченной и частично размеченной выборки, а не только как вспомогательный этап обработки изображений.

Предложенный подход может быть использован как основа для создания прикладных систем мониторинга состояния посевов, ориентированных на раннее выявление засоренности и поддержку решений по дифференцированному внесению средств защиты растений. Важным преимуществом методики является ее воспроизводимость: она не требует специализированных сенсоров, а базируется на стандартной RGB-съемке с доступного БПЛА и использовании современных методов аннотирования и предобработки данных. Это создает предпосылки для дальнейшего переноса разработанного подхода на другие культуры и расширения его применения в практических задачах цифрового земледелия.

Заключение. Методы компьютерного зрения, особенно те, которые связаны с обнаружением, локализацией и классификацией объектов, помогают в процедурах борьбы с сорняками на конкретных участках поля. Они позволяют внедрять практики переменной нормы внесения гербицидов в пределах поля. Однако основной проблемой при разработке системы обнаружения сорняков и подсчета культурных растений является необходимость в правильно аннотированной базе данных для различения сорняков и культур в полевых условиях. В настоящей работе использован разработанный авторами набор аннотированных изображений культур на соевом поле, содержащий более 370 высоко детальных изображений, аннотированных для применения фреймворков YOLO.

Общее число аннотированных растений приближается к 50 тысячам. На первом этапе экспериментов достигнуты результаты в целом сопоставимые с описанными в литературе. Дальнейшие эксперименты с применением *Bbox* сегментации, позволили существенно улучшить эти результаты. Достигнут высокий результат детектирования сои на снимках полученных с борта БПЛА.

Разработанная методика подготовки данных и обучения позволила превзойти результаты, описанные ранее (Li et al., 2024) ($F1_score=0.93$)

и достичь результатов получаемых с помощью систем машинного зрения, установленных на наземных роботах ($F1_score=0.98$). Другими словами, предложенная и реализованная нами методика предобработки обучающего набора изображений позволила достичь лучших чем известные из литературы результаты детектирования сои с помощью БПЛА.

Однако, несмотря на хорошие результаты, достигнутые в процессе исследования, имеются некоторые ограничения, которые нужно будет учесть в процессе проведения следующих этапов исследования:

1. Наборы данных не сбалансированы и содержат много изображений культурных растений и намного меньше сорняков. В результате качество классификации сорняков значительно уступает качеству детектирования полезного растения.

2. Все эксперименты были проведены только на первом этапе роста сои, когда растения хорошо идентифицируются и почти не перекрываются. Расширение набора данных снимками полей на последующих этапах роста сои повысит практическую применимость набора данных и соответствующего метода детектирования.

3. Использовалась модель YOLO предобученная на изображениях данных СОСО, которые не содержат специфических изображений культурных растений и сорняков. Увеличение набора размеченных снимков и применение методов аугментации может повысить качество классификации и локализации не только полезных растений, но и всех классов сорняков.

Перечисленные задачи являются целями следующих этапов исследований, планируемых авторами.

Приложение А. Технические характеристики дрона разведчика

Технические характеристики дрона разведчика приведены в таблице А.1:

Таблица А.1 - Технические характеристики дрона разведчика (DJI, 2024a)

Параметр	Значение
Взлетный вес	<249 г
Размер по диагонали (без пропеллеров)	213 мм
Максимальная скорость набора высоты	5 м/с (режим S); 3 м/с (режим N); 2 м/с (режим C)
Максимальная скорость снижения	3,5 м/с (режим S); 3 м/с (режим P); 1,5 м/с (режим C)
Максимальная скорость полета	16 м/с (режим S); 10 м/с (режим N); 6 м/с (режим C)
Максимальный угол наклона	40° (режим S); 25° (режим N); 25° (режим C) (*до 40° при сильном ветре)
Максимальная угловая скорость	130°/с (режим S); 60°/с (режим N); 30°/с (режим C)* (*можно настроить до 250°/с в приложении DJI Fly)
Максимальная высота полета	2000 - 4000 м
Максимальная скорость ветра	8,5–10,5 м/с (до 5 баллов по шкале Бофорта)

Максимальное время полета	31 минута (при скорости 4,7 м/с в безветренную погоду)
Диапазон рабочих температур	От 0° до 40°C
Точность позиционирования в вертикальной плоскости	±0,1 м (визуальное позиционирование), ±0,5 м (спутниковое позиционирование)
Точность позиционирования в горизонтальной плоскости	±0,3 м (визуальное позиционирование), ±1,5 м (спутниковое позиционирование)
Рабочая частота	2,4–2,4835 ГГц
Параметр	Значение
Взлетный вес	<249 г
Размер по диагонали (без пропеллеров)	213 мм
Максимальная скорость набора высоты	5 м/с (режим S); 3 м/с (режим N); 2 м/с (режим C)
Максимальная скорость снижения	3,5 м/с (режим S); 3 м/с (режим P); 1,5 м/с (режим C)
Максимальная скорость полета	16 м/с (режим S); 10 м/с (режим N); 6 м/с (режим C)
Максимальный угол наклона	40° (режим S); 25° (режим N); 25° (режим C) (*до 40° при сильном ветре)
Максимальная угловая скорость	130°/с (режим S); 60°/с (режим N); 30°/с (режим C)* (*можно настроить до 250°/с в приложении DJI Fly)
Максимальная высота полета	2000 - 4000 м
Максимальная скорость ветра	8,5–10,5 м/с (до 5 баллов по шкале Бофорта)
Максимальное время полета	31 минута (при скорости 4,7 м/с в безветренную погоду)
Диапазон рабочих температур	От 0° до 40°C
Точность позиционирования в вертикальной плоскости	±0,1 м (визуальное позиционирование), ±0,5 м (спутниковое позиционирование)
Точность позиционирования в горизонтальной плоскости	±0,3 м (визуальное позиционирование), ±1,5 м (спутниковое позиционирование)
Рабочая частота	2,4–2,4835 ГГц

Технические характеристики камеры дрона разведчика приведены в таблице А.2.

Таблица А.2 - Технические характеристики камеры дрона разведчика DJI Mavic Mini 2 (DJI, 2024b)

Параметр	Значение
Эффективное количество пикселей	12 Мп
Тип сенсора	CMOS, размер 1/2,3 дюйма
Объектив	Угол обзора: 83°, эквивалент 24 мм (формат 35 мм); Диафрагма: f/2.8; Фокусировка: от 1 м до бесконечности
Выдержка	Электронный затвор, скорость: 4–1/8000 с
ISO	Видео: 100–3200 (авто), 100–3200 (ручной); Фото: 100–3200 (авто), 100–3200 (ручной)
Размер изображения	4:3: 4000x3000; 16:9: 4000x2250
Поддерживаемые файловые системы	FAT32 (до 32 Гбайт); exFAT (более 32 Гбайт)

Максимальный битрейт видео	100 Мбит/с
Поддерживаемые форматы файлов	Фото: JPEG/DNG (RAW); Видео: MP4 (H.264/MPEG-4 AVC)
Поддерживаемые карты памяти	UHS-I Speed Class 3 или выше
Разрешение видео	4K: 3840×2160 при 24/25/30 кадров в секунду; 2.7K: 2720×1530 при 24/25/30/48/50/60 кадров в секунду; FHD: 1920×1080 при 24/25/30/48/50/60 кадров в секунду.

References

- CVAT (2024). <https://www.cvat.ai> Retrieved October 24, 2025 (in Eng.).
- DJI (2024a). Consumer drones comparison. — <https://www.dji.com/products/comparison-consumer-drones?from=store-product-page-comparison> (Retrieved October 24, 2025). (in Eng.).
- DJI (2024b) Support for DJI Mini 2. — <https://www.dji.com/global/support/product/mini-2> Retrieved October 24, 2024). (in Eng.).
- Dwivedi A. (2017) Precision agriculture. *Parmar Publ. Distrib.*, 5. – P. 83–105. (in Eng.).
- Kashyap B., Kumar R. (2021) Sensing methodologies in agriculture for soil moisture and nutrient monitoring. *IEEE Access*, 9. – P.14095–14121. <https://doi.org/https://doi.org/10.1109/access.2021.3052478> (in Eng.).
- Li J., Zhang W., Zhou H., Yu C., Li Q. (2024) Weed detection in soybean fields using improved YOLOv7 and evaluating herbicide reduction efficacy. *Frontiers in Plant Science*, 14. – P.1284338. <https://doi.org/https://doi.org/10.3389/fpls.2023.1284338> (in Eng.).
- Masi M., Di Pasquale J., Vecchio Y., Capitanio F. (2023) Precision farming: Barriers of variable rate technology adoption in Italy. *Land*, 12. – P.1084. <https://doi.org/https://doi.org/10.3390/land12051084> (in Eng.).
- Mukhamediev R.I., Symagulov A., Kuchin Y.I., Zaitseva E.N., Bekbotayeva A., Yakunin K., Assanov I., Levashenko V.G., Popova Y., Akzhalova A.Z., Bastaubayeva S., Tabyimbaeva L. (2021) Review of some applications of unmanned aerial vehicles technology in the resource-rich country. *Applied Sciences*, 11(21). – P.10171. <https://doi.org/https://doi.org/10.3390/app112110171> (in Eng.).
- Okseneko A.A., Yerimbetova A.S., Kuanaev A., Mukhamediev R.I., Kuchin Y.I. (2024) Tekhnicheskie sredstva distantsionnogo monitoringa s pomoshch'yu bespilotnykh letatel'nykh platform [Technical remote monitoring systems using unmanned aerial vehicles]. *Izvestiya Natsional'noi Akademii Nauk Respubliki Kazakhstan*, 3(351). – P.152–173. <https://doi.org/https://doi.org/10.32014/2024.2518-1726.298> (in Russian).
- Osman Y., Dennis R., Elgazzar K. (2021) Yield estimation and visualization solution for precision agriculture. *Sensors*, 21(19). -P.6657. <https://doi.org/https://doi.org/10.3390/s21196657> (in Eng.).
- Peteinatos G., Reichel P., Karouta J., Andújar D., Gerhards R. (2020) Weed identification in maize, sunflower, and potatoes with the aid of convolutional neural networks. *Remote Sens.*, 12. – P.4185. <https://doi.org/https://doi.org/10.3390/rs12244185> (in Eng.).
- PlantCV (2024). *Plant Computer Vision*. — <https://plantcv.org/> (Retrieved October 24, 2024). (in Eng.).
- Symagulov A., Kuchin Y., Yakunin K., Murzakhmetov S., Yelis M., Oxenenko A., Assanov I., Bastaubayeva Sh., Tabyimbaeva L., Rabčan J. Mukhamediev R. (2022) Recognition of soybean crops and weeds with YOLO v4 and UAV. *International Conference on Internet and Modern Society*. Cham: Springer Nature Switzerland, 3–14. https://doi.org/https://doi.org/10.1007/978-3-031-50609-3_1 (in Eng.).
- Tetila E. C., Moro B.L., Astolfi G. (2024) Real-time detection of weeds by species in soybean using UAV images. *Crop Protection*, 184. – P. 106846. <https://doi.org/https://doi.org/10.1016/j.cropro.2024.106846> (in Eng.).
- Silva J.A.O.S., Siqueira V.S., Mesquita M., Vale L.S.R., Marques T.N.B., Silva J.L.B., Silva M.V.,

Lacerda L.N., Oliveira-Júnior J. F., Pedroso de Lima J.L., Oliveira H.F.E. (2024) Deep learning for weed detection and segmentation in agricultural crops using images captured by an unmanned aerial vehicle. *Remote Sensing*, 16. – P. 4394. <https://doi.org/https://doi.org/10.3390/rs16204394> (in Eng.).

Steinmetz R., Kich V. A., Krever H., Mazzarolo J. D. R., Grando R. B., Marini V., Trois C., Nieuwenhuizen A. (2024) From Seedling to Harvest: The GrowingSoy Dataset for weed detection in soy crops via instance segmentation. In *Proceedings of the IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE International Conference on Robotics, Automation and Mechatronics (RAM)*. – P. 502–507. <https://doi.org/https://doi.org/10.1109/cis-ram61939.2024.10673082> (in Eng.).

Zhang J., Yu F., Zhang Q., Wang M., Yu J., Tan Y. (2024) Advancements of UAV and deep learning technologies for weed management in farmland. *Agronomy*, 14. – P. 494. DOI:10.3390/agronomy14030494 (in Eng.).

Ultralytics (2024) YOLOv8 Label Format: A stepbystep guide. — <https://docs.ultralytics.com/ru/datasets/detect/coco/> (Retrieved October 24, 2025). (in Eng.).

Umar M., Altaf S., Ahmad S., Mahmoud H., Mohame A. S. N., Ayub R. (2024) Precision agriculture through deep learning: Tomato plant multiple diseases recognition with CNN and improved YOLOv7. *IEEE Access*, 12. – P. 9167–49183. <https://doi.org/https://doi.org/10.1109/access.2024.xxxxx> (in Eng.).

Zenn(2024) Explanation of all of YOLO series, 11. <https://doi.org/https://doi.org/14a87a0db17dfa> (Retrieved October 24, 2025). (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 368–390

<https://doi.org/10.32014/2026.2518-1726.444>

IRSTI 27.47.19

UDC 512.647

© **Tashenova Zh.^{*}, Gabdullin A.R.¹, Abdugulova Zh.¹, Amanzholova Sh.²,
Santeyeva S.¹, 2026.**

Department of Information Security, L.N. Gumilyov Eurasian National
University, Astana, Kazakhstan.
E-mail: zhuldyz_tm@mail.ru

SECURITY EVALUATION OF WPA3 WIRELESS NETWORKS UNDER DEAUTHENTICATION ATTACK SCENARIOS

Tashenova Zhuldyz — PhD, Department of Information Security, L. N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: zhuldyz_tm@mail.ru, <https://orcid.org/0000-0003-3051-1605>;

Gabdullin Abzal — Department of Information Security System, Faculty of Information Technologies, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: anchorite.exe@gmail.com, <https://orcid.org/0000-0003-3051-1605>;

Abdugulova Zhanat — Associated Professor, Department of Information Technologies, L. N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: janat_6767@mail.ru, <https://orcid.org/0000-0001-7462-4623>;

Amanzholova Shirin — PhD, Kurmangazy Kazakh National Conservatory, Almaty, Kazakhstan,

E-mail: schirin75@mail.ru; <https://orcid.org/0000-0002-6674-2766>;

Santeyeva Saya — PhD, Department of Information Security, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: saya_santeeva@mail.ru, <https://orcid.org/0000-0001-9426-6704>.

Abstract. Wireless networks are now ubiquitous as part of the global digital infrastructure, providing communications in residential, commercial and public contexts. The advent of the Wi-Fi Protected Access III (WPA3) protocol has represented a major advance in wireless network security through improved authentication and better cryptographic protections. Nevertheless, despite the advances provided by WPA3, there continue to be vulnerabilities in wireless network availability that attackers can leverage. Specifically, de-authentication attacks exploit vulnerabilities in IEEE 802.11 management frames that do not compromise the integrity of encryption, but can disrupt client access to the network. This study aims at assessing the resistance of WPA3 wireless networks against de-authentication attacks, as well as analyzing the effects of these attacks on the availability of the network. The study employed an experimental approach

based on a controlled wireless testbed where both WPA2 and WPA3 network configurations were created and tested with simulated de-authentication attacks. The results demonstrated that while WPA3 provides significant improvements in authentication and encryption, it is possible to negatively impact its availability through de-authentication attacks under specific circumstances. The study identifies potential avenues to improve the protection of wireless networks against availability attacks, including enhancing the protection of IEEE 802.11 management frames and developing more sophisticated intrusion detection systems. The purpose of this study is to evaluate the behavior of modern wireless networks under deauthentication attacks and to assess the impact of these attacks on wireless network availability in a network infrastructure that utilizes modern security standards. The experiments demonstrate that deauthentication attacks can seriously impact wireless network stability, disrupting communications between wireless devices and access points.

Keywords: Wireless network security, WPA3, Deauthentication attack, Wi-Fi security, Wireless intrusion detection

For citations: Tashenova Zh., Gabdullin A.R., Abdugulova Zh., Amanzholova Sh., Santeyeva S. Security evaluation of WPA3 wireless networks under deauthentication attack scenarios. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 368–390. DOI <https://doi.org/10.32014/2026.2518-1726.444>

© Ташенова Ж.М.^{*}, Габдуллин А.Р.¹, Абдугулова Ж.К.¹,
Аманжолова Ш.А.², Сантеева С.Ә.¹, 2026.

Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан.
E-mail: zhuldyz_tm@mail.ru

ДЕАТЕНТИФИКАЦИЯЛАУ ШАБУЫЛЫ СЦЕНАРИЙЛЕРІНДЕГІ WPA3 СЫМСЫЗ ЖЕЛІЛЕРІНІҢ ҚАУІПСІЗДІГІН БАҒАЛАУ

Ташенова Жұлдыз — PhD, Ақпараттық қауіпсіздік кафедрасы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: zhuldyz_tm@mail.ru, <https://orcid.org/0000-0003-3051-1605>;

Габдуллин Абзал — Ақпараттық қауіпсіздік кафедрасы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: anchorite.exe@gmail.com, <https://orcid.org/0000-0003-3051-1605>;

Абдугулова Жанат — экономика ғылымдарының кандидаты, қауымдастырылған профессор, Л.Н. Гумилев Атындағы Еуразия Ұлттық университеті, ақпараттық технологиялар факультеті, Астана, Қазақстан,
E-mail: janat_6767@mail.ru, <https://orcid.org/0000-0001-7462-4623>;

Аманжолова Ширин — PhD, Құрманғазы атындағы Қазақ ұлттық консерваториясы, Алматы, Қазақстан,
E-mail: schirin75@mail.ru, [https:// orcid.org/0000-0002-6674-2766](https://orcid.org/0000-0002-6674-2766);

Сантеева Сая — PhD, Ақпараттық қауіпсіздік кафедрасы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail:: saya_santeeva@mail.ru, <https://orcid.org/0000-0001-9426-6704>.

Аннотация. Сымсыз желілер қазіргі уақытта тұрғын үй, коммерциялық және қоғамдық контексте байланысты қамтамасыз ететін жаһандық цифрлық инфрақұрылымның бөлігі ретінде кең таралған. Wi-Fi Protected Access III (WPA3) протоколының пайда болуы аутентификацияны жақсарту және криптографиялық қорғауды жақсарту арқылы сымсыз желі қауіпсіздігінің айтарлықтай ілгерілеуін көрсетті. Дегенмен, WPA3 ұсынған жетістіктерге қарамастан, сымсыз желінің қолжетімділігінде шабуылдаушылар пайдалана алатын осалдықтар әлі де бар. Нақтырақ айтсақ, аутентификациядан бас тарту шабуылдары IEEE 802.11 басқару шеңберіндегі осалдықтарды пайдаланады, олар шифрлаудың тұтастығына нұқсан келтірмейді, бірақ клиенттің желіге кіруіне кедергі келтіруі мүмкін. Бұл зерттеу wpa3 сымсыз желілерінің аутентификация шабуылдарына төзімділігін бағалауға, сондай-ақ осы шабуылдардың желінің қолжетімділігіне әсерін талдауға бағытталған. Зерттеу барысында wpa2 және WPA3 желілік конфигурациялары жасалған және имитацияланған аутентификация шабуылдары арқылы сыналған басқарылатын сымсыз сынақ алаңына негізделген эксперименттік тәсіл қолданылды. Нәтижелер WPA3 аутентификация мен шифрлауда айтарлықтай жақсартуларды қамтамасыз еткенімен, белгілі бір жағдайларда аутентификациядан бас тарту шабуылдары арқылы оның қолжетімділігіне теріс әсер етуі мүмкін екенін көрсетті. Зерттеу сымсыз желілерді қолжетімділік шабуылдарынан қорғауды жақсартудың ықтимал жолдарын, соның ішінде IEEE 802.11 басқару жүйелерін қорғауды күшейтуді және интрузияны анықтаудың жетілдірілген жүйелерін әзірлеуді анықтайды. Бұл зерттеудің мақсаты - аутентификациясыз шабуылдар кезіндегі заманауи сымсыз желілердің әрекетін бағалау және осы шабуылдардың заманауи қауіпсіздік стандарттарын пайдаланатын желілік инфрақұрылымдағы сымсыз желінің қолжетімділігіне әсерін бағалау. Тәжірибелер аутентификациясыз шабуылдар сымсыз желінің тұрақтылығына айтарлықтай әсер етуі мүмкін екенін, сымсыз құрылғылар мен кіру нүктелері арасындағы байланысты бұзуы мүмкін екенін көрсетеді.

Түйін сөздер: Сымсыз желі қауіпсіздігі, WPA3, Деаутентификация шабуылы, Wi-Fi қауіпсіздігі, Сымсыз желілерде интрузияны анықтау

© Ташенова Ж.М.^{*}, Габдуллин А.Р.¹, Абдугулова Ж.К.¹,
Аманжолова Ш.А.², Сантеева С.Ә.¹, 2026.

¹ Евразийский национальный университет имени Л.Н. Гумилева,
Астана, Казахстан;

² Казахская национальная консерватория имени Курмангазы,
Алматы, Казахстан.

E-mail: zhuldyz_tm@mail.ru

ОЦЕНКА БЕЗОПАСНОСТИ БЕСПРОВОДНЫХ СЕТЕЙ WPA3 В УСЛОВИЯХ АТАКИ С ДЕАВТЕНТИКАЦИЕЙ

Ташенова Жулдыз — PhD, кафедра информационной безопасности, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: zhuldyz_tm@mail.ru, <https://orcid.org/0000-0003-3051-1605>;

Габдуллин Абзал — кафедра информационной безопасности, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: anchorite.exe@gmail.com, <https://orcid.org/0000-0003-3051-1605>;

Абдугулова Жанат — доцент факультета информационных технологий, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: janat_6767@mail.ru, <https://orcid.org/0000-0001-7462-4623>;

Аманжолова Ширин — PhD, Казахская национальная консерватория имени Курмангазы, Алматы, Казахстан,
E-mail: schirin75@mail.ru, <https://orcid.org/0000-0002-6674-2766>;

Сантеева Сая — PhD, кафедра информационной безопасности, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: saya_santeeva@mail.ru, <https://orcid.org/0000-0001-9426-6704>.

Аннотация. *Актуальность.* Беспроводные сети являются важной частью современной цифровой инфраструктуры и обеспечивают связь в жилых, коммерческих, образовательных и общественных пространствах. Появление протокола Wi-Fi Protected Access III (WPA3) стало значимым этапом в развитии безопасности беспроводных сетей благодаря усовершенствованной аутентификации и криптографической защите. Вместе с тем даже при использовании WPA3 сохраняются риски, связанные с доступностью сети. В частности, атаки деаутентификации используют особенности кадров управления IEEE 802.11 и могут нарушать подключение клиентов к сети, не затрагивая напрямую криптографическую целостность передаваемых данных. *Цель.* Оценить устойчивость беспроводных сетей WPA3 к атакам деаутентификации и определить влияние таких атак на доступность беспроводной сети в условиях современной сетевой инфраструктуры. *Методы.* Исследование основано на экспериментальном подходе с использованием контролируемого беспроводного испытательного стенда. В рамках эксперимента были созданы и протестированы конфигурации сетей WPA2 и WPA3, после чего проведена имитация атак деаутентификации для оценки влияния на стабильность соединения, доступность сети и взаимодействие клиентских устройств с точкой доступа. Особое внимание

уделено анализу поведения сети при воздействии на кадры управления IEEE 802.11 и сравнению устойчивости различных конфигураций безопасности. *Результаты и выводы.* Результаты показали, что WPA3 обеспечивает существенные улучшения в части аутентификации и шифрования по сравнению с предыдущими стандартами, однако при определенных условиях атаки деаутентификации могут негативно влиять на доступность беспроводной сети. Проведенные эксперименты подтвердили, что такие атаки способны нарушать стабильность беспроводной связи, препятствовать корректному взаимодействию клиентских устройств с точками доступа и снижать надежность сетевой инфраструктуры. В исследовании определены потенциальные направления повышения защищенности беспроводных сетей от атак на доступность, включая усиление защиты кадров управления IEEE 802.11, применение механизмов мониторинга сетевых аномалий и разработку более эффективных систем обнаружения вторжений в беспроводных сетях. Практическая значимость исследования заключается в возможности использования полученных результатов при проектировании и оценке безопасности Wi-Fi-инфраструктуры в организациях, образовательных учреждениях и общественных сетях.

Ключевые слова: безопасность беспроводных сетей, WPA3, атака деаутентификации, безопасность Wi-Fi, IEEE 802.11, кадры управления, доступность сети, обнаружение вторжений в беспроводные сети

Introduction. Wireless Networks are now a central element in contemporary communication infrastructures, enabling both residential and enterprise users to transmit data via Wi-Fi. With the ever-increasing number of devices being added to wireless networks, and their increasing role as a means to provide internet access and support digital services, the security of wireless networks has emerged as a very important topic of research in modern network engineering and cybersecurity studies (Stallings et al., 2018).

The mechanism for communication in Wi-Fi networks, defined by the IEEE 802.11 standard, provides for communication between wireless clients and access points. Due to the open nature of radio transmission, wireless communication is inherently more susceptible to threats than wired networks. A threat actor who is within range of a wireless network's radio coverage can attempt to intercept data being transmitted, impersonate legitimate network devices, or disrupt communication between wireless clients and access points. This makes wireless networks particularly appealing to various types of cyberattacks, including eavesdropping, unauthorized access, and denial-of-service attacks (Mishra et al.).

There have been many different methods developed over the years to secure wireless communications. Security enhancements began with the Wired Equivalent Privacy (WEP) protocol, then evolved to Wi-Fi Protected Access (WPA) and Wi-Fi Protected Access II (WPA2), which included stronger encryption and authentication mechanisms. Although there have been significant enhancements

to the WPA2 protocol family, numerous researchers have found that WPA2 networks remain vulnerable to several security attacks. One example is the Key Reinstallation Attack (KRACK) that was discovered due to vulnerabilities in the WPA2 four-way handshake, allowing an attacker to manipulate encrypted traffic under specific circumstances (Vanhoeft et al., 2017).

As a response to the KRACK vulnerability, the Wi-Fi Alliance released the new WPA3 protocol, which includes better authentication mechanisms and improved cryptographic protections compared to its predecessor WPA2. Specifically, WPA3 replaces the Pre-Shared Key authentication method with the Simultaneous Authentication of Equals (SAE) protocol, designed to prevent offline dictionary attacks and improve password-based authentication. Unfortunately, subsequent research demonstrated that WPA3 is not invulnerable to all security vulnerabilities. Research published in “Dragonblood” demonstrated that weaknesses in the SAE handshake can be exploited using downgrade attacks and side-channel attacks (Vanhoeft et al., 2020).

In addition to authentication vulnerabilities, researchers have also identified vulnerabilities in the IEEE 802.11 standard itself. Design flaws in frame aggregation and fragmentation mechanisms allow an attacker to manipulate encrypted frames and inject malicious traffic into wireless networks, potentially impacting multiple generations of Wi-Fi security protocols, including WPA3 (Vanhoeft et al., 2021). These findings demonstrate that the security of wireless networks depends not only on the strength of encryption algorithms but also on the strength of the communication protocols.

Another type of attack that continues to exist in modern wireless networks are availability attacks. Deauthentication attacks utilize vulnerabilities in management frames to communicate between clients and access points. An attacker can cause clients to be disconnected from a wireless network by sending forged deauthentication frames. Because management frames have historically been transmitted without authentication, deauthentication attacks can continue to negatively impact modern wireless networks under certain configurations (Gonçalves et al., 2018).

Although WPA3 includes new features such as Protected Management Frames (PMF), research has indicated that availability attacks can continue to negatively impact wireless networks depending on how the network is configured and implemented. Therefore, evaluating the resiliency of WPA3 networks to deauthentication attack scenarios remains an important problem in wireless network security research.

This project evaluates the resiliency of WPA3 wireless networks to deauthentication attack scenarios and evaluates the effects of such attacks on network availability. Additionally, this project identifies limitations in modern wireless security mechanisms and outlines potential avenues for enhancing wireless network protection mechanisms in future wireless network architectures.

Literary review

Background of Wireless Network Security

Wireless networks have become an essential element of contemporary digital infrastructure; they provide flexibility and scalability for end-users, businesses and public services to communicate using internet connections, mobile communications, and to transfer data. Many Wi-Fi technologies (IEEE 802.11 standards) support this process. With the rapid proliferation of wireless communication technologies comes an increased number of connected devices including mobile phones, laptops, smart home systems and IoT devices. Thus, the security of wireless networks has emerged as a significant area of study in today's network engineering and cybersecurity fields (Stallings et al., 2018).

Wireless communication systems differ from wired counterparts since they use radio signals to transmit information, thus making them susceptible to potential security threats due to being physically exposed and accessible to anyone who falls within the wireless network's signal footprint. Those individuals can intercept wireless network transmissions, impersonate legitimate network devices, or prevent communication between clients and access points. As a result, wireless networks face many different types of security threats such as unauthorized access, interception of traffic, spoofing of devices, and DoS attacks (Bianchi et al., 2018).

In order to minimize these security risks, Wi-Fi networks implement security protocols to ensure that three fundamental security characteristics are met: confidentiality, integrity, and availability. Confidentiality ensures that data transmitted via the wireless network is encrypted. Integrity ensures that transmitted data cannot be altered or tampered with. Availability provides assurance to legitimate users that access will not be interrupted. Together, these security attributes represent the basis of wireless communication protection frameworks (Stallings et al., 2018).

Evolution of Wi-Fi Security Protocols

The advancements made in Wireless Communication Technologies have been matched by the continual enhancement of Security Mechanisms aimed at securing Wireless Networks. The First Wide-Scale Deployed Wireless Security Solution was Wired Equivalent Privacy (WEP), which was included in the Original IEEE 802.11 Standards. WEP was intended to provide confidentiality via the utilization of the RC4 Stream Cipher and Shared Encryption Keys. However, subsequent research revealed multiple flaws in WEP, including weaknesses in Key Management and the Reuse of Initialization Vectors; these flaws enabled an attacker to recover encryption keys by analyzing captured traffic, ultimately rendering WEP insecure and unsuitable for modern wireless networks (Mishra et al., 2018).

As an interim wireless security solution, Wi-Fi Protected Access (WPA) was developed to mitigate the limitations of WEP. WPA incorporated the Temporal Key Integrity Protocol (TKIP), improved key management and introduced message integrity protection solutions. Although WPA significantly enhanced wireless security compared to WEP, it was designed as a transitional protocol and retained many architectural limitations.

A significant improvement in wireless network protection was provided with the development of Wi-Fi Protected Access II (WPA2). WPA2 implemented the Advanced Encryption Standard (AES) using the Counter Mode Cipher Block Chaining Message Authentication Code Protocol (CCMP), providing stronger protection for the confidentiality and integrity of wireless communications. WPA2 became the dominant security standard for wireless networks for many years. However, further research demonstrated that WPA2 was not immune to security vulnerabilities, as demonstrated by the KRACK attack (Vanhoeft et al., 2017).

In response to these security challenges, the Wi-Fi Alliance developed Wi-Fi Protected Access III (WPA3), which incorporates improvements in authentication and encryption mechanisms. One of the main innovations of WPA3 is the Simultaneous Authentication of Equals (SAE) protocol, which replaces the traditional Pre-Shared Key authentication mechanism and provides stronger resistance against offline dictionary attacks. WPA3 also introduces enhanced encryption mechanisms and improved protection for management frames (Vanhoeft et al., 2020).

Although improvements have been made, recent studies indicate that modern wireless security protocols still contain potential vulnerabilities. The “DragonBlood” study demonstrated that weaknesses in the SAE authentication mechanism could be exploited through downgrade attacks and side-channel attacks, highlighting limitations in the current design of WPA3 authentication procedures. In addition, further research has identified structural weaknesses in IEEE 802.11 mechanisms such as frame aggregation and fragmentation (Hamad et al., 2021).

Existing Research on Wi-Fi Security

Wi-Fi Network Security has been a subject of study for over two decades, with extensive study regarding vulnerabilities in wireless protocols, authentication methods, and network architecture. Studies on Wi-Fi Security in the past analyzed vulnerabilities in the IEEE 802.11 protocol and reviewed the efficacy of current protective mechanisms. The discovery of the KRACK vulnerability is one of the greatest contributions to Wi-Fi Security Research, demonstrating that weaknesses in the WPA2 four-way handshake could allow an attacker to manipulate cryptographic keys (Ali et al., 2022).

After the release of WPA3, researchers started studying the security aspects of the new authentication methods. The Dragonblood research found that the SAE handshake in WPA3 was susceptible to downgrade and side-channel attacks when specific conditions existed (Vanhoeft et al., 2020). Further studies investigated the security features of the Dragonfly Handshake and demonstrated that both implementation details and the configuration of the protocol can greatly impact the overall security of a WPA3 network (Vanhoeft et al., 2020).

Another area of research is focused on attacks that target the availability of wireless networks. Deauthentication attacks use weaknesses in management frames utilized by wireless clients and access points to maintain communication.

An attacker can send forged deauthentication frames to cause the forced disconnection of wireless devices from their networks. Recently, researchers have been exploring more sophisticated methods for detecting and preventing wireless attacks, including machine learning-based intrusion detection systems (Hamad S. et al.).

Research Problem

Despite progress in development of Wi-Fi security protocols, modern Wi-Fi networks continue to experience a number of security challenges. The transition from earlier Wi-Fi protocols including WEP and WPA, to more advanced mechanisms including WPA2 and WPA3 has greatly enhanced confidentiality and authentication features of wireless communication. However, numerous studies have demonstrated that even the latest Wi-Fi security protocols may still contain weaknesses that can be exploited by attackers.

In addition to authentication related vulnerabilities, wireless networks are vulnerable to attacks targeting the operational features of wireless communication. One of the most well-known examples is the deauthentication attack which exploits weakness in management frames used in the IEEE 802.11 protocol (Wang et al., 2023). By transmitting forged deauthentication frames, an attacker can forcefully disconnect wireless clients from access points without directly breaking the encryption mechanisms of the security protocol (Carbajal et al., 2023).

Several studies have analyzed effectiveness of existing protection mechanisms against management frame-based attacks and found that deauthentication countermeasures may still be bypassed under certain conditions (Schepers et al., 2022). Though WPA3 introduces protection mechanisms such as Simultaneous Authentication of Equals (SAE) and Protected Management Frames (PMF), the practical resilience of modern Wi-Fi networks against availability-based attacks remains an important research challenge.

Research Objective

The main objective of this research is to evaluate how well modern wireless networks can recover from disruptions by means of a deauthentication attack, specifically for wireless networks secured using WPA3. The primary purpose was to test and assess the effect of a deauthentication attack on the ability of a wireless network to provide continuous connections and stable communications in both WPA2 and WPA3 environments. Furthermore, this research determines if current security protections are sufficient, and provides researchers and industry professionals with information that can be used to develop more secure wireless networks (Chatzoglou et al., 2022).

Scientific Novelty

The innovative aspect of this study are the results from the first overall assessment of resilience of current wireless networks against deauthentication attacks; particularly, it evaluated the performance of wireless networks using WPA3 security protocols. In contrast to many past studies based upon protocol design defects or theoretical models, this study conducted an experimentally-

based assessment of the impacts of deauthentication attacks in real-world wireless networks, allowing for a better understanding of how resilient today's wireless security mechanisms are.

A second novel contribution is the combination of experimental assessments with an evaluation of the prior literature on wireless security vulnerabilities and attack detection methods. The research contributes to the ongoing debate on the effectiveness of current Wi-Fi security protocols by showing that improved authentication and encryption protocols alone do not preclude all types of wireless attacks.

Materials and Methods

The current section explains the experimental method used to measure the resistance of modern wireless networks under deauthentication attacks. In particular, it describes the study of the behaviors of wireless networks working under WPA2 and WPA3 security protocols which are submitted to management-frame-based attacks that target their availability (Chatzoglou et al., 2021).

The experimental study was carried out in a controlled laboratory setting so as to reproduce the results of the tests and accurately monitor wireless traffic. The research method included the establishment of a wireless testbed, the implementation of deauthentication attacks and the collection and analysis of network traffic through packet inspections (Natkaniec et al., 2022).

Research Methodology

The research methodology employed within this study is based upon an experimental approach aimed at studying the effect of deauthentication attacks on modern wireless networks. The experimental methodology consists of multiple phases. At first, a controlled wireless network setting was established in order to simulate normal wireless communications settings. The experimental setup consisted of a wireless access point that worked under both WPA2 and WPA3 security protocols; a wireless client device connected to the wireless network and a monitoring device used to capture and analyze wireless traffic.

Then deauthentication attacks were implemented in order to study the responses of the wireless network to the malicious management frames. The deauthentication attacks were implemented by sending fake deauthentication frames to disrupt the communication between the wireless client and the access point. During the experiment, wireless traffic was recorded in monitoring mode so as to be able to record all the frames exchanged in the wireless network (Suryadi et al., 2021).

All the captures have been analyzed using packet inspection tools in order to identify authentication frames, management frames and deauthentication events that occurred during the attack. The results of the experiments have been used to study the resistance of wireless networks under different security configurations (Carbajal et al., 2021).

Experimental Testbed

The experimental laboratory setting simulated a typical wireless communication setting and was developed to provide an experimental testbed that could replicate

the wireless network operations found in typical networks but allow for close examination and monitoring of wireless network traffic during an active attack.

There were three primary elements to the experimental setup: (1) A Wireless Access Point (WAP), which was configured to operate with current Wi-Fi security protocols; (2) A Client Device that was attached to the wireless network via a wireless interface on the WAP; (3) A Monitoring Device, which was placed into “monitoring mode” so that the wireless traffic could be captured.

The wireless infrastructure in this experimental environment was provided by a wireless access point utilizing the OpenWrt firmware. OpenWrt is an open source operating system for network devices, such as wireless routers, that provides users with flexibility in configuring their wireless devices and also provides options for configuring their wireless security settings and protocols. The wireless access point utilized in this experimental testbed is shown in Figure 1.



Figure 1- OpenWrt-based wireless access point used in the experimental environment.

During the normal operational phase, as well as during the Deauthentication Attack phases, the monitoring device was able to capture the wireless traffic being transmitted through the wireless network. All of the captured packets were then saved in PCAP format and were analyzed using various packet inspection tools to determine the behavior of the authentication exchanges and the behavior of the management frames and deauthentication events.

Network Configuration

Two configurations were set up in the experimental testing environment, each based on different security protocols: a network secured by the WPA2 protocol, and a network operating under the WPA3 security protocol. The network configurations were set up to allow researchers to analyze how the current generation of wireless security mechanisms respond to de-authentication attacks.

The WPA2 network was configured to operate under the pre-shared key (PSK) method of authentication using Advanced Encryption Standard (AES) for data encryption. To verify the WPA2 network configuration parameters, the authors analyzed beacon frames and RSN Information Elements captured in the wireless traffic. The decoded RSN parameters indicated that the PSK authentication suite was used. An example of RSN parameter decoding for the WPA2 network is illustrated in Figure 2.

```

@bzeldgabdellin: ~$ cd ~/pcaps_2025
PCAP=wpa2_psk_12345678_2025-10.pcapng
tshark -r "$PCAP" -Y "wlan.fc.type_subtype==0x08 && wlan.rsn.akms" \
-v | egrep -n "SSID:RSN capabilities|AKM Suites|Authentication Key Management|Suite|WPA version" | head -n 80
154:      SSID: "test-wmm-rsn"
198:      Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
199:      Group Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
200:      Group Cipher Suite type: AES (CCM) (4)
201:      Pairwise Cipher Suite Count: 1
202:      Pairwise Cipher Suite List 00:0f:ac (Ieee 802.11) AES (CCM)
203:      Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
204:      Pairwise Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
205:      Pairwise Cipher Suite type: AES (CCM) (4)
206:      Auth Key Management (AKM) Suite Count: 1
208:      Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) PSK
367:      .0. .... = Multiple BSSID: Not supported
395:      ....0 = UTF-8 SSID: Not supported
634:      SSID: "test-wmm-rsn"
678:      Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
679:      Group Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
680:      Group Cipher Suite type: AES (CCM) (4)
681:      Pairwise Cipher Suite Count: 1
682:      Pairwise Cipher Suite List 00:0f:ac (Ieee 802.11) AES (CCM)
683:      Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
684:      Pairwise Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
685:      Pairwise Cipher Suite type: AES (CCM) (4)
686:      Auth Key Management (AKM) Suite Count: 1
688:      Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) PSK
847:      .0. .... = Multiple BSSID: Not supported
875:      ....0 = UTF-8 SSID: Not supported
1114:      SSID: "test-wmm-rsn"
1158:      Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
1159:      Group Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
1160:      Group Cipher Suite type: AES (CCM) (4)
1161:      Pairwise Cipher Suite Count: 1
1162:      Pairwise Cipher Suite List 00:0f:ac (Ieee 802.11) AES (CCM)
1163:      Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
1164:      Pairwise Cipher Suite OUI: 00:0f:ac (Ieee 802.11)
1165:      Pairwise Cipher Suite type: AES (CCM) (4)
1166:      Auth Key Management (AKM) Suite Count: 1
1168:      Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) PSK
1327:      .0. .... = Multiple BSSID: Not supported
1355:      ....0 = UTF-8 SSID: Not supported

```

Figure 2- RSN information elements of the WPA2 network indicating the use of the PSK authentication suite.

A second wireless network was established using the WPA3 protocol, which builds upon prior wireless security technologies with additional enhancements including the Simultaneous Authentication of Equals (SAE) protocol (Vanhoef et al., 2017). In the experimental environment, the WPA3 network configuration was set-up using SAE authentication with Protected Management Frames enabled. The authors verified the RSN parameters of the WPA3 network by analyzing beacon frames captured during operation. An example of RSN parameter decoding for the WPA3 configuration is illustrated in Figure 3.

```
shazal@abdullin:~/pcaps_2025$ cd ~/pcaps_2025
PCAP==wpa3_2025-10.pcapng

tshark -r "PCAP" -Y "wlan.fc.type_subtype==0x00 && wlan.rsn.akms" \
-V | egrep -n "SSID:[Auth Key Management|AKM|SAE|Suite: 00:0f:ac]" | head -n 120
183:   SSID: "WPA3-Network"
197:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
202:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
205:   Auth Key Management (AKM) Suite Count: 1
206:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
207:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
208:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
209:   Auth Key Management (AKM) type: SAE (SHA256) (0)
233:   .1. .... = Multiple BSSID: Supported
281:   ....0 = UTF-8 SSID: Not supported
305:   SSID: "WPA3-Network"
349:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
354:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
357:   Auth Key Management (AKM) Suite Count: 1
358:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
359:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
360:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
361:   Auth Key Management (AKM) type: SAE (SHA256) (0)
365:   .1. .... = Multiple BSSID: Supported
367:   ....0 = UTF-8 SSID: Not supported
391:   SSID: "WPA3-Network"
396:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
399:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
400:   Auth Key Management (AKM) Suite Count: 1
401:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
402:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
403:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
404:   Auth Key Management (AKM) type: SAE (SHA256) (0)
407:   .1. .... = Multiple BSSID: Supported
409:   ....0 = UTF-8 SSID: Not supported
433:   SSID: "WPA3-Network"
438:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
441:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
442:   Auth Key Management (AKM) Suite Count: 1
443:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
444:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
445:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
446:   Auth Key Management (AKM) type: SAE (SHA256) (0)
449:   .1. .... = Multiple BSSID: Supported
451:   ....0 = UTF-8 SSID: Not supported
475:   SSID: "WPA3-Network"
480:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
483:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
484:   Auth Key Management (AKM) Suite Count: 1
485:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
486:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
487:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
488:   Auth Key Management (AKM) type: SAE (SHA256) (0)
491:   .1. .... = Multiple BSSID: Supported
493:   ....0 = UTF-8 SSID: Not supported
517:   SSID: "WPA3-Network"
522:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
525:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
526:   Auth Key Management (AKM) Suite Count: 1
527:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
528:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
529:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
530:   Auth Key Management (AKM) type: SAE (SHA256) (0)
533:   .1. .... = Multiple BSSID: Supported
535:   ....0 = UTF-8 SSID: Not supported
559:   SSID: "WPA3-Network"
564:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
567:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
568:   Auth Key Management (AKM) Suite Count: 1
569:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
570:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
571:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
572:   Auth Key Management (AKM) type: SAE (SHA256) (0)
575:   .1. .... = Multiple BSSID: Supported
577:   ....0 = UTF-8 SSID: Not supported
601:   SSID: "WPA3-Network"
606:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
609:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
610:   Auth Key Management (AKM) Suite Count: 1
611:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
612:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
613:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
614:   Auth Key Management (AKM) type: SAE (SHA256) (0)
617:   .1. .... = Multiple BSSID: Supported
619:   ....0 = UTF-8 SSID: Not supported
643:   SSID: "WPA3-Network"
648:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
651:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
652:   Auth Key Management (AKM) Suite Count: 1
653:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
654:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
655:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
656:   Auth Key Management (AKM) type: SAE (SHA256) (0)
659:   .1. .... = Multiple BSSID: Supported
661:   ....0 = UTF-8 SSID: Not supported
685:   SSID: "WPA3-Network"
690:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
693:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
694:   Auth Key Management (AKM) Suite Count: 1
695:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
696:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
697:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
698:   Auth Key Management (AKM) type: SAE (SHA256) (0)
699:   .1. .... = Multiple BSSID: Supported
701:   ....0 = UTF-8 SSID: Not supported
725:   SSID: "WPA3-Network"
730:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
733:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
734:   Auth Key Management (AKM) Suite Count: 1
735:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
736:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
737:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
738:   Auth Key Management (AKM) type: SAE (SHA256) (0)
741:   .1. .... = Multiple BSSID: Supported
743:   ....0 = UTF-8 SSID: Not supported
767:   SSID: "WPA3-Network"
772:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
775:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
776:   Auth Key Management (AKM) Suite Count: 1
777:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
778:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
779:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
780:   Auth Key Management (AKM) type: SAE (SHA256) (0)
783:   .1. .... = Multiple BSSID: Supported
785:   ....0 = UTF-8 SSID: Not supported
809:   SSID: "WPA3-Network"
814:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
817:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
818:   Auth Key Management (AKM) Suite Count: 1
819:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
820:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
821:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
822:   Auth Key Management (AKM) type: SAE (SHA256) (0)
825:   .1. .... = Multiple BSSID: Supported
827:   ....0 = UTF-8 SSID: Not supported
851:   SSID: "WPA3-Network"
856:   Group Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
859:   Pairwise Cipher Suite: 00:0f:ac (Ieee 802.11) AES (CCM)
860:   Auth Key Management (AKM) Suite Count: 1
861:   Auth Key Management (AKM) List 00:0f:ac (Ieee 802.11) SAE (SHA256)
862:   Auth Key Management (AKM) Suite: 00:0f:ac (Ieee 802.11) SAE (SHA256)
863:   Auth Key Management (AKM) OUI: 00:0f:ac (Ieee 802.11)
864:   Auth Key Management (AKM) type: SAE (SHA256) (0)
867:   .1. .... = Multiple BSSID: Supported
869:   ....0 = UTF-8 SSID: Not supported
```

Figure 3- RSN information elements of the WPA3 network showing the SAE authentication suite.

The wireless network configurations were set up and managed through the OpenWrt administration interface, where it is possible to have complete control of all wireless parameters including security mode, channel configuration, and authentication settings. The OpenWrt administration interface for setting up the wireless experimental network configuration is shown in Figure 4.

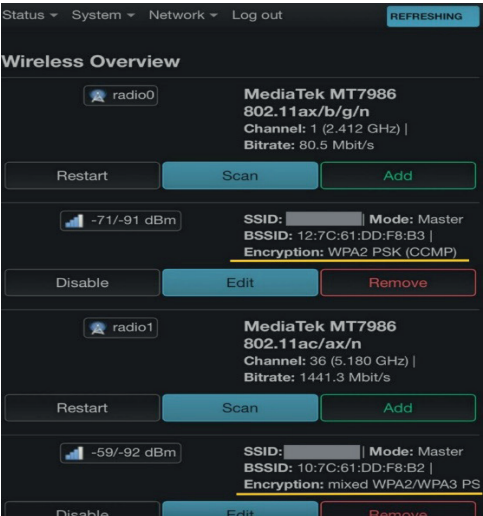


Figure 4- Wireless configuration interface of the OpenWrt-based access point used in the experimental testbed.

Attack Modeling

This study's attack focus area is a deauthentication attack focused on disrupting the availability of wireless communication. Deauthentication attacks take advantage of the characteristics of IEEE 802.11 Management Frames, which handle the association between wireless clients and access points and are unencrypted.

Wireless clients periodically exchange management frames with their associated access point to maintain their active association. When a wireless client intentionally disconnects from a network, it sends a deauthentication frame. Due to the nature of IEEE 802.11 protocols, however, an attacker may create a deauthentication frame using forged management frames that represents itself as either the client or the access point. The target device will then see the forged frame as a valid request and subsequently disconnect from the network.

To model the attack scenario, the study employed a wireless network consisting of a single access point and one wireless client device. The attack was simulated by transmitting forged deauthentication management frames toward the target wireless client device. The attack was performed using a wireless adapter designed for packet injection. During the transmission of the deauthentication frames, a monitoring device captured all of the wireless traffic that occurred between devices on the network.

Data Collection and Traffic Analysis

The collection and analysis of Wi-Fi traffic were an important component of the methodology. To assess the effect of deauth-attacks on the performance of a Wi-Fi network, all of the 802.11 wireless frames sent or received within the scope of the experiments were captured and analyzed through packet inspection techniques. Wi-Fi traffic was captured via a monitor station containing a wireless NIC functioning in "monitor" mode, which allows a device to capture all of the IEEE 802.11 frames including Management Frames, Control Frames and Data Frames.

Frames that were captured were saved as PCAP or PCAPNG files, which preserve detailed metadata for each frame, including timestamps, frame type information, and protocol hierarchy information. A depiction of metadata collected during a capture of traffic is illustrated in Figure 5.

```

azrat@gebdellin:~$ pcaps_2025$ cd ~/pcaps_2025
for f in wpa2_psk_12345678_2025-10.pcapng wpa3_2025-10.pcapng krack_example_ft_2025-10.pcapng pmf_deauth_example_2025-10.pcapng; do
  echo "----- $f -----"
  capinfos -sf | egrep -i "File name|Number of packets|Capture duration|Earliest packet time|Latest packet time"
  echo
done
----- wpa2_psk_12345678_2025-10.pcapng -----
File name:      wpa2_psk_12345678_2025-10.pcapng
Number of packets: 39
Capture duration: 0.275304454 seconds
Earliest packet time: 2025-10-15 12:00:00.000000072
Latest packet time: 2025-10-15 12:00:00.275304526
Number of packets = 39
----- wpa3_2025-10.pcapng -----
File name:      wpa3_2025-10.pcapng
Number of packets: 167
Capture duration: 13.414401502 seconds
Earliest packet time: 2025-10-15 12:00:00.000000057
Latest packet time: 2025-10-15 12:00:00.134401639
Number of packets = 167
----- krack_example_ft_2025-10.pcapng -----
File name:      krack_example_ft_2025-10.pcapng
Number of packets: 1378
Capture duration: 48.459441261 seconds
Earliest packet time: 2025-10-15 12:00:00.000000001
Latest packet time: 2025-10-15 12:00:00.459441262
Number of packets = 1378
----- pmf_deauth_example_2025-10.pcapng -----
File name:      pmf_deauth_example_2025-10.pcapng
Number of packets: 37
Capture duration: 6.778007177 seconds
Earliest packet time: 2025-10-15 12:00:00.000000070
Latest packet time: 2025-10-15 12:00:06.778007247
Number of packets = 37
azrat@gebdellin:~$ pcaps_2025$

```

Figure 5- PCAP/PCAPNG capture metadata and timestamp parameters.

When analyzing the WPA3 network configuration, the presence of EAPOL (Extensible Authentication Protocol Over LAN) frames served as indicators that authentication was active on the wireless network. These frames represent the key exchange procedures and authentication processes that take place between the client device and the Access Point (AP). A representation of the EAPOL frames captured in the WPA3 network traffic is depicted in Figure 6.

```
abzal@gabdullin: ~/pcaps_2025
PCAP="wpa3_2025-10.pcapng"

echo -n "EAPOL frames count (WPA3): "
tshark -r "$PCAP" -Y "eapol" | wc -l

tshark -r "$PCAP" -Y "eapol" \
-T fields -E header=y -E separator='\t' \
-e frame.number -e frame.time -e wlan.sa -e wlan.da -e eapol.type | head -n 25
EAPOL frames count (WPA3): 4
frame.number    frame.time      wlan.sa         wlan.da         eapol.type      d2:c6:b4:ab:58:88  3
92      Oct 15, 2025 12:00:06.711381602 +05      e2:20:ae:cb:03:04      d2:c6:b4:ab:58:88  3
94      Oct 15, 2025 12:00:06.715514396 +05      e2:20:ae:cb:03:04      d2:c6:b4:ab:58:88  3
96      Oct 15, 2025 12:00:06.715832796 +05      e2:20:ae:cb:03:04      d2:c6:b4:ab:58:88  3
98      Oct 15, 2025 12:00:06.716112352 +05      e2:20:ae:cb:03:04      d2:c6:b4:ab:58:88  3
abzal@gabdullin:~/pcaps_2025$
```

Figure 6- Captured EAPOL frames in the WPA3 traffic record that indicate the presence of handshake artifacts during authentication.

Results

This section presents the traffic captured from the wireless network while the deauthentication attack was being conducted. The traffic was analyzed for its effects upon the operation of the network. Specifically, this work analyzes traffic traces collected by the monitoring system during the execution of the deauthentication attack for the presence of deauthentication frames and assesses the effectiveness of these frames in interrupting or disrupting the wireless connection established by the client.

Execution of Deauthentication Attack

The deauthentication attack simulated in this study consisted of transmitting forged IEEE 802.11 management frames toward a wireless client that had previously been authenticated to the experimental access point. The objective of the attack was to provide a realistic model of an availability-based attack and to determine how a wireless network would respond to malicious management frames.

Deauthentication frames were transmitted toward the client device utilizing a wireless adapter configured to perform packet injection. These frames were forged to appear as if they were legitimate communication between the access point and the client device and caused the client device to interpret them as a valid request to terminate the wireless association.

Examples of the captured deauthentication management frames are illustrated in Figures 7 and 8. These figures represent the decoding of the IEEE 802.11 frame fields observed in the captured wireless traffic trace.

```

abzalagabullin:~pcaps_2025$ cd ~/pcaps_2025
PCAP=pmf_deauth_example_2025-10.pcapng

tshark -r "SPCAP" -Y "frame.number==36" -V | head -n 140
Frame 36: 64 bytes on wire (512 bits), 64 bytes captured (512 bits) on interface hwsim0, id 0
Section number: 1
Interface id: 0 (hwsim0)
Interface name: hwsim0
Encapsulation type: IEEE 802.11 plus radiotap radio header (23)
Arrival Time: Oct 15, 2025 12:00:06.778804566 +05
UTC Arrival Time: Oct 15, 2025 07:00:06.778804566 UTC
Epoch Arrival Time: 1760511606.778804566
[Time shift for this packet: 0.000000000 seconds]
[Time delta from previous captured frame: 0.002377856 seconds]
[Time delta from previous displayed frame: 0.000000000 seconds]
[Time since reference or first frame: 6.778804496 seconds]
Frame Number: 36
Frame length: 64 bytes (512 bits)
Capture Length: 64 bytes (512 bits)
[Frame is marked: False]
[Frame is ignored: False]
[Protocols in frame: radiotap:wlan_radio:wlan:data]
Radiotap Header v0, Length 22
Header revision: 0
Header pad: 0
Header length: 22
Present flags
  Present flags word: 0x0000000f
  .....0.....1 = TSFT: Present
  .....0.....1 = Flags: Present
  .....0.....1 = Rate: Present
  .....0.....1 = Channel: Present
  .....0.....0 = FHSS: Absent
  .....0.....0 = dBm Antenna Signal: Absent
  .....0.....0 = dBm Antenna Noise: Absent
  .....0.....0 = Lock Quality: Absent
  .....0.....0 = TX Attenuation: Absent
  .....0.....0 = dB TX Attenuation: Absent
  .....0.....0 = dBm TX Power: Absent
  .....0.....0 = Antenna: Absent
  .....0.....0 = dB Antenna Signal: Absent
  .....0.....0 = dB Antenna Noise: Absent
  .....0.....0 = RX flags: Absent
  .....0.....0 = TX flags: Absent
  .....0.....0 = data retries: Absent
  .....0.....0 = Channels: Absent
  .....0.....0 = MCS information: Absent
  .....0.....0 = A-MPDU Status: Absent
  .....0.....0 = VHT information: Absent
  .....0.....0 = frame timestamp: Absent
  .....0.....0 = HE information: Absent
  .....0.....0 = HE-MU information: Absent
  .....0.....0 = Length PDU: Absent
  .....0.....0 = L-SIG: Absent
  .....0.....0 = Reserved: 0x0
  .....0.....0 = TIVS: Absent
  .....0.....0 = Radiotap NS next: False
  .....0.....0 = Vendor NS next: False
  .....0.....0 = Ext: Absent
MAC timestamp: 1681570102201666
Flags: 0x00
  .....0 = CFP: False
  .....0 = Preamble: Long
  .....0 = WEP: False
  .....0 = Fragmentation: False
  .....0 = FCS at end: False
  .....0 = Data Pad: False
  .....0 = End FCS: False
  .....0 = Short GI: False

```

Figure 7- Detailed decoding of deauthentication management frame (part 1).

```

Data Rate: 1.0 Mb/s
Channel Frequency: 2412 [2.4 GHz 1]
Channel Flags: 0x0000, Complementary Code Keying (CCX), 2 GHz spectrum
  .....0 = 700 MHz spectrum: False
  .....0 = 800 MHz spectrum: False
  .....0 = 900 MHz spectrum: False
  .....0 = Turbo: False
  .....1 = Complementary Code Keying (CCX): True
  .....0 = Orthogonal Frequency-Division Multiplexing (OFDM): False
  .....1 = 2 GHz spectrum: True
  .....0 = 5 GHz spectrum: False
  .....0 = Passive: False
  .....0 = Dynamic CCK-OFDM: False
  .....0 = Gaussian Frequency Shift Keying (GFSK): False
  .....0 = GFSK (nonpipe): False
  .....0 = Static Turbo: False
  .....0 = Half rate Channel (10MHz Channel Width): False
  .....0 = Quarter Rate Channel (5MHz Channel Width): False
802.11 radio information
  Mfr type: 802.11b (HR/DSSS) (4)
  Short preamble: False
  Data rate: 1.0 Mb/s
  Channel: 1
  Frequency: 2412MHz
  TSF timestamp: 1681570102201666
  [Duration: 529us]
  [Preamble: 192us]
  [TSF: 1681570102201139us]
  [Start: 1681570102201139us]
  [End: 1681570102201666us]
IEEE 802.11 deauthentication, Flags: p.....
Type/Subtype: Deauthentication (ex000c)
Frame Control field: 0x0000
  .....00 = Version: 0
  .....00 = Type: Management frame (0)
  .....1100 = Subtype: 12
  Flags: 0x00
  .....00 = DS status: Not leaving DS or network is operating in AD-HOC mode (To DS: 0 From DS: 0) (0x0)
  .....0 = More Fragments: This is the last fragment
  .....0 = Retry: Frame is not being retransmitted
  .....0 = PMF Mfr: STA will stay up
  .....0 = More Data: No data buffered
  .....1 = Protected flag: Data is protected
  .....0 = HT/Order flag: Not strictly ordered
  .....0000 0001 0011 1010 = Duration: 314 microseconds
Receiver address: 02:00:00:00:00:00 (02:00:00:00:00:00)
  .....0.....0 = IG bit: Locally administered address (this is NOT the factory default)
  .....0.....0 = IG bit: Individual address (unicast)
Destination address: 02:00:00:00:00:00 (02:00:00:00:00:00)
  .....0.....0 = IG bit: Locally administered address (this is NOT the factory default)
  .....0.....0 = IG bit: Individual address (unicast)
Transmitter address: 02:00:00:00:00:00 (02:00:00:00:00:00)
  .....0.....0 = IG bit: Locally administered address (this is NOT the factory default)
  .....0.....0 = IG bit: Individual address (unicast)
Source address: 02:00:00:00:00:00 (02:00:00:00:00:00)
  .....0.....0 = IG bit: Locally administered address (this is NOT the factory default)
  .....0.....0 = IG bit: Individual address (unicast)
BSS id: 02:00:00:00:00:00 (02:00:00:00:00:00)
  .....0.....0 = IG bit: Locally administered address (this is NOT the factory default)
  .....0.....0 = IG bit: Individual address (unicast)
.....0000 = Fragment number: 0
0000 0101 0011 = Sequence number: 03
[WLAN Flags: p.....]
COW parameters
  COW Ext. Initialization Vector: 0x000000000001
  Key Index: 0
Data (10 bytes)
0000 28 95 09 04 76 a1 ed 6b 3f 08 .....v.k?
Data: 2895090476a1ed6b3f08
[Length: 10]
abzalagabullin:~pcaps_2025$

```

Figure 8- Detailed decoding of deauthentication management frame (part 2).

The decoded frame fields verify that the transmitted frames are IEEE 802.11 deauthentication management frames that include frame control fields, transmitter and receiver MAC addresses, and reason codes that indicate the termination of a wireless association. Analysis of the captured wireless traffic indicated that the injected deauthentication frames successfully terminated communication between the wireless client and the access point.

Results for WPA2 Network

The initial test case used a wireless network operating under the WPA2 security standard. In normal network operation, the wireless client remained associated to the access point, and the captured trace indicated typical IEEE 802.11 network communications, including beacon frames, authentication dialogues and encrypted data transfer.

During the de-authentication attack, forged de-authentication management frames were sent towards the wireless client device, which was currently connected to the access point. The received frames led to the client device losing its association with the network, and attempting to re-associate again. An examination of the captured trace demonstrated that the WPA2 network was susceptible to the described de-authentication attack.

The numerous de-authentication frames captured during the attack could be easily identified through the analysis of the IEEE 802.11 management frame fields and reason codes present within each packet. Therefore, the results obtained from this study demonstrate that the WPA2 network is still vulnerable to de-authentication attacks on network availability, but unlike traditional man-in-the-middle (MitM) attacks, no encryption keys nor authentication credentials were compromised.

Results for WPA3 Network

The second set of experiments used a network configured with the WPA3 security protocol. During normal operation the network demonstrated stable connectivity between the wireless client and access point. Analysis of captured traffic traces from the PCAP file showed standard IEEE 802.11 communication patterns including beacon frames, authentication exchanges and encrypted data transmissions.

The WPA3 protocol has several advantages over earlier wireless security methods, notably in the area of authentication. The SAE protocol was designed to improve password-based authentication by providing better protection against offline dictionary attacks (Vanhoeft M. et al., 2020). Despite these improvements, the experimental results show that attacks targeting management frame behavior may still affect the availability of Wi-Fi networks.

During the execution of the deauthentication attack, forged IEEE 802.11 deauthentication frames were sent towards the Wi-Fi client device connected to the access point. Analysis of the captured traffic traces confirmed that the deauthentication frames caused the connection termination events on the client device. The Wi-Fi client interpreted the forged management frames as legitimate

requests to terminate the connection and subsequently attempted to reconnect to the access point.

Impact on Network Availability

Experimental results from deauthentication attack simulations show significant disruption in availability of wireless network communications. Analysis of captured traffic showed that injecting forged deauthentication management frames caused the wireless client to be disconnected from the access point multiple times and attempt to reconnect. During the simulation, the client continuously lost and then tried to regain connection to the AP through WPA2 and WPA3 network configurations.

Although the WPA3 protocol includes enhanced authentication mechanisms compared to past wireless security protocols, the experimental results indicated that it does not provide complete protection against availability-based attacks. The results confirmed that management-frame based attacks can disrupt wireless connectivity even when using current generation wireless security.

```
abzal@gabdullin: ~/pcaps_2025$ cd ~/pcaps_2025
PCAP="pmf_deauth_example_2025-10.pcapng"
echo -n "Deauth frames: "
tshark -r "$PCAP" -Y "wlan.fc.type_subtype==0x0c" | wc -l
Deauth frames: 1
abzal@gabdullin: ~/pcaps_2025$
```

Figure 9- Deauthentication frame activity observed in captured wireless traffic during attack execution.

Table 1 – Impact of deauthentication attack on wireless network availability

Parameter	WPA2 Network	WPA3 Network
Successful client disconnection	Yes	Yes
Reconnection attempts observed	Yes	Yes
Presence of deauthentication frames in traffic	High	High
Impact on network availability	Significant	Significant
Authentication mechanism affected	No	No
Encryption compromised	No	No

Table 1 shows the results of the experiment clearly indicating that the main effect of the deauthentication attack is to interfere with the operation of the wireless network, rather than to break either the authentication mechanism or the encryption mechanism. Thus, the deauthentication attack targeted the availability of the network rather than the confidentiality or integrity of the transmitted data.

```

abzal@gabdullin: ~/pcaps_2025$ set -euo pipefail
cd ~/pcaps_2025
mkdir -p ~/pcaps_out

OUT=~/pcaps_out/results_table.txt

{
  printf "%-34s %8s %12s %10s %8s %8s %8s %8s\n" \
    "FILE" "PKTS" "DUR(s)" "PPS" "MGMT" "DATA" "EAPOL" "DEAUTH"

  for f in \
    wpa2_psk_12345678_2025-10.pcapng \
    wpa3_2025-10.pcapng \
    pmf_deauth_example_2025-10.pcapng \
    krack_example_ft_2025-10.pcapng
  do
    # PKTS with commas/spaces removed (e.g., "1,378" -> "1378")
    pkts=$(capinfos "$f" | awk -F: '/Number of packets/ {gsub(/[ ,]/, ""); print $2}')
    # duration seconds
    dur=$(capinfos "$f" | awk -F: '/Capture duration/ {gsub(/ seconds/, ""); gsub(/ ^ +/, ""); print $2}')
    # PPS (safe if dur=0)
    pps=$(python3 - <<PY
pkts=float("$pkts")
dur=float("$dur") if float("$dur")>0 else 1.0
print(f"{pkts/dur:.2f}")
PY
)
    # Counts from tshark (never crash on parse errors)
    mgmt=$(tshark -r "$f" -Y "wlan.fc.type==0" 2>/dev/null | wc -l | tr -d ' ')
    data=$(tshark -r "$f" -Y "wlan.fc.type==2" 2>/dev/null | wc -l | tr -d ' ')
    eapol=$(tshark -r "$f" -Y "eapol" 2>/dev/null | wc -l | tr -d ' ')
    deauth=$(tshark -r "$f" -Y "wlan.fc.type_subtype==0x0c" 2>/dev/null | wc -l | tr -d ' ')

    printf "%-34s %8s %12s %10s %8s %8s %8s %8s\n" \
      "$f" "$pkts" "$dur" "$pps" "$mgmt" "$data" "$eapol" "$deauth"
  done
} | tee "$OUT"

echo
echo "Saved: $OUT"

```

FILE	PKTS	DUR(s)	PPS	MGMT	DATA	EAPOL	DEAUTH
wpa2_psk_12345678_2025-10.pcapng	39	0.275304454	141.66	16	6	4	2
wpa3_2025-10.pcapng	167	13.414401582	12.45	151	4	4	1
pmf_deauth_example_2025-10.pcapng	37	6.778807177	5.46	19	5	5	1
krack_example_ft_2025-10.pcapng	1378	48.459441261	28.44	1089	141	4	1

Saved: /home/abzal/pcaps_out/results_table.txt

Figure 10 – Comparative packet statistics observed during normal operation and deauthentication attack scenarios.

These results identify an important aspect of contemporary wireless security mechanisms: that although recent generations of Wi-Fi standards have greatly improved the level of authentication and encryption security, management-frame based attacks could potentially affect the stability of wireless networks.

Discussion

In order to better understand the behaviors of wireless communications in response to an attack on the deauthentications, the researchers analyzed the captured packets from their experiment and identified how much of an impact an attack that uses forged deauthenticating frames can cause to the availability of wireless communications. The forged deauthenticating frames resulted in the access point repeatedly disconnecting from the client device, which resulted in a disruption in the network's availability until the access point was able to reconnect back to the client device.

Additionally, according to the researcher's findings, most of the improvements in the WPA3 protocol were focused on the authentication and cryptographic aspects. The Simultaneous Authentication of Equals (SAE) protocol is designed to improve the security of the authentication mechanism and prevent a user's password from being guessed (Vanhoeft et al., 2020). However, the results found in this study demonstrate that even though the SAE protocol provides improved security against authentication-related attacks, there are still ways to disrupt the

normal operation of a wireless network through the use of a deauthenticating frame.

Interpretation of Experimental Results

The experimental results clearly demonstrate that deauthentication attacks can disrupt wireless network availability regardless of the Wi-Fi security protocol in use. This behavior is in-line with the way the IEEE 802.11 protocol works; the 802.11 standard states that management frames are used by network devices to determine if they are associated with each other (Dalal et al., 2023).

Previous studies have shown that similar types of attacks, including deauthenticating frames, are still possible and can disrupt the normal functioning of a wireless network, due to the fact that these attacks target the control mechanisms of the wireless protocol and not the encryption methods (Chadee et al., 2022). Therefore, even if a wireless network has implemented strong authentication mechanisms, a deauthenticating frame attack can potentially disrupt its operation.

Furthermore, the researchers identified that the deauthenticating frames produced visible anomalies in the normal patterns of wireless traffic. The anomalies that occur as a result of deauthenticating frames may be useful for developing intrusion detection mechanisms and wireless monitoring systems that can detect these types of attacks (Moharam et al., 2024).

Limitations of WPA3 Security Mechanisms

WPA3 introduces significant improvements to wireless network security; however, it does not fully eliminate all potential vulnerabilities. The protocol primarily focuses on strengthening authentication mechanisms and improving protection against password-based attacks. These improvements represent an important advancement compared to previous wireless security standards, but they do not address all possible attack vectors within wireless communication systems.

The experimental results indicate that attackers may still influence wireless network connectivity by targeting management frame exchanges, particularly through deauthentication attacks. These attacks exploit the control mechanisms responsible for maintaining associations between wireless clients and access points. Because these mechanisms operate independently from the main encryption process, they may still be used to disrupt the availability of wireless communication.

Beyond authentication-related limitations, potential vulnerabilities may also arise from other aspects of the IEEE 802.11 protocol design. Previous studies have identified weaknesses related to frame aggregation and fragmentation mechanisms, which were originally introduced to improve network performance but may introduce additional security risks. These observations suggest that improving wireless network security requires a broader approach that extends beyond authentication improvements alone.

Comparison with Previous Research

The results obtained in this study are consistent with findings reported in previous research examining how wireless networks behave under attack conditions. Earlier

studies have shown that attackers frequently use deauthentication techniques to disrupt wireless communication by exploiting weaknesses in the way wireless networks manage device connections and exchange management frames (Wang et al., 2022). These attacks remain one of the most common methods used to interfere with wireless network availability.

Large-scale traffic analysis studies, such as the AWID3 dataset, verify that management frame-based attacks still exist in modern wireless environments. The research methodology adopted in this study is similar to those presently being adopted in wireless network security research, using a testbed environment to study how wireless networks behave under attack (Carbajal et al., 2022). The findings of this study provide additional evidence of how availability-based attacks can impact the stability of wireless networks even when security protocols like WPA3 are being adopted.

Future Directions for Wireless Network Security

The results obtained from the study demonstrate that the development of the resilience of wireless communication systems against availability-based attacks is an important area for further research. One promising area is the development of sophisticated wireless monitoring systems able to identify abnormal management frame traffic in real-time. The application of traffic analysis for detecting abnormal wireless traffic that could indicate attack activities has already been identified as having considerable potential (Suryadi et al., 2024).

Another important area for further research is the development of wireless intrusion detection systems that take into account the application of modern security protocols such as WPA3. The application of machine learning could provide the solution for the accurate identification of complex attack patterns in wireless network environments (Wang et al., 2023).

Moreover, the development of wireless resilience against other sophisticated attack mechanisms that compromise wireless communication protocols, such as side-channel attacks and packet-level manipulation attacks, has been identified in emerging wireless security research (Wang et al., 2022). Therefore, the development of wireless security mechanisms in the future must consider the integration of stronger authentication protocols, sophisticated monitoring mechanisms, and advanced detection mechanisms.

Conclusion

This study aimed to evaluate the behavior of modern wireless networks under deauthentication attack scenarios and evaluate the impacts of these attacks on the availability of wireless networks within a network infrastructure that uses modern security standards. The experiments conducted have shown that deauthentication attacks can severely affect the stability of wireless communication by disrupting the association between wireless devices and wireless access points.

Although the WPA3 security protocol has shown significant improvements in wireless network authentication security using Simultaneous Authentication of Equals (SAE), the experiments have shown that these improvements do not entirely

remove the impacts of attacks on the behavior of wireless network management frames. This shows a severe limitation of modern wireless security standards, which mostly attempt to strengthen wireless network authentication security.

The contribution of this work to the scientific community lies in the experimental assessment of the resilience of the wireless network under deauthentication attack scenarios, conducted in controlled environments using traffic analysis. This work adds further weight to the fact that attacks involving the use of management frames still need to be considered in modern wireless networks.

This work further stresses the need to ensure that future wireless security protocols not only focus on improving existing authentication protocols but also include effective detection mechanisms. Improving the security of management frames, improving the capabilities of existing wireless intrusion detection systems, and improving traffic analysis techniques can greatly improve the resilience of modern wireless networks against availability attacks.

Future research will thus be focused on the development of stronger wireless monitoring systems that can effectively identify any abnormalities in the management frames in real time, as well as the improvement of the design of wireless security protocols to effectively combat threats to network availability. Continued investigation of wireless attack scenarios in controlled experimental environments is necessary to evaluate the effectiveness of newly developed wireless security mechanisms.

Generally, the current research seeks to provide a deeper understanding of the behavior of modern wireless networks when subjected to deauthentication attack scenarios, as well as the challenges that must be overcome to effectively advance the security of future wireless communication systems.

References

- W. Stallings (2018) *Wireless Communications and Networks*. Upper Saddle River, NJ, USA: Pearson Education. (in English)
- A. Mishra and W.A. Arbaugh (2002) "An Initial Security Analysis of the IEEE 802.11 Wireless Network Protocol," Univ. Maryland, College Park, MD, USA, Tech. Rep. (in English)
- M. Vanhoef and F. Piessens (2017) "Key Reinstallation Attacks: Forcing Nonce Reuse in WPA2," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*. — P. 1313–1328, doi: 10.1145/3133956.3134027. (in English)
- M. Vanhoef and E. Ronen (2020) "Dragonblood: A Security Analysis of WPA3's SAE Handshake," in *Proc. IEEE Symp. Security and Privacy*. — P. 517–533, doi: 10.1109/SP40000.2020.00037 (in English)
- M. Vanhoef (2021) "Fragment and Forge: Breaking Wi-Fi Through Frame Aggregation and Fragmentation," in *Proc. USENIX Security Symp.* — P. 183–199. (in English)
- R. Gonçalves, M. E. Correia, and P. Brandão (2018) "A Flexible Framework for Rogue Access Point Detection," in *Proc. Int. Conf. Security and Cryptography (SECRYPT)*. — P. 1–8. (in English)
- G. Bianchi (2000) "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3. — P. 535–547, Mar. doi: 10.1109/49.840210. (in English)
- S. Hamad et al. (2021) "Machine Learning-Based Intrusion Detection for Wireless Networks," *IEEE Access*, vol. 9. — P. 123456–123468. (in English)

M. Vanhoef and C. Matte (2020) “Dragonblood: Analyzing the Dragonfly Handshake of WPA3,” *IEEE Security Privacy*, vol. 18, no. 3. — P. 51–57, May/Jun. doi: 10.1109/MSEC.2020.2979196. (in English)

A. Ali, S. Khan, and I. Ahmad (2022) “Security Analysis of Modern Wi-Fi Networks: Attacks and Countermeasures,” *IEEE Access*, vol. 10. — P. 65432–65450. doi: 10.1109/ACCESS.2022.3176543. (in English)

Z. Wang, X. Feng, Q. Li, K. Sun, and Y. Yang (2023) “Real-Time Detection of Wi-Fi Attacks Using Hybrid Deep Learning Models,” *IEEE Access*, vol. 11. — P. 11234–11248. (in English)

A. Carbajal and A. Akbarfam (2023) “Evaluating the Effectiveness of WPA3 Protocol Against Advanced Hacking Attacks,” in *Proc. Int. Conf. Cyber Security*. — P. 45–52. (in English)

D. Schepers, A. Ranganathan, and M. Vanhoef (2021) “On the Robustness of Wi-Fi Deauthentication Countermeasures,” in *Proc. ACM Conf. Security Privacy Wireless Mobile Netw. (WiSec)*. — P. 101–112, doi: 10.1145/3507657.3528540. (in English)

E. Chatzoglou, G. Kambourakis, and C. Kolias (2021) “How Is Your Wi-Fi Connection Today? DoS Attacks on WPA3-SAE,” *J. Inf. Security Appl.*, vol. 63, Art. no. 103020. doi: 10.1016/j.jisa.2021.103020. (in English)

E. Chatzoglou, G. Kambourakis, and C. Kolias (2021) “Empirical Evaluation of Attacks Against IEEE 802.11 Enterprise Networks: The AWID3 Dataset,” *IEEE Access*, vol. 9. — P. 34188–34205. doi: 10.1109/ACCESS.2021.3061211. (in English)

M. Natkaniec and M. Bednarz (2023) “Wireless Local Area Networks Threat Detection Using 1D-CNN,” *Sensors*, vol. 23, no. 4, Art. no. 2145. doi: 10.3390/s23042145. (in English)

M.T. Suryadi et al. (2024) “Machine Learning-Based Policing Models for Cyber-Attack Detection in Wi-Fi Networks,” *Electronics*, vol. 13, no. 2, Art. no. 415. doi: 10.3390/electronics13020415. (in English)

A. Carbajal and A. Akbarfam (2023) “A Software-Defined Testbed for Quantifying Deauthentication Resilience in Modern Wi-Fi Networks,” in *Proc. IEEE Int. Conf. Commun. (ICC)*. — P. 1–6. (in English)

N. Dalal et al. (2023) “A Wireless Intrusion Detection System for 802.11 WPA3 Networks,” in *Proc. IEEE Int. Conf. Cyber Security and Resilience*. — P. 78–85. (in English)

K. Chadee, W. Goodridge, and K. Khan (2022) “Recovering WPA3 Network Password by Bypassing the Simultaneous Authentication of Equals Handshake Using a Social Engineering Captive Portal,” in *Proc. Int. Conf. Cyber Warfare Security*. — P. 55–63. (in English)

M.H. Moharam et al. (2024) “Real-Time Detection of Wi-Fi Attacks Using Hybrid Deep Learning Models on NodeMCU,” *Sensors*, vol. 24, no. 1, Art. no. 145. doi: 10.3390/s24010145. (in English)

Z. Wang et al. (2022) “Off-Path TCP Hijacking in Wi-Fi Networks: A Packet-Size Side-Channel Attack,” in *Proc. USENIX Security Symp.* — P. 987–1004. (in English)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 391–413

<https://doi.org/10.32014/2026.2518-1726.445>

IRSTI 28.17.23

UDC 629.7.052:004.056

© **Tursunbayeva G.U.***, **Satybaldina D.Zh.**, **Tleuberdin S.T.**,
Tashatov N.N., **Egamberdiyev E.E.**, 2026.

Research Institute of Information Security and Cryptology, L.N. Gumilyov

Eurasian National University, Astana, Kazakhstan.

E-mail: t.gulzhamal@outlook.com

ANOMALY DETECTION IN UAV TELEMETRY SYSTEMS BASED ON SIMULATION MODELING

Tursunbayeva Gulzhamal — PhD Student in Information Security System, Junior Researcher at the Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: t.gulzhamal@outlook.com, <https://orcid.org/0000-0002-2044-8027>;

Satybaldina Dina — Candidate of Physical and Mathematical Sciences, Professor, Head of Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: satybaldina_dzh@enu.kz, <https://orcid.org/0000-0003-0291-4685>;

Tleuberdin Saken — PhD Student in Information Security System, Junior Researcher at the Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: sakentleuberdin@gmail.com, <https://orcid.org/0000-0001-9170-9936>;

Tashatov Nurlan — Candidate of Physical and Mathematical Sciences, Researcher of Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: tashatov_nn@enu.kz, <https://orcid.org/0000-0002-3271-2163>;

Egamberdiyev Eldor — PhD, Researcher of Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: eeldoru@gmail.com, <https://orcid.org/0000-0001-5289-6580>.

Abstract. Currently, the rapid development and widespread implementation of unmanned technologies have led to significant changes in various fields of human activity. At the same time, the risks associated with the unauthorized use of aircraft systems have increased. This has resulted in the emergence of a separate research area focused on the protection and counteraction of various components and platforms of aerial systems. Despite the existence of modern attack and anomaly detection methods, their effectiveness is significantly reduced under complex operating scenarios, including dynamically changing environments, interference

effects, small target sizes, and low radar visibility. To address this issue, this study presents the main results of a comprehensive analysis of modern cyber threats and vulnerabilities arising in unmanned aerial vehicle (UAV) systems. Based on the analysis, an up-to-date classification of existing attack types targeting the basic UAV architecture was developed. This made it possible to investigate the main protection methods for ensuring the security of UAV systems and components, as well as to classify cyberattack detection methods for these systems. Based on the obtained data, a multi-level protection architecture was developed, including three main levels: a secure communication channel, a secure flight controller, and a secure ground control station. The developed software environment for telemetry stream simulation in Python 3.12 enabled the generation of MAVLink/UDP/TCP packets, as well as the emulation of attacks and detection of network anomalies in UAV telemetry systems. The obtained results include the processing of 97 MAVLink packets, where the anomaly injection rate amounted to 10% with a total of 118 anomalies detected. The average MAVLink packet delay was 0.037 seconds, indicating stable operation of the telemetry channel. Experimental verification consisting of 100 cycles demonstrated the ability to detect packet structure violations, false identifiers, coordinate substitution, and delay anomalies.

Keywords: unmanned aerial vehicles (UAVs); telemetry channel; UAV cybersecurity; anomaly detection; MAVLink; attacks; vulnerability

For citations: Tursunbayeva G.U., Satybaldina D.Zh., Tleuberdin S.T., Tashatov N.N., Egamberdiyev E.E. Anomaly detection in uav telemetry systems based on simulation modeling. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 391-413. DOI <https://doi.org/10.32014/2026.2518-1726.445>*

© Турсунбаева Г.*, Сатыбалдина Д., Тлеубердин С., Ташатов Н.,
Эгамбердиев Э., 2026.

«Ақпараттық қауіпсіздік және криптология» ғылыми-зерттеу институты,
Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан.
E-mail: t.gulzhamal@outlook.com

СИМУЛЯЦИЯЛЫҚ МОДЕЛЬДЕУ НЕГІЗІНДЕ ҰШҚЫШСЫЗ ҰШУ АППАРАТТАРЫНЫҢ ТЕЛЕМЕТРИЯЛЫҚ ЖҮЙЕЛЕРІНДЕГІ АНОМАЛИЯЛАРДЫ АНЫҚТАУ

Турсунбаева Гулжамал — Ақпараттық қауіпсіздік жүйелері білім беру бағдарламасының докторанты, «Ақпараттық қауіпсіздік және криптология» ғылыми-зерттеу институтының кіші ғылыми қызметкері, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: t.gulzhamal@outlook.com, <https://orcid.org/0000-0002-2044-8027>;

Сатыбалдина Дина — физика-математика ғылымдарының кандидаты, профессор, «Ақпараттық қауіпсіздік және криптология» ғылыми-зерттеу институтының директоры, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: satybaldina_dzh@enu.kz, <https://orcid.org/0000-0003-0291-4685>;

Тлеубердин Сакен — Ақпараттық қауіпсіздік жүйелері білім беру бағдарламасының докторанты, «Ақпараттық қауіпсіздік және криптология» ғылыми-зерттеу институтының кіші ғылыми қызметкері, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: sakentleuberдин@gmail.com, <https://orcid.org/0000-0001-9170-9936>;

Ташатов Нурлан — физика-математика ғылымдарының кандидаты, доцент, «Ақпараттық қауіпсіздік және криптология» ғылыми-зерттеу институтының ғылыми қызметкері, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: tashatov_nn@enu.kz, <https://orcid.org/0000-0002-3271-2163>;

Эгамбердиев Эльдор — PhD, «Ақпараттық қауіпсіздік және криптология» ғылыми-зерттеу институтының ғылыми қызметкері, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: eeldoru@gmail.com, <https://orcid.org/0000-0001-5289-6580>.

Аннотация. Қазіргі уақытта ұшқышсыз технологиялардың қарқынды дамуы мен кеңінен енгізілуі адам қызметінің әртүрлі салаларында елеулі өзгерістерге алып келді. Сонымен қатар ұшу аппараттары жүйелерін рұқсатсыз пайдалануға байланысты тәуекелдер де артты. Бұл ұшу жүйелерінің әртүрлі компоненттері мен платформаларын қорғау және қарсы әрекет ету саласында жеке ғылыми бағыттың қалыптасуына себеп болды. Қолданыстағы шабуылдар мен аномалияларды анықтау әдістерінің болуына қарамастан, олардың тиімділігі динамикалық өзгеретін орта, кедергілердің әсері, нысандардың шағын өлшемдері және төмен радиолокациялық байқалуы сияқты күрделі пайдалану сценарийлерінде айтарлықтай төмендейді. Осы мәселені шешу мақсатында зерттеуде ұшқышсыз ұшу аппараттары (ҰҰА) жүйелерінде туындайтын заманауи киберқауіптер мен осалдықтарды кешенді талдау нәтижелері ұсынылған. Талдау негізінде ҰҰА базалық архитектурасына бағытталған шабуылдардың өзекті жіктелімі жасалды. Бұл ҰҰА жүйелері мен компоненттерінің қауіпсіздігін қамтамасыз ету әдістерін зерттеуге, сондай-ақ олардың жүйелеріне бағытталған кибершабуылдарды анықтау тәсілдерін жіктеуге мүмкіндік берді. Алынған мәліметтер негізінде қорғалған байланыс арнасын, қорғалған ұшу контроллерін және қорғалған жерүсті басқару станциясын қамтитын көпдеңгейлі қорғаныс архитектурасы әзірленді. Python 3.12 ортасында жасалған телеметриялық ағындарды модельдеу бағдарламалық ортасы MAVLink/UDP/TCP форматындағы пакеттерді генерациялауға, шабуылдарды имитациялауға және ҰҰА телеметрия жүйесіндегі желілік аномалияларды тіркеуге мүмкіндік берді. Алынған нәтижелерге сәйкес 97 MAVLink пакеті өңделді, мұнда аномалия инъекцияларының үлесі 10%-ды, ал олардың жалпы саны 118 бірлікті құрады. MAVLink пакетінің орташа кідіріс уақыты 0,037 секундты құрады, бұл телеметриялық арнаның тұрақты жұмысын көрсетеді. 100 циклден тұратын эксперименттік верификация деректер пакеттері құрылымының бұзылуын, жалған идентификаторларды, координаталарды алмастыруды және кідіріс аномалияларын анықтау мүмкіндігін көрсетті.

Түйін сөздер: ұшқышсыз ұшу аппараттары (ҰҰА); телеметрия арнасы;

ҰҰА киберқорғанысы; аномалияларды анықтау; MAVLink; шабуылдар; осалдық

©Турсунбаева Г.*, Сатыбалдина Д., Тлеубердин С., Ташатов Н.,
Эгамбердиев Э., 2026.

НИИ информационной безопасности и криптологии Евразийского
национального университета имени Л.Н. Гумилева, Астана, Казахстан.

E-mail: t.gulzhamal@outlook.com

ОБНАРУЖЕНИЕ АНОМАЛИЙ В ТЕЛЕМЕТРИЧЕСКИХ СИСТЕМАХ БПЛА НА ОСНОВЕ СИММУЛЯЦИОННОГО МОДЕЛИРОВАНИЯ

Турсунбаева Гулжамал — докторант образовательной программы «Системы информационной безопасности», младший научный сотрудник НИИ информационной безопасности и криптологии, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: t.gulzhamal@outlook.com, <https://orcid.org/0000-0002-2044-8027>;

Сатыбалдина Дина — кандидат физико-математических наук, профессор, директор НИИ информационной безопасности и криптологии, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: satybalдина_dzh@enu.kz, <https://orcid.org/0000-0003-0291-4685>;

Тлеубердин Сакен — докторант образовательной программы «Системы информационной безопасности», младший научный сотрудник НИИ информационной безопасности и криптологии, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: sakentleuberдин@gmail.com, <https://orcid.org/0000-0001-9170-9936>;

Ташатов Нурлан — кандидат физико-математических наук, доцент, научный сотрудник НИИ информационной безопасности и криптологии, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: tashatov_nn@enu.kz, <https://orcid.org/0000-0002-3271-2163>;

Эгамбердиев Эльдор — PhD, научный сотрудник НИИ информационной безопасности и криптологии, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,

E-mail: eeldoru@gmail.com, <https://orcid.org/0000-0001-5289-6580>.

Аннотация. *Актуальность.* Стремительное развитие и широкое внедрение беспилотных технологий привели к значительным изменениям в различных сферах деятельности человека. Одновременно возросли риски, связанные с несанкционированным использованием беспилотных летательных аппаратов и нарушением работы их телеметрических систем. Это обусловило формирование отдельного направления исследований, связанного с противодействием киберугрозам и защитой компонентов, каналов связи и платформ БПЛА. Несмотря на наличие существующих методов обнаружения атак и аномалий, их эффективность может снижаться в условиях сложных сценариев эксплуатации, включая динамически изменяющуюся среду, помеховое воздействие, малые размеры целей и низкую радиолокационную заметность. *Цель.* Разработать и апробировать

подход к обнаружению аномалий в телеметрических системах БПЛА на основе симуляционного моделирования телеметрических потоков, имитации атак и анализа сетевых аномалий. *Методы.* В исследовании проведен комплексный анализ современных киберугроз и уязвимостей, возникающих в системах беспилотных летательных аппаратов. На основе анализа составлена классификация существующих типов атак на базовую архитектуру БПЛА, что позволило определить основные методы защиты систем и компонентов БПЛА, а также классифицировать подходы к обнаружению кибератак. На основе полученных данных разработана многоуровневая архитектура защиты, включающая три основных уровня: защищенный канал связи, защищенный контроллер полета и защищенную наземную контрольную станцию. Для экспериментальной проверки создана программная среда моделирования телеметрических потоков на Python 3.12, позволяющая генерировать пакеты в формате MAVLink/UDP/TCP, имитировать атаки и фиксировать сетевые аномалии в системе телеметрии БПЛА. *Результаты и выводы.* Полученные результаты включают обработку 97 пакетов MAVLink, при этом доля инъекций аномалий составила 10%, а общее количество зафиксированных событий достигло 118 единиц. Среднее время задержки пакета MAVLink составило 0,037 секунды, что свидетельствует о стабильной работе телеметрического канала в условиях моделирования. Экспериментальная верификация, состоящая из 100 циклов, продемонстрировала способность предложенного подхода обнаруживать нарушения структуры пакетов данных, ложные идентификаторы, подстановку координат и аномалии задержек. Практическая значимость исследования заключается в возможности использования разработанной симуляционной среды и многоуровневой архитектуры защиты при проектировании систем кибербезопасности БПЛА, мониторинге телеметрических каналов и разработке средств обнаружения атак на беспилотные платформы.

Ключевые слова: беспилотные летательные аппараты, БПЛА, канал телеметрии, киберзащита БПЛА, обнаружение аномалий, MAVLink, UDP, TCP, атаки, уязвимости, симуляционное моделирование.

Финансирование: Данное исследование было профинансировано Комитетом науки Министерства науки и высшего образования Республики Казахстан в рамках реализации проекта ГФ (ИРН: AP26101065).

Введение. На сегодняшний день широкое применение беспилотных летательных аппаратов (БПЛА) для коммерческих, гражданских и военных целях значительно повысило потребность безопасности их систем, механизмов и компонентов (Islam et al., 2025). Учитывая, что они все чаще используются в выполнении самых разнообразных задачах начиная от инспекции инфраструктуры, аэрофотосъемки, мониторинг окружающей среды, дистанционного зондирования Земли (ДЗЗ) и до доставки посылок в

труднодоступные места делают их крайне уязвимыми для кибервторжений. Эти угрозы могут нарушить выполнение различных миссий и поставить под угрозу целостность данных, включая причинение физического вреда БПЛА (Alsumayt et al., 2026). По мере расширения масштабов и усложнения задач и миссий в различных секторах, растущая зависимость от цифровых платформ и систем подвергает их воздействию широкого спектра угроз кибербезопасности.

Операционная целостность платформы БПЛА напрямую зависит от безопасности базовых подсистем, таких как телеметрическая связь, каналы управления и контроля, навигация на основе GNSS/GPS, механизмы удалённой идентификации, облачные сервисы мониторинга и хранения данных (Tlili et al., 2024).

Структурная схема базовой архитектуры БПЛА состоит из трех основных компонентов:

1. Бортовой сегмент БПЛА обеспечивает автономное функционирование БПЛА в воздушном пространстве. В данном блоке реализация алгоритмов автоматического управления непосредственно выполняются в подсистеме Flight Controller, а для определения координат, скорости и высоты воздушного пространства применяется подсистема Navigation System. Передача телеметрических данных и приём управляющих команд по заданному каналу обеспечивается через подсистему Communication Module. Для выполнения различных миссий БПЛА может в составе иметь Payload. Питание всех бортовых компонентов БПЛА и поддержание их стабильной работы в режиме реального времени обеспечивается через подсистему Power System.

2. Наземная система управления БПЛА обеспечивает планирование заданной миссии, мониторинг и контроль параметров телеметрии, включая командное управление БПЛА.

3. Облачная платформа БПЛА предназначена для обработки и хранения данных.

Связь между БПЛА и наземной станцией осуществляется посредством радиоканала или сотовой сети с использованием протоколов передачи данных. Обмен данными между наземной системой управления БПЛА и облачной платформой осуществляется через сеть Интернет.

Нарушение функциональности любого из них может привести к потере устойчивости управления полетом БПЛА в воздушном пространстве, снижению точности навигационных решений и полной потере связи с наземной станцией что в конечном итоге повышает вероятность возникновения аварийных ситуаций и создаёт риск невыполнения поставленной миссии.

По своей природе, БПЛА являются более уязвимы к атакам, чем пилотируемые летательные аппараты. Причиной тому является отсутствие вмешательства человека, включая сильную зависимость от беспроводной связи и наличие не защищенной аппаратной архитектуры. В настоящее время наиболее распространены атаки, осуществляемые с четырёх сторон: датчики

БПЛА, связь, программное обеспечение и кибербезопасность. Типовые кибератаки на БПЛА охватывают каналы связи, навигационную подсистему, полётный контроллер, полезную нагрузку, энергетическую систему и наземно-облачную инфраструктуру, формируя многоуровневую поверхность атак, что повышает актуальность разработки подходящих решений для обеспечения безопасности БПЛА.

Целью исследования является формирование концептуальной модели многоуровневой системы киберзащиты БПЛА и её программная реализация в виде симуляционной платформы, предназначенной для моделирования телеметрического обмена, инъекции аномалий и оценки эффективности детектирования нарушений целостности, аутентичности и непрерывности передачи данных без использования методов машинного обучения.

Обзор литературы. Разработка подходящих оптимальных решений для обеспечения безопасности БПЛА затруднена из-за множества факторов (Alshamrani et al., 2026). Самым распространённым из них можно отметить крайне ограниченную вычислительную мощность на борту БПЛА (Mohammed et al., 2025). Это ограничивает возможности внедрения программного обеспечения и их компонентов для обеспечения безопасности и целостности передаваемых данных. Более того наземные станции управления БПЛА часто представляют собой простых устройств дистанционного управления, которые имеют ограничения в установке дополнительных программных функций безопасности (Burbank et al., 2026). Проблемы, связанные с ограничениями по размеру, весу БПЛА делают многие аппаратные решения для обеспечения безопасности непрактичными. В том числе многие механизмы безопасности, разработанные для традиционных компьютерных систем и сетей передачи данных, не могут быть напрямую применены к БПЛА (Sharifi et al., 2026). Еще одним самым распространяющим фактором нарушения безопасности может быть физический захват БПЛА из-за повышенной заметности БПЛА в воздушном пространстве (Yoo et al., 2025). Эти факторы дают злоумышленникам больше возможностей для взаимодействия с БПЛА для создания помех, подмены сигналов и перехвата данных нарушая связь управления и контроля, потенциально приводя к приостановке полета или непредсказуемому изменению траектории полета БПЛА.

В данном направлении множество исследований была посвящена классификации уязвимостей БПЛА, типам атак (Tang et al., 2025), включая разработки методов защиты (Romagnoli et al., 2023). При этом существующие подходы имеют разные точки зрения с учетом архитектуры системы БПЛА. Большинство из них демонстрируют основных векторов атак, направленные на полетный контроллер, наземную станцию управления и канал передачи данных, включая GPS-спуфинг, глушение сигналов, атаки типа DoS, подмену видеопотока и внедрение вредоносного кода. Множество работ за последние десятилетия были посвящены выявлению потенциальных слабых мест и классификации типов атак на основе уязвимостей архитектуры системы

БПЛА. В основном киберугрозы БПЛА классифицировались по целевым компонентам системы (полетный контроллер, наземная станция управления и канал передачи данных) и по возможностям злоумышленника (раскрытие данных, получение знаний о системе и нарушение работы) (Qiu et al., 2022). Множество уязвимых мест для атак со стороны злоумышленников в значительной степени связаны с беспроводными протоколами связи, бортовых датчиков и автоматизированными контурами управления (Yang et al., 2026).

Во многих исследованиях были рассмотрены современные виды атак на БПЛА и среди них самые распространённые виды были отмечены подмена GPS-сигнала, внедрение команд, подавление каналов передачи данных или отказ в обслуживании (DoS) (Tahavori et al., 2020). В последствии несвоевременного обнаружения и предотвращения атак на компоненты БПЛА могут привести к срыву миссий, нарушению конфиденциальности или даже нанесение физического вреда. При этом авторы (Zuev et al., 2018) отмечают о значительном повышении риска, в случаях, когда БПЛА работают автономно или взаимодействуют с наземными или облачными системами управления. В работе (Wang et al 2026), авторы утверждают, что атаки в виде глушения или подмены GPS-сигналов осуществляются для отключения системы навигации БПЛА. При этом на сегодняшний день GPS-сигналы в системах БПЛА часто не подлежат шифрованию, что существенно упрощает реализацию атаки и в последствии подмена сигнала ложными хаотичными координатами могут вызвать дезориентацию и падение БПЛА. Во многих случаях нарушение вещания позиций (аналог ADS-B) происходит из-за глушения сигналов управления GCS (Rezaee et al 2024). Нарушение связи приводит к активации режимов отказа, вызывая хаотичный полет БПЛА и может вызвать процесс столкновения, что приводит к аварийному исходу. Подмена сигналов в системе телеметрии и видеонаблюдения может ввести в заблуждение оператора, который управляет БПЛА дистанционно. Также существуют атаки в виде манипуляции захваченным видео. В данном случае перехват системных данных для замены реального видео на фальсифицированное позволяет злоумышленнику захватить контроль. Отмечается (Wang et al 2026) что данный способ атаки можно реализовать в комбинации с подменой GPS координат. Внедрение фальсифицированных данных с сенсоров БПЛА осуществляется с помощью внешнего вмешательства или внутренний доступ в целях дестабилизации режима полета. Злонамеренные аппаратные и вредоносные программные обеспечения (трояны, бэкдоры, командные инъекции и т.д.) могут быть внедрены для кражи конфиденциальных данных или же для намеренного отключения системы защиты чтобы вызвать сбой в системе управления БПЛА. Также отмечается важность своевременного предотвращения несанкционированного раскрытия коммуникаций системы БПЛА. Данный вид атак реализуется с целью несанкционированного получения и утечки персональных данных и может привести к компрометации

программного обеспечения диспетчеризации, формированию некорректных заданий, нарушению маршрутизации, операционным сбоям. Одним из наиболее распространённых типов атак на БПЛА считается атака типа «Отказ в обслуживании» (DoS). В большинстве случаев незапланированная внезапная перегрузка вычислительных ресурсов или сетевых каналов БПЛА приводит к остановке операций, снижению производительности либо переводу системы в непредусмотренное состояние функционирования.

Общий анализ существующих исследований в данном направлении сводится к тому, что кибератаки на БПЛА могут охватывать все компоненты и подсистемы базовой архитектуры БПЛА и классифицируются следующим образом, представленной на рисунке 1.

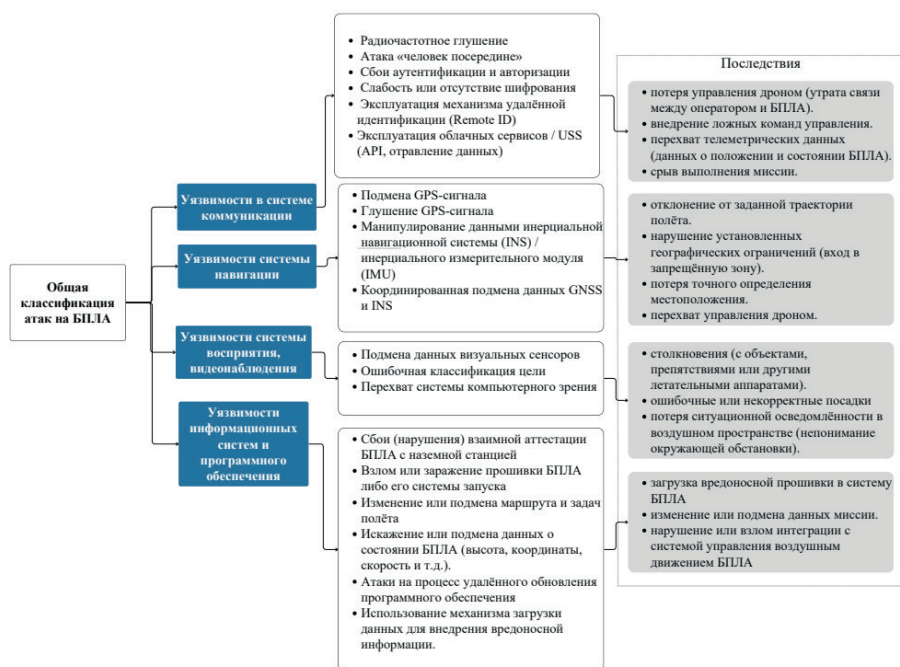


Рисунок 1 – Основная классификация существующих видов атак на базовую архитектуру БПЛА

За последние годы, рост использования БПЛА в значительной степени повышает их киберустойчивость. Отсутствие надёжной защиты делает их более уязвимыми для эксплуатации уязвимостей в программном обеспечении, внедрения вредоносных программ и несанкционированных модификаций программного обеспечения. Формирование надёжной системы защиты и безопасность их компонентов становится необходимым для обеспечения его функциональности при выполнении в различных миссиях.

В данном направлении были проанализированы различные виды угроз безопасности на БПЛА. Демонстрируя всесторонний анализ существующих

атак исследователями (Qiu et al., 2022), была предложена модель киберугроз и были обсуждены методы повышения безопасности компонентов БПЛА. Также схожие работы были посвящены разработкам усовершенствованных методов шифрования канала связи, защиты от атак типа «отказ в обслуживании» (DoS) и систем обнаружения вторжений. В некоторых из них были получены интересные выводы. Одни авторы утверждают, если своевременное обнаружение пассивного перехвата и внедрение надежных протоколов шифрования и защищенных каналов связи может способствовать значительному повышению уровня безопасности БПЛА. Другая группа авторов в своих исследованиях (Yang et al., 2026) основное внимание уделили построению эффективного метода оценки состояния на основе полученных данных от систем интеллектуальных датчиков для обнаружения атак на данные связи БПЛА. Это позволило предотвратить перехвата конфиденциальной информации осуществляемые через навигационные команды, системы телеметрии. Также для защиты критически важных сведений об архитектуре и работе сети БПЛА и для обеспечения безопасной и надежной интеграции БПЛА в воздушное пространство многими регулирующими органами различных стран были разработаны руководящие принципы кибербезопасности. Изучая установленные нормативы кибербезопасности БПЛА, авторы (Tucker et al., 2026) пришли к тому, что в большей степени регламенты Federal Aviation Administration (FAA) предназначено для защиты каналов связи БПЛА от подмены, помех и несанкционированного доступа. В этом направлении аналогичный регламент European Union Aviation Safety Agency (EASA) описывает меры кибербезопасности и защиты данных для операций с БПЛА. Для защиты сетей беспилотных летательных аппаратов (БПЛА) от киберугроз криптографические методы, обеспечивающие конфиденциальность, целостность и аутентичность передаваемых данных, включая управление идентификацией на основе блокчейна. В дополнение к регламентам отмечается важность соответствия методом защиты к стандартам кибербезопасности ISO/IEC 27001.

На сегодняшний день были разработаны и проанализированы множество способов и методов, направленные на борьбу с атаками на БПЛА. Анализируя существующие передовые отраслевые практики формирования устойчивой системы кибербезопасности БПЛА применяются различные методы, которые классифицируются следующим образом, согласно таблице 1.

Таблица 1 – Классификация методов обнаружения кибератак в БПЛА

№	Метод обнаружения	Принцип работы	Примеры применения	Преимущества	Ограничения
1	Signature-Based	Сравнение трафика или поведения с известными шаблонами атак	IDS по правилам, проверка MAVLink-команд	Высокая точность для известных атак, низкая вычислительная нагрузка	Не выявляет новые (zero-day) атаки
2	Anomaly-Based	Обнаружение отклонений от нормального профиля работы	Анализ телеметрии, контроль задержек, частоты пакетов	Позволяет выявлять неизвестные атаки	Возможны ложные срабатывания
3	ML-Based	Классификация поведения на основе обученных моделей	Детекция GPS-спуфинга, DoS, инъекции телеметрии	Высокая адаптивность, выявление сложных атак	Требует обучающих данных и вычислительных ресурсов
4	Cryptographic verification	Проверка целостности и аутентичности данных	Цифровые подписи, HMAC, secure boot	Высокая защита целостности и подлинности	Не выявляет поведенческие аномалии
5	Behavioral Detection	Сравнение фактического поведения с физической моделью полёта	Кросс-проверка GNSS и IMU, контроль траектории	Эффективен для навигационных атак	Требует точной модели движения

В сравнительном анализе современных методов обнаружения кибератак из Таблицы 1 были рассмотрены сигнатурные, аномальные, основанные на машинном обучении, криптографические и поведенческие методы обнаружения атак. Проведенный анализ показывает, что традиционные сигнатурные методы обеспечивают высокую точность обнаружения известных угроз, однако недостаточно эффективны против новых типов атак. В свою очередь, методы анализа аномалий и машинного обучения обладают более высокой адаптивностью и способны выявлять ранее неизвестные угрозы, но требуют значительных вычислительных ресурсов и качественных обучающих данных. Криптографические методы обеспечивают контроль целостности и подлинности данных, а поведенческий анализ позволяет выявлять навигационные атаки на основе сравнения фактических и ожидаемых параметров движения БПЛА.

Материалы и методы. Разностороннее изучение существующих

кибератак на БПЛА и эффективность применения существующих методов позволило разработать многоуровневую систему киберзащиты БПЛА, которая представлена на рисунке 2.

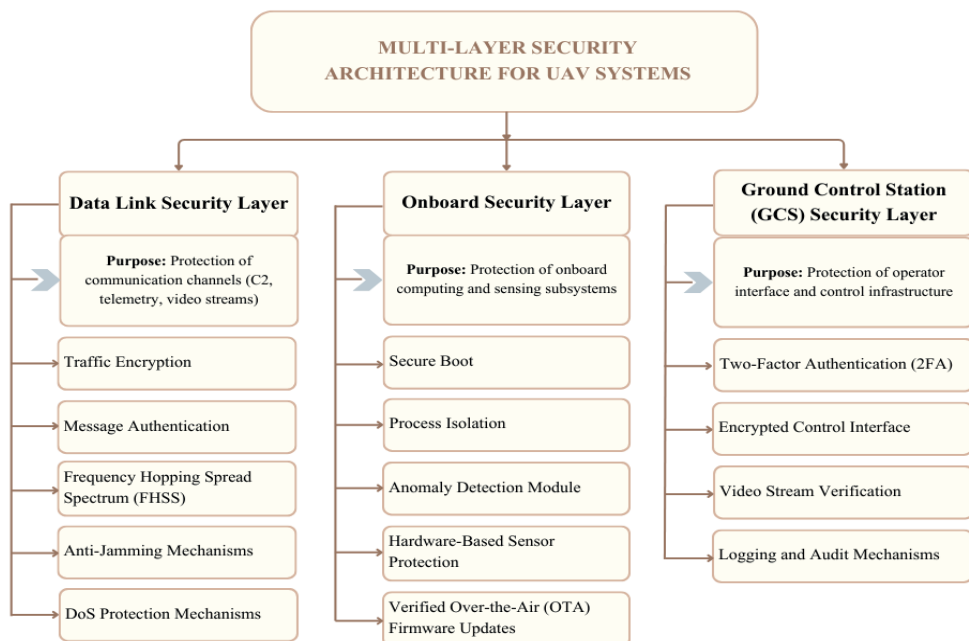


Рисунок 2 – Концептуальная архитектура системы киберзащиты БПЛА

В разработанной Концептуальной архитектуре системы киберзащиты БПЛА безопасность обеспечивается одновременно на уровне канала связи, бортовой вычислительной платформы и наземной станции управления. Представленный подход в отличие от существующих решений интегрирует защиту навигации, телеметрических потоков, сенсорного комплекса, исполнительных модулей и операторского интерфейса в единую непрерывную цепочку.

Для детализированного анализа эффективности предложенной многоуровневой архитектуры целесообразно рассмотреть её функциональные компоненты по отдельности. Среди них особое значение имеет телеметрический канал связи, поскольку именно он обеспечивает непрерывный обмен управляющими командами и параметрами полёта между бортовой системой и наземной станцией управления.

Телеметрический канал, обеспечивающий передачу управляющих команд и параметров полёта, является критическим компонентом БПЛА, уязвимым к инъекциям, подмене идентификаторов и нарушениям целостности пакетов. Структура телеметрической системы БПЛА представлена на рисунке 3.

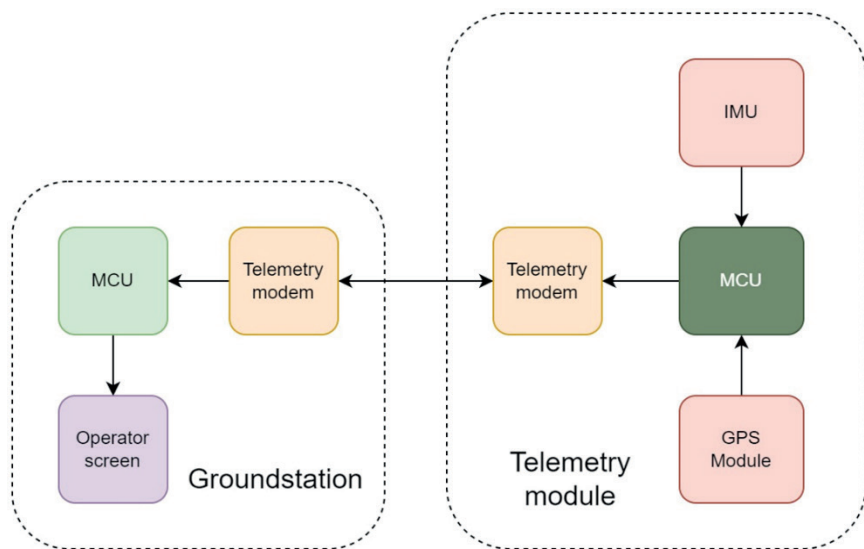


Рисунок 3 – Структура телеметрической системы БПЛА

На схеме представлена структурная организация телеметрического взаимодействия между наземной станцией управления и бортовым телеметрическим модулем БПЛА. Наземная станция включает микроконтроллер (MCU), операторский интерфейс и телеметрический модем, обеспечивающий двусторонний обмен данными. Управляющие команды от оператора передаются через MCU к радиомодему и далее по беспроводному каналу связи. Бортовой телеметрический модуль содержит собственный MCU, подключённый к навигационным и инерциальным датчикам (GPS Module, IMU). Телеметрический модем осуществляет приём управляющих команд и передачу параметров полёта (координаты, ориентация, скорость, состояние системы) обратно на наземную станцию. Анализ исследований выявил наиболее распространённых протоколов телеметрической связи, используемых БПЛА является MAVLink, User Datagram Protocol (UDP), Transmission Control Protocol (TCP) (Yu et al., 2025).

Протокол Micro Air Vehicle Link (MAVLink) в системах БПЛА применяется для обеспечения радиосвязи в пределах ограниченной зоны действия при средней скорости передачи данных в режиме прямой видимости (Dad et al., 2024). Особенностью MAVLink является низкая избыточность и малый размер пакетов, что делает его эффективным для радиоканалов с ограниченной пропускной способностью. Однако в базовой версии протокол не обеспечивает встроенного шифрования, что создаёт риски перехвата и инъекции сообщений. Протокола управления передачей TCP и UDP используется как транспортный протокол для передачи данных телеметрии в реальном времени. Они широкое применение нашли как в проводных, так и в беспроводных сетях. Применение TCP обеспечивает надежность,

гарантию доставки данных в определенном порядке и контроль перегрузки. Однако за счёт механизма подтверждений увеличивается время задержки, что ограничивает его применение в критичных к времени каналах управления. UDP в меньшей степени учитывает параметры надежности чем TCP, чтобы сосредоточиться на доставке пакетов данных в реальном времени. Это делает UDP предпочтительным для потоков телеметрии и видео, но одновременно повышает уязвимость к spoofing-атакам и packet injection (Yu et al., 2025).

На данном этапе исследования детальное изучение структуры сообщений, включая поля серийных номеров, логику проверки контрольной суммы и формальные определения полей данных предоставило возможность выявить три основных типа уязвимостей:

- отсутствие или слабость шифрования, замена пакетов;
- вызванная использованием предсказуемых идентификаторов;
- несогласованность в проверке контрольной суммы.

Эти результаты стали эмпирической основой для представления моделей угроз в телеметрическом модуле. Для более детального исследования была представлена структурно-функциональная схема программно-моделируемой системы анализа телеметрического обмена БПЛА.

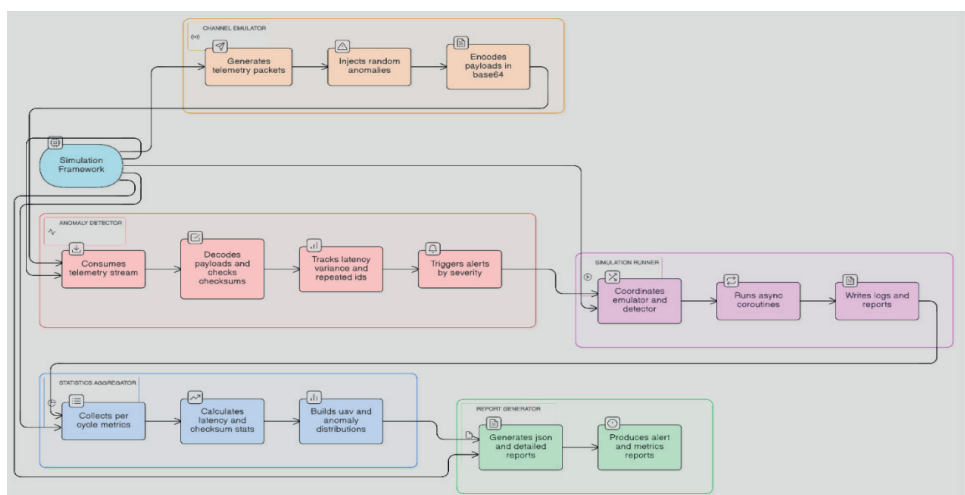


Рисунок 4 – Функциональная архитектура системы эмуляции и обнаружения аномалий телеметрического канала БПЛА

Функциональная архитектура системы эмуляции и обнаружения аномалий телеметрического канала БПЛА состоит из следующих модулей:

Для формирования телеметрических пакетов, инжектирования случайных аномалии и координации полезную нагрузку (base64) применяется Channel Emulator. Оно применяется для моделирования реальных условий передачи данных. Выполнение декодирования полезной нагрузки и проверку контрольных сумм применяется модуль Anomaly Detector. Оно осуществляет

анализ повторяющихся идентификаторов и вариаций задержки и инициирует оповещения в зависимости от уровня критичности. Сборка метрик по циклам обработки, вычисления статистики задержек и контрольных сумм, включая формирование распределения аномалий осуществляет модуль Statistics Aggregator. Координация взаимодействия существующих модулей осуществляется через Event Loop Runner. Report Generator формирует JSON-отчёты, журналы событий и сводные метрики системы.

В представленной функциональной архитектуре системы эмуляции и обнаружения аномалий телеметрического канала БПЛА механизм проверки контрольной суммы для пакетов MAVLink можно выразить следующим образом:

$$CRC_{valid} = \sum_{i=1}^n b_i \bmod 256 \quad (1)$$

где, CRC_{valid} – ожидаемое значение контрольной суммы, которая определяется вычислением из содержимого пакета MAVLink;

b_i – значение i -го байта в полезной нагрузке пакета телеметрии;

n – общее количество байт в полезной нагрузке пакета MAVLink;

$\bmod 256$ – операция модуля, гарантирующая, что контрольная сумма соответствует одному байту (0–255), которая соответствует структуре пакета MAVLink.

Для количественной оценки частоты нарушений в контрольной сумме коэффициент несоответствия определяется:

$$R_{CS} = \frac{N_{mismatch}}{N_{total}} \cdot 100\%, \quad (2)$$

где, $N_{mismatch}$ – количество пакетов, для которых вычисленная контрольная сумма не совпала с полученной контрольной суммой;

N_{total} – общее количество обработанных телеметрических пакетов в окне симуляции.

В условиях нормального режима функционирования типичное время доставки пакета от бортового модуля к наземной станции характеризуется средним значением задержки. Математическая модель задержек канала в данном случае определяется следующим образом:

$$\mu = \frac{1}{N} \sum_{i=1}^N d_i \quad (3)$$

где d_i – задержка передачи i – го телеметрического пакета;

N – общее количество обработанных пакетов.

Индикатор нестабильности телеметрического обмена характеризуется дисперсией задержки, которая определяется:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (d_i - \mu)^2 \quad (4)$$

Статистический критерий отклонения передачи пакетов от нормального режима оценивается критерием аномалий по задержке пакета и определяется следующим образом:

$$|d_i - \mu| > k\sigma \quad (5)$$

где k – порог чувствительности, обычно принимаются значения 2 (доверительный интервал ~95%) или 3 (доверительный интервал ~99.73%) и зависит от чувствительности детектора и вероятности возникновения ложных тревог;

Эффективность алгоритма детектирования характеризуется моделью оценки вероятности обнаружения атак:

$$P_D = \frac{N_{detected}}{N_{injected}} \quad (6)$$

где, $N_{detected}$ – количество искусственно внедрённых аномалий;

$N_{injected}$ – число корректно обнаруженных аномалий.

Вероятность правильного обнаружения атаки оценивается в пределах от 0 до 1.

$$0 \leq P_D \leq 1 \quad (7)$$

Вероятность ложного срабатывания определяется:

$$P_{FA} = \frac{N_{false}}{N_{normal}} \quad (8)$$

N_{normal} – число пакетов в нормальном режиме

N_{false} – число ошибочно классифицированных пакетов

Этот показатель оценивает вероятность того, что система объявляет аномалию при её отсутствии. Если значение чрезмерно высокий, то может привести к ненужным реакциям защиты и снижению доверия к системе.

Пороги для задержки, ошибок контрольной суммы и повторяющихся идентификаторов можно задать с применением параметров `latency_threshold`, `checksum_threshold`, `repeat_id_threshold`. Они позволяют более детально настраивать чувствительность обнаружения аномалий. В случае если пороги превышают установленную норму, то событие регистрируется в системе с следующих категориях уровней тяжести: средний, высокий и критический. Тогда оценка серьезности угрозы в системе многоуровневой модели будет выглядеть следующим образом:

$$S_i = \begin{cases} \text{critical} & \text{if } R_{CS} > \theta_{CS}^{high} \text{ AND } f_j > \theta_{id} \\ \text{high} & \text{if } R_{CS} > \theta_{CS}^{high} \text{ OR } |d_i - \mu| > 3\sigma_d \\ \text{medium} & \text{if } \text{payload}_{malformed} = \text{true} \text{ OR } |d_i - \mu| > 2\sigma_d \end{cases} \quad (9)$$

где S_i – классификация серьезности, присваиваемая i -му обнаруженному аномалии;

θ_{CS}^{high} – параметр порога несоответствия верхней контрольной суммы (`checksum_threshold`);

f_j – частота повторения идентификатора БПЛА в текущем окне обработки определяется соотношением числа появлений идентификатора в общее число пакетов N :

$$f_j = \frac{n_j}{N} \quad (10)$$

θ_{id} – параметр порога повторения идентификатора (`repeat_id_threshold`);

$\text{payload}_{malformed}$ – булев индикатор, равный моменту, когда полезная нагрузка пакета, декодированная в Base64.

Для верификации предложенной архитектурной модели была выполнена её практическая реализация в среде Python 3.12, в рамках которой разработана программная платформа симуляции телеметрического канала БПЛА.

Реализация осуществляется на совместной работе модулей генерации телеметрических пакетов, механизм инъекции аномалий, включая подсистему детектирования нарушений и блоков статистической обработки результатов.

Результаты. Обработка телеметрии рассматривается как непрерывный процесс с постепенным накоплением статистики, что позволяет фиксировать даже слабовыраженные признаки попыток атаки. Предъявленные требования и технические спецификации к безопасности и киберзащите основаны на программных возможностях, которые необходимы соблюдать в системе БПЛА для обнаружения основных форм помех или повреждений. В частности, система должна уметь выявлять отсутствующие участки, ошибочно сгенерированные полезные нагрузки и поддельные идентификаторы БПЛА.

Характер атак реализуемые в телеметрическом канале нарушают целостность, подлинность или непрерывность функциональности БПЛА что является критически опасным.

Разработанная архитектура системы обеспечивает реализацию модульной программной платформы, которая должна имитировать реалистичную телеметрическую среду для БПЛА. Структура защиты основана на трёх функциональных уровнях: профилактика; обнаружение; реагирование. Каждый модуль концептуально отличался друг от друга, но был разработан для взаимодействия в общей моделирующей среде. В этом цикле отчётности акцент делался на уровне обнаружения, и планировалось более глубоко интегрировать уровни профилактики и реагирования на следующих этапах.

Уровень предотвращения обеспечит криптографическую защиту и защищённые протоколы передачи, однако с точки зрения управления ключами он сложнее, и его реализация рассматривается после завершения разработки механизмов обнаружения аномалий. Уровень отклика позволяет выдавать предупреждения и вызывать системные реакции на основе подтверждения наличия телеметрических аномалий. На данном этапе этот модуль разработан в базовом виде с наличием элементарных функций. Этот процесс будет реализован в зависимости от результатов тестирования и возможностей обнаружения атак в более сложных сценариях. Рабочий процесс сформирован визуальным образом и представлено на рисунке 5.

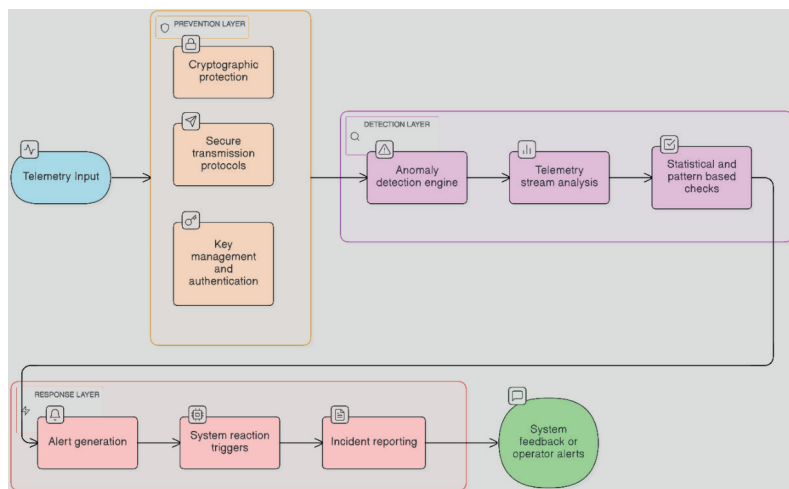


Рисунок 5 – Многоуровневая архитектура киберзащиты телеметрического канала БПЛА

Многоуровневая архитектура киберзащиты телеметрического канала БПЛА обеспечивает непрерывную цепочку защиты телеметрического обмена БПЛА.

Платформа симуляции в программной среде Python 3.12 состоит из выжних трех модулей ChannelEmulator, AnomalyDetector и SimulationRunner.

Эти модули разрабатывались в итерациях, каждый из которых независимо проходил модульное тестирование и верифицировался для обеспечения внутренней согласованности результатов и прослеживаемости.

Эмулятор канала является источником телеметрических данных. Он генерирует синтетические пакеты с реалистичными идентификаторами БПЛА, временными метками и другими параметрами полёта, такими как высота, скорость, курс и уровень заряда батареи. Зашифрованный канал связи и полезные нагрузки кодируются в блоке Base64. Система также обладает контролируемой инъекцией аномалий: случайно введённые ошибки включают потерю пакетов, искажённые полезные нагрузки и поддельные идентификаторы БПЛА. Это позволяет наблюдать эффективность системы в повторяемых, полуслучайных условиях.

Другой важный компонент, AnomalyDetector, обеспечивает функциональность основного анализа. Он потребляет поток телеметрии в реальном времени, декодирует каждую полезную нагрузку и проверяет встроенные контрольные суммы. Также он рассчитывает и отслеживает статистические показатели, такие как дисперсия задержки и повторяющиеся идентификаторы БПЛА, чтобы выявлять необычные закономерности.

Пороги для задержки, ошибок контрольной суммы и повторяющихся идентификаторов можно задать с помощью параметров: `latency_threshold`, `checksum_threshold`, `repeat_id_threshold`. Они позволяют тонко калибровать чувствительность обнаружения. Если эти пороги превышаются, событие регистрируется системой с назначенным уровнем тяжести: средний, высокий или критический.

SimulationRunner связывает оба компонента вместе, управляя процессом выполнения. Он гарантирует, что ChannelEmulator и AnomalyDetector работают одновременно с использованием `asyncio` корутинов Python, эмулируя реальный канал связи в реальном времени. Взаимосвязь представленных модулей и компонентов отображена на рисунке 6.

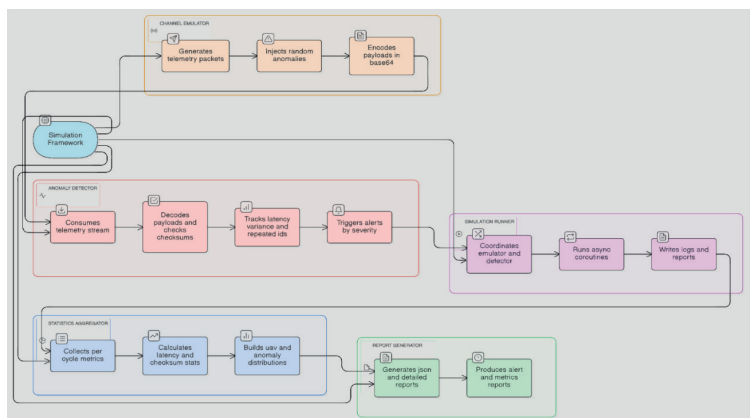


Рисунок 6 – Взаимосвязь модулей и компонентов

Детерминированные случайные начальные значения используются для обеспечения повторяемых результатов очень важной особенностью для верификации и регрессионного тестирования. После каждого прогона SimulationRunner выдает два вида выходных данных: файл .log с меткой времени, записывающий все обнаруженные аномалии, и структурированный отчет .json, обобщающий ключевые метрики из выполнения, тем не менее, как заключение записанных модулей и компонентов, следующая диаграмма объясняет взаимосвязь между компонентами, которые были закодированы.

Обсуждение. Для проверки работоспособности архитектуры была проведена серия экспериментов в разработанной среде симуляции. Основные результаты работы симуляционной платформы телеметрического канала БПЛА представлены на таблице 2.

Таблица 2 – Основные результаты работы симуляционной платформы телеметрического канала БПЛА

№	Показатель	Значение	Интерпретация
1	Количество циклов симуляции	100	Представительный прогон модели
2	Доля инъекции аномалий	10%	Контролируемая нагрузка на систему детекции
3	Обработано пакетов	97	Незначительные потери, обусловленные моделируемыми сбоями
4	Обнаружено несовпадений контрольных сумм	6	Выявлены нарушения целостности данных
5	Общее число зафиксированных аномалий	118	Система демонстрирует высокую чувствительность
6	Средняя задержка пакета	0.037 сек	Стабильная работа канала при нагрузке
7	Дисперсия задержки	1.45×10^{-4}	Низкая вариативность передачи
8	Форматы выходных данных	.log, .json	Двойная верификация и трассируемость
9	Повторяемость эксперимента	Deterministic random seeds	Обеспечение воспроизводимости результатов
10	Среда реализации	Python 3.12 (asyncio)	Асинхронная обработка телеметрии

Большинство событий в полученных результатах симуляций классифицировано как значимые и распределение по уровню критичности составило 100 high, 14 medium, 4 critical.

С точки зрения обработанных данных, большинство аномалий были вызваны высокими показателями отказа контрольной суммы, необычно частыми повторениями идентификаторов БПЛА или идентификацией поддельных меток БПЛА, таких как UAV_612, UAV_162 и UAV_678. Все обнаруженные аномалии регистрировались в файле с названием

anomalies_20251112_163329.log. Журнал событий этого .log файла представлена на рисунке 7.

```

99 [Cycle 19] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.5972352, "rate": 0.058823529411764705, "severity": "high"}
100 [Cycle 20] Anomaly detected: malformed_payload - UAV: UAV_800 - Packet ID: 20
101 [Cycle 20] Alert: spoofed_id - Severity: critical - {"type": "spoofed_id", "timestamp": 1762947210.5508788, "uav_id": "UAV_012", "severity": "critical"}
102 [Cycle 20] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.5508788, "rate": 0.0625, "severity": "high"}
103 [Cycle 20] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.5972352, "rate": 0.058823529411764705, "severity": "high"}
104 [Cycle 20] Alert: malformed_payload - Severity: medium - {"type": "malformed_payload", "timestamp": 1762947210.6299398, "packet_id": 20, "severity": "medium"}
105 [Cycle 20] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.6299398, "rate": 0.11111111111111111, "severity": "high"}
106 [Cycle 21] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.5972352, "rate": 0.058823529411764705, "severity": "high"}
107 [Cycle 21] Alert: malformed_payload - Severity: medium - {"type": "malformed_payload", "timestamp": 1762947210.6299398, "packet_id": 20, "severity": "medium"}
108 [Cycle 21] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.6299398, "rate": 0.11111111111111111, "severity": "high"}
109 [Cycle 21] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.6607776, "rate": 0.10526315789473684, "severity": "high"}
110 [Cycle 22] Alert: malformed_payload - Severity: medium - {"type": "malformed_payload", "timestamp": 1762947210.6299398, "packet_id": 20, "severity": "medium"}
111 [Cycle 22] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.6299398, "rate": 0.11111111111111111, "severity": "high"}
112 [Cycle 23] Alert: high_checksum_mismatch_rate - Severity: high - {"type": "high_checksum_mismatch_rate", "timestamp": 1762947210.6607776, "rate": 0.10526315789473684, "severity": "high"}

```

Рисунок 7 – Log view

Соответствующий отчет .json содержит общее количество обработанных пакетов, статистику задержек и сведения о серьезности, а также временные метки выполнения, которые дают количественное представление каждого прогона моделирования. Структура регистрации, разработанная на этом этапе, оказалась очень полезной для устранения неполадок и проверки достоверности данных. Фактически, неожиданное поведение, такое как отложенные оповещения или пропуски записей об аномалиях, было гораздо проще диагностировать путем сравнения выходных данных в форматах .log и .json. Поддержание этой системы с двойным выходом гарантирует, что каждая аномалия будет хронологически задокументирована и аналитически обобщена, как показано на следующем рисунке 8.

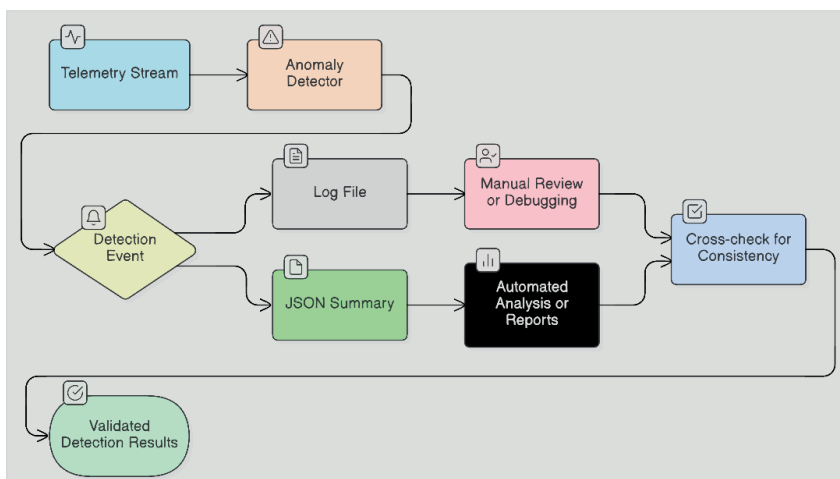


Рисунок 8 – Механизм двойного выхода

Если статистический телеметрический анализ будет постоянно применяться и поддерживаться модульными инструментами моделирования

то в дальнейшем оно имеет возможность предоставить практическую основу для раннего обнаружения аномалий даже при отсутствии алгоритмов машинного обучения. Этот вклад является важным шагом к созданию полноценной системы киберзащиты для платформ БПЛА.

Выводы. Результаты проведённых симуляционных экспериментов раскрывают характер реакций бортовой системы в условиях навигационных и телеметрических искажений, формируемых в контролируемой среде моделирования. Симуляция и тестирование проводилось в последовательном порядке наблюдения за системой, работающей в режиме непрерывной передачи данных. Оно не просто дает возможность фиксировать факт нарушения нормальной работы, но и может анализировать траекторию возникновения аномалий, их развитие и момент перехода системы из устойчивого состояния в предаварийное или аварийное.

Практическая реализация архитектуры в виде программной симуляционной платформы в среде Python 3.12 обеспечила воспроизводимое моделирование телеметрического обмена и контролируемую инъекцию аномалий. Результаты серии экспериментов подтвердили эффективность статистического метода обнаружения аномалий на основе анализа задержек, повторяемости идентификаторов и ошибок контрольных сумм. Система продемонстрировала устойчивость к моделируемым воздействиям при сохранении стабильных временных характеристик передачи данных. Детерминированный статистический анализ телеметрических потоков способен обеспечить раннее выявление киберугроз без привлечения ресурсоёмких алгоритмов машинного обучения, что особенно важно для систем с ограниченными вычислительными ресурсами.

Перспективы дальнейших исследований связаны с углублённой интеграцией криптографического уровня защиты, расширением сценариев атак, а также тестированием предложенного подхода в аппаратно-реальных условиях функционирования БПЛА.

References

- Alshamrani A., and Alghamdi A.M. (2026) Zero-Shot Attack Detection in UAV Networks Using Foundation Models. *Alexandria Engineering Journal*, 136. – P. 105–124. <https://doi.org/10.1016/j.aej.2025.12.065> (in Eng.).
- Alsumayt A., Nagy N., Alsharyofi S., Alahmadi R., Al-Rabie R., Alesse R., Alibrahim N., Alahmadi A., Alghamedy F. H., and Alfawaer Z. (2026) Cutting-Edge DoS Attack Detection in Drone Networks: Leveraging Machine Learning for Robust Security. *Sci*, 8(1). – P. 20. <https://doi.org/10.3390/sci8010020> (in Eng.).
- Burbank J., Caleb T., Andam E., and Kaabouch N. (2026) Detection and Mitigation of Cyber Attacks on UAV Networks. *Electronics*, 15(2). – P. 317. <https://doi.org/10.3390/electronics15020317> (in Eng.).
- Islam M. S., Mahmoud A. S., and Sheltami T. R. (2025) AI-Enhanced Intrusion Detection for UAV Systems: A Taxonomy and Comparative Review. *Drones*, 9(10). – P. 682. <https://doi.org/10.3390/drones9100682> (in Eng.).
- Mohammed U. M., Omolara A. E., Abiodun O. I., Rasheed J., Osman O., Lar P. M., Adeyinka P.

O., and Olugbenga A. G. (2025) Cyber Threat in Drone Systems: Bridging Real-Time Security, Legal Admissibility, and Digital Forensic Solution Readiness. *Frontiers in Communication and Networks*, 6. – P. 1661928. <https://doi.org/10.3389/frcmn.2025.1661928> (in Eng.).

Qiu S., and Liu H. (2022) A New Zonotope-Based Attack Detection Method for UAV. 2022 41st Chinese Control Conference (CCC), Hefei, China. – P. 4276–4280. <https://doi.org/10.23919/CCC55666.2022.9902124> (in Eng.).

Rezaee H., Salvato E., Fenu G., and Parisini T. (2024) Resilient Coverage by Teams of Quadrotor UAVs: Theory and Experiments. *IEEE Transactions on Control Systems Technology*, 32(6). – P. 2009–2022. <https://doi.org/10.1109/TCST.2024.3389350> (in Eng.).

Romagnoli R., Krogh B. H., de Niz D., Hristozov A. D., and Sinopoli B. (2023) Software Rejuvenation for Safe Operation of Cyber-Physical Systems in the Presence of Run-Time Cyberattacks. *IEEE Transactions on Control Systems Technology*, 31(4). – P. 1565–1580. <https://doi.org/10.1109/TCST.2023.3236470> (in Eng.).

Sharifi I., Ghazanfari M., Taye A., Wei P., Ahmed M. H., Kim H. T., Ghasemi M., Gupta V., Dahle N., Canady R., Diaz Gonzalez A., Coursey A., Bjorkman B., Lemieux-Mack C., Ward B. C., Koutsoukos X., Biswas G., Herencia-Zapana H., Hasan S., Amundson I., Fotiadis F., Topcu U., Lu J., Chen Q. A., Aryal N., Ibrahim A., Ras A. K., and Shirkhodaie A. (2026) A Survey of Security Challenges and Solutions for UAS Traffic Management (UTM) and Small Unmanned Aerial Systems (sUAS). *arXiv*. – P. 2601.08229. <https://arxiv.org/abs/2601.08229> (in Eng.).

Tahavori M. (2020) A System-Theoretic Measure for Quantification of Vulnerabilities to Cyber Attacks with Application to Unmanned Aerial Vehicles. 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT), Prague, Czech Republic. – P. 492–495. <https://doi.org/10.1109/CoDIT49905.2020.9263905> (in Eng.).

Tang H., and Chen Y. (2025) Composite Observer-Based Resilient MPC for Heterogeneous UAV-UGV Systems Under Hybrid Cyberattacks. *IEEE Transactions on Aerospace and Electronic Systems*, 61(4). – P. 8277–8290. <https://doi.org/10.1109/TAES.2025.3542737> (in Eng.).

Tlili F., Ayed S., and Fourati L. C. (2024) Exhaustive Distributed Intrusion Detection System for UAVs Attacks Detection and Security Enforcement (E-DIDS). *Computers & Security*, 142. – P. 103878. <https://doi.org/10.1016/j.cose.2024.103878> (in Eng.).

Tucker R. S., Nadeem M., and Pervez S. (2025) Real-Time Detection and Mitigation of GPS Spoofing in UAV Systems. 2025 12th International Conference on Information Technology (ICIT), Amman, Jordan. – P. 154–160. <https://doi.org/10.1109/ICIT64950.2025.11049153> (in Eng.).

Wang P., et al. (2026) QUADFormer: Learning-Based Detection of Cyber Attacks in Quadrotor UAVs. *IEEE Transactions on Control Systems Technology*, 34(1). – P. 59–73. <https://doi.org/10.1109/TCST.2025.3598255> (in Eng.).

Yang H., Yu Z., and Zhang Y. (2026) Observer-Based Adaptive Resilient Fault-Tolerant Cooperative Control for Multiple Fixed-Wing UAVs Subject to Cyberattacks and Actuator Faults. *IEEE Internet of Things Journal*, 13(3). – P. 5179–5192. <https://doi.org/10.1109/JIOT.2025.3642912> (in Eng.).

Yoo J. D., Kim G. M., Song M. G., and Kim H. K. (2025) MeNU: Memorizing Normality for UAV Anomaly Detection with a Few Sensor Values. *Computers & Security*, 150. – P. 104248. <https://doi.org/10.1016/j.cose.2024.104248> (in Eng.).

Yu A., Kolotylo I., Hashim H. A., and Eltoukhy A. E. E. (2025) Electronic Warfare Cyberattacks, Countermeasures, and Modern Defensive Strategies of UAV Avionics: A Survey. *IEEE Access*, 13. – P. 68660–68681. <https://doi.org/10.1109/ACCESS.2025.3561068> (in Eng.).

Zuev A., Gryb O., Shvets S., and Makarov V. (2018) Evaluating and Ensuring the Cybersecurity of Power Line Remote Monitoring Systems. 2018 IEEE 3rd International Conference on Intelligent Energy and Power Systems (IEPS), Kharkiv, Ukraine. – P. 271–274. <https://doi.org/10.1109/IEPS.2018.8559572> (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 414–432

<https://doi.org/10.32014/2026.2518-1726.446>

IRSTI 28.23.15

UDC 004.89

© **Tursynova N.^{1,2}, Yerimbetova A.^{1,3}, Amangeldy N.^{1,2}, Zhumabayeva A.²,
Daiyrbayeva E.^{1,4*}, 2026.**

¹Institute of Information and Computational Technologies, Almaty, Kazakhstan;

²L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;

³META University, Astana, Kazakhstan;

⁴Satbayev University, Almaty, Kazakhstan.

E-mail: nurbekkyzyelmira@gmail.com

COMPARATIVE ANALYSIS OF MULTILINGUAL TRANSFORMER MODELS FOR KAZAKH-TO-GLOSS TRANSLATION

Tursynova Nazira — Software Engineer, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD student, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: ntursynova000@gmail.com, <https://orcid.org/0000-0002-1857-9028>;

Yerimbetova Aigerim — Leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD, Cand. Tech. Sci., Associate Professor, Research Professor, META University, Astana, Kazakhstan,

E-mail: aigerian8888@gmail.com, <https://orcid.org/0000-0002-2013-1513>;

Amangeldy Nurzada — Senior Researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD, Researcher, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: amangeldi_n_3@enu.kz, <https://orcid.org/0000-0002-4669-9254>;

Zhumabayeva Assel — PhD, Senior Lecturer, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

E-mail: zhumabayeva_as@enu.kz, <https://orcid.org/0000-0002-8984-1492>;

Daiyrbayeva Elmira — Researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; Master, senior lecturer, Satbayev University, Almaty, Kazakhstan,

E-mail: nurbekkyzyelmira@gmail.com, <https://orcid.org/0000-0002-4255-5456>.

Abstract. Kazakh sign language is the focus of attention among scientists. In recent decades, increasing attention has been paid to the creation of technologies capable of removing barriers to communication and interaction between people with disabilities and society. The lack of a parallel corpus in sign language opens up the topic of text translation into gloss for urgent and in-depth research. Unfortunately, digital technologies are not accessible to people with hearing and speech impairments. Text-to-gloss translation is still relevant for Kazakh sign language. In particular, because one sign has several meanings, synthesizing it

into gloss in a software environment takes a long time. This article discusses the translation of text into gloss using five different models based on a multilingual transformer. That is, it examines whether the models are faster or slower to read and whether they consume more or less resources. The mBART-large-50, M2M100-418M, mT5-base, MarianMT, and T5-base were selected for the experiment. All of these models were trained using the same parameters. The results were evaluated using widely used standard metrics (BLEU, ROUGE-L, and chrF++), and a qualitative analysis of the models was also performed. The study found that larger multilingual models provide more accurate translations. Although the mBART-large-50 and M2M100-418M models performed well, the other smaller multilingual models were faster and less resource-intensive. Translating text into a glossary is still relevant for Kazakh sign language. In particular, because one sign language has several meanings, synthesizing it into a glossary in a programming environment takes a significant amount of time.

Keywords: Kazakh sign language, text translation to gloss, multilingual Transformer platform, BLEU score, training

Financing. This work was carried out with the financial support of the Science Committee of the Ministry of Science and higher education of the Republic of Kazakhstan (Grant No. BR24992875).

For citations: Tursynova N., Yerimbetova A., Amangeldy N., Zhumabayeva A., Daiyrbayeva E. Comparative analysis of multilingual transformer models for Kazakh-to-gloss translation. *Academic Scientific Journal of Computer Science*, 2026. — No.2. — P. 414-432. DOI <https://doi.org/10.32014/2026.2518-1726.446>

© Турсынова Н.^{1,2}, Еримбетова А.^{1,3}, Амангелді Н.^{1,2}, Жумабаева А.²,
Дайырбаева Э.^{1,4*}, 2026.

¹ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы, Қазақстан;

²Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана, Қазақстан;

³META University, Алматы, Қазақстан;

⁴Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан.

E-mail: nurbekkyzyelmira@gmail.com

КАЗАҚ ТІЛІНЕН ГЛОССҚА АУДАРУ ҮШІН КӨПТІЛДІ ТРАНСФОРМЕРЛІК МОДЕЛЬДЕРДІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ

Турсынова Назира — инженер-бағдарламашы, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», PhD докторант, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: ntursynova000@gmail.com, <https://orcid.org/0000-0002-1857-9028>;

Еримбетова Айгерим — жетекші ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», PhD, т.ғ.к., қауымдастырылған профессор, профессор – зерттеуші, META University, Алматы, Қазақстан,

E-mail: aigerian8888@gmail.com, <https://orcid.org/0000-0002-2013-1513>;

Амангелді Нұрзада — аға ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», PhD, зерттеуші, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: amangeldi_n_3@enu.kz, <https://orcid.org/0000-0002-4669-9254>;

Жумабаева Асель Сагнаевна — PhD, аға оқытушысы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: zhumabayeva_as@enu.kz, <https://orcid.org/0000-0002-8984-1492>;

Дайырбаева Эльмира Нурбекқызы – ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы, Қазақстан; магистр, аға оқытушы, Қ.И.Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан, E-mail: nurbekkyzyelmira@gmail.com, <https://orcid.org/0000-0002-4255-5456>.

Аннотация. Қазақ ым тілі әлі де ғалымдардың назарында. Соңғы онжылдықтарда мүмкіндігі шектеулі жандар мен қоғамның қарым-қатынасы мен өзара әрекеттесуіндегі кедергілерді жоюға қабілетті технологияларды құруға көбірек көңіл бөлінді. Ым тілінде параллельді корпустың аз болуы, мәтінді глоссқа аудару тақырыбы өзекті және терең зерттеуге жол ашады. Себебі есту және сөйлеу қаблетінде шектеу бар адамдар үшін цифрлық технология өкінішке орай қолжетімсіз болып отыр. Қазақ ым тілі үшін әліде мәтіннен глоссқа аудару өзекті болып тұр. Нақтырақ айтқанда бір ым тілінің бірнеше мағнасы болуына байланысты оны программалау ортасында глоссқа синтездеу айтарлықтай уақытты талап етеді. Бұл мақалада көптілді Transformer-ға негізделген бес түрлі модельдерді қолдану арқылы мәтінің глоссқа аударылуы қарастырылады. Яғни модельдердің жылдамырақ немесе әлсіз оқылатыны және ресурстың көп немесе аз жұмсалыуы зерттеледі. Тәжірбиені іске асыру үшін mBART-large-50, M2M100-418M, mT5-base, MarianMT және T5-base модельдері таңдалды. Бұл модельдердің барлығына бірдей параметрлер қолдану арқылы оқытылды. Нәтижелерін кең таралған стандарттар BLEU, ROUGE-L, және chrF++) метрикаларымен бағаланып, модельдерге сапалы талдау жасалды. Зерттеу барысында үлкен көптілді модельдердің дәлірек аударма жасайтыны анықталды. mBART-large-50 және M2M100-418M модельдері жақсы өнімділікті көрсетсе, қалған кішірек көптілді модельдер жылдамырақ оқылады, аз ресурс жұмсайтыны дәлелденді. Қазақ ым тілі үшін әліде мәтіннен глоссқа аудару өзекті болып тұр. Нақтырақ айтқанда бір ым тілінің бірнеше мағынасы болуына байланысты оны программалау ортасында глоссқа синтездеу айтарлықтай уақытты талап етеді.

Түйін сөздер: Қазақ ым тілі, мәтінді глоссқа аудару, көптілді Transformer негізі, BLEU бағалау, оқыту

© Турсынова Н.^{1,2}, Еримбетова А.^{1,3}, Амангелді Н.^{1,2}, Жумабаева А.²,
Дайырбаева Э.^{1,4*}, 2026.

¹Институт информационных и вычислительных технологий,
Алматы, Казахстан;

²Евразийский национальный университет имени Л.Н. Гумилева,
Астана, Казахстан;

³META University, Алматы, Қазақстан;

⁴Казахский национальный исследовательский технический университет им.
К.И. Сатпаева, Алматы, Казахстан.
E-mail: nurbekkyzymira@gmail.com

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МНОГОЯЗЫЧНЫХ ТРАНСФОРМЕРНЫХ МОДЕЛЕЙ ДЛЯ ПЕРЕВОДА С КАЗАХСКОГО ЯЗЫКА НА ГЛОССИРОВАННОЕ ПРЕДСТАВЛЕНИЕ

Турсынова Назира — инженер-программист, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; PhD-докторант, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: ntursynova000@gmail.com, <https://orcid.org/0000-0002-1857-9028>;

Еримбетова Айгерим — ведущий научный сотрудник, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; PhD, кандидат технических наук, ассоциированный профессор, профессор-исследователь META University, Алматы, Казахстан,
E-mail: aigerian8888@gmail.com, <https://orcid.org/0000-0002-2013-1513>;

Амангелді Нұрзада — старший научный сотрудник, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; PhD, исследователь, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: amangeldi_n_3@enu.kz, <https://orcid.org/0000-0002-4669-9254>;

Жумабаева Асель — PhD, старший преподаватель, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: zhumabayeva_as@enu.kz, <https://orcid.org/0000-0002-8984-1492>;

Дайырбаева Эльмира — научный сотрудник, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; магистр, старший преподаватель, Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан,
E-mail: nurbekkyzymira@gmail.com, <https://orcid.org/0000-0002-4255-5456>.

Аннотация. *Актуальность.* Казахский язык жестов является важным направлением исследований в области инклюзивных цифровых технологий и обработки естественного языка. В последние десятилетия возрастает потребность в создании интеллектуальных систем, способных снижать барьеры коммуникации между людьми с нарушениями слуха и речи и обществом. Одной из актуальных задач является перевод текста на глоссированное представление, поскольку для казахского языка жестов отсутствуют полноразмерные параллельные корпуса, а неоднозначность жестовых единиц усложняет автоматическую обработку и синтез голосов.

Цель. Провести сравнительный анализ многоязычных трансформных моделей для перевода с казахского языка на глоссированное представление и оценить их качество, скорость работы и ресурсоемкость. Методы. В исследовании рассмотрен перевод текста в глоссы с использованием пяти моделей на основе многоязычных трансформеров: mBART-large-50, M2M100-418M, mT5-base, MarianMT и T5-base. Все модели обучались с использованием одинаковых параметров, что обеспечило сопоставимость экспериментальных результатов. Оценка качества перевода проводилась с применением распространенных метрик BLEU, ROUGE-L и chrF++, а также качественного анализа выходных последовательностей. Дополнительно анализировались вычислительная скорость моделей и уровень потребления ресурсов. Результаты и выводы. Результаты исследования показали, что более крупные многоязычные модели обеспечивают более точный перевод с казахского языка на глоссированное представление. Модели mBART-large-50 и M2M100-418M продемонстрировали более высокое качество перевода, тогда как меньшие по размеру модели оказались быстрее и менее ресурсоемкими. Полученные результаты подтверждают наличие компромисса между точностью перевода и вычислительной эффективностью. Практическая значимость исследования заключается в возможности использования сравнительного анализа при выборе архитектуры для систем автоматического перевода казахского текста в глоссы, образовательных и инклюзивных цифровых платформ, а также технологий поддержки коммуникации для людей с нарушениями слуха и речи.

Ключевые слова: казахский язык жестов, перевод текста в глоссы, глоссированное представление, многоязычные трансформеры, mBART-large-50, M2M100-418M, mT5-base, MarianMT, T5-base, BLEU, ROUGE-L, chrF++

Introduction. Kazakh sign language (KSL) is the main method of communication for deaf citizens in Kazakhstan. Most digital services still uses only Kazakh written language. Usually it is not problematic, but problem is occurred when deaf citizens have to deal with urgent information without assistance. Fact that often surprises non-professional saqwrsqwerrdfts is that the sign language is not “visual version” of the spoken language: it is the separate linguistic system with own grammar. That is why providing written information (even if it uses official language) may not always mean that it is samely available for KSL users.

Practical approach to ensuring accessibility is to convert written text into intermediate symbolic layer. It is easier to deal with them then dealing with signs. In this study glossary is used as such layer. Here glossary is the sequence of the standardized tokens which corresponds to the signs. But gloss is not complete written form of KSL, rather it is the compact notation which is used for annotation and modelling. The main importance of glossary is that it allows to represent complex multimodal target to discrete sequence which simplifies usage

of sequence-to-sequence methods preserving main linguistical cues. Notable that some small lexemes carry disproportionate large meaning, omitting them the meaning is changed definitely: questions may become statement or tenses may be changed. And this is not a bug but it changes the communicative intent.

Thus, translation from text to gloss can become a base for extension of technological capabilities of KSL. After development of the glossary it may be used to synthesis of signs, avatar animations, extracting sign content or editing it manually without recording new videos. Intermediate stage makes more interpretable the behavior of the system: not translation not completed successfully, output gloss may illustrate if the model lost the important phrases, incorrectly deals with negative sentences or repeatedly uses same patterns. However, progress in KSL is restricted by two structural bottlenecks. Firstly, the count of parallel corpora to translate Kazakh language to glossary is limited. Secondly, annotation process is expensive and long as it requires professional signers and consistent labeling convention: even a little difference may break control and reduce generalization.

Neural Machine Translations (NMT) provides a potential solution to these constraints through transfer learning. Of all the available methods, the Transformer architecture stands out. Transformer architecture is an encoder-decoder model based on a self-attention mechanism. Currently, Transformer is becoming very popular because it is effective in dealing with long-distance dependencies (Vaswani et al., 2017). Multilingual pretrained models provide a unified system for multiple languages. It is very well suited for languages with limited resources, such as Kazakh (Joshi et al., 2020). It can be easily applied to solve various problems by configuring it for a specific task.

The models discussed in this article, such as mBART, M2M-100, and mT5, differ in their language coverage, tokenization methods, and optimization. These solutions can help in the subsequent generations, especially when the amount of training data is limited (Liu et al., 2020; Fan et al., 2020; Xue et al., 2021). This makes them particularly attractive for Kazakh sign language, which has a small amount of parallel corpora.

However, despite the rapid development of multilingual NMT, less attention is currently being paid to sign language translation. Most of the work on sign language focuses on video-to-text translation, where the main challenges lie in recognizing images from video rather than processing text data (Camgöz et al., 2018; Camgöz et al., 2020).

On the other hand, it raises additional questions when translating from text to gloss. For example, how can existing models be adapted, to what extent can they be adjusted, and how will the size of the model affect performance?

There is also a limitation related to experimental research. In the absence of carefully prepared tests, the claimed improvement may be due to preprocessing, optimization, or decoding methods rather than the actual capabilities of a particular model.

To address the above issues, this article presents the results of a comparative

analysis of five pretrained Transformer models (mBART-large-50, M2M100-418M, mT5-base, MarianMT, and T5-base). These five models were tested with similar configurations and fine tuning. They were evaluated using standard metrics such as BLEU, ROUGE-L, and chrF++. The main goal of this experimental study is to determine how multilingual pretraining performs for Kazakh sign language. The following questions were also considered: (1) Which model shows the best results in gloss generation? (2) How different are the models in terms of performance for the Kazakh language? (3) Which types of errors are most resistant to correction and why?

Literary review. Sign language is fully developed and consists of a natural language that is expressed through movements of the face, body, and hands. Sign languages have their own distinct grammatical systems and vocabularies (Bragg et al., 2019). Government agencies provide written information to deaf and hard-of-hearing people. As a result, an asymmetrical communication environment arises everywhere, as they do not have access to all the necessary government services. This is because deaf and hard-of-hearing people do not have a standardized writing system.

Therefore, sign language translation should be viewed not only as a convenience for deaf and hard-of-hearing people, but also as an important tool for improving access and reducing inequality in healthcare, education, and government digital services. Real-time sign language translation systems must meet the high latency requirements of practical applications while maintaining semantic accuracy when trained on real-world data.

These limitations are particularly important when inaccuracies in translation can have various consequences (Papatsimouli et al., 2023). From a scientific point of view, accuracy and adaptability to internal and external changes in sign language translation are considered interrelated, as they promote the use of a modeling approach rather than simple conversion.

Most modern sign language translation systems, including the SLT system, use intermediate spaces to facilitate multimodal correspondence processing (Liang et al., 2023). Using a set of methods and techniques, sign2gloss2text converts a sign language image into a text gloss, which is then translated into spoken language. Thus, a gloss can be understood as a compact text representation, since words are extracted from it and translated into sign language.

The difference between standard text form and gloss text in sign language has practical significance. Gloss sequences are subject to systematic sources of noise. In particular, annotators may differ in how they segment or mark up a given text in a given sign language. Gloss dictionaries vary across different datasets and research groups (Liang et al., 2023). However, despite their shortcomings, glosses are used continuously in practice. That is, they convert continuously moving video into a sequence of characters. This allows the use of traditional “chain-to-chain” models. Manual glossing is a costly and labor-intensive process, and there is growing interest in model architectures and training schedules that reduce the

amount of costly intermediate supervision (Camgöz et al., 2020; Lin et al., 2023). Therefore, researchers are looking for an optimal way to introduce glosses that does not require a lot of labor and resources.

A lot of researchers worked closely on CSLR even before it became possible to develop truly reliable sign language translation. This is because continuous conversion of all data through recognition into translation streams based on a glossary is a necessary condition. During an experiment on continuous sign language recognition using the CSLR method, it was found that visual space and time are closely related in sign language. That is, the study found that finger movements, facial expressions indicating internal states, and head movements are interrelated. It is important to note that each of these movements conveys different grammatical information (Rastgoo et al., 2021). In the process of recognizing sign language and individual signs, the CSLR system immediately detects small changes in external gestures. In particular, changes in gestures during speech, changes in posture, as well as shading or overlapping of objects. These variations make it difficult for the system to distinguish between linguistically meaningful movements and non-linguistic movements (Cheok et al., 2019).

The researcher (Koller et al., 2018) believes that the application of deep learning has significantly improved it by systematically organizing the spatiotemporal representation of sign language data. To prove this, the Deep Sign project demonstrated that large-scale training combined with appropriate temporal modeling can lead to a significant increase in the accuracy of continuous recognition. Modern methods have been proposed to reduce the number of simultaneously perceived meaningful texts, as well as the additional gestures required to understand the meaning of these texts. However, this approach is problematic in design due to the large amount of data (Li & Meng, 2022). As a result, continuous translation work aimed at maximizing efficiency suffers in real-world conditions. That is, errors made in the perception of ambiguous texts are systematic in nature, as they have a high potential for direct propagation to subsequent translations.

Transformer-based models have been found to be highly effective in sign language translation. In particular their mechanism is so effective that it allows for modeling long distances and understanding and coordinating information from various sources, such as video and text, at a high level. The “chain-to-chain” rule is used in neurotranslation research on sign language. According to this rule, it is considered more effective to link video sequences with spoken language. Its advantage is that cascading recognition allows for systematic analysis in accordance with a single training and evaluation rule in the translation process (Camgöz et al., 2018). In response to the above-mentioned problems, Sign Language Transformers has developed a scalable structure for sign language translation. The main purpose of the device is to reduce the accumulation of errors inherent in rigid cascade pipelines by combining recognition and translation (Camgöz et al., 2018).

Some researchers have shown that glosses are not essential for the system. A

method has been proposed for representing the compatibility of video and sign language translation in text without using glosses and without losing the main meaning. This method provides high-quality translation by teaching the system the relationship between sign language and speech. It uses additional signals to establish semantic links between video and text. This means that intermediate labels, which are often checked when using the model, are clearly unnecessary (Li et al., 2025). One of the main reasons for abandoning the use of glosses is to use convenient solutions and prevent data corruption (Lin et al., 2023).

On the other hand, the ability to translate sign language without the help of gloss has been proven to be poor. Yes, although it is possible to translate sign language into text using video, the second problem is that written text can only be translated with the help of gloss. To use it to convert the translated gloss into sign language (Liang et al., 2023). The use of glosses has contributed significantly to the development of sign language. The text is intended to be used to reveal meaning and is interpreted using gestures. An example of this is the Text2Sign system. The system clearly demonstrates the possibilities of creating gestures for intermediate scenes (Stoll et al., 2020). Even within a single sign language, different glossary standards are of particular interest for different language units. It is considered more appropriate to view a glossary as a program interface rather than a linguistic transcription. The conversion of text into a glossary depends not only on the architecture of the model, but also on the correct construction of the dataset. Otherwise, data truncation and other obstacles may arise (Matsumoto et al., 2006).

There are certain requirements for the gloss-to-text conversion model to work correctly. With fewer parallel corpora, the translation capabilities are more limited. This is because the quality of the translation depends on the amount of data collected. For example, suppose there is little data, but the translation of a chemical subject common to both languages into gloss will be of low quality due to the small amount of data, i.e., the parallel corpus (Haddow et al., 2022; Ranathunga et al., 2023). This requirement also applies to neural machine translation. This is because high-quality translation is not possible when working with a small amount of data in neural machine translation. (Dabre et al. , 2021).

There is also an approach to modeling based on subwords. It is used for languages with complex morphology. In this method, words taken from the dictionary are removed, and words that have not been encountered before are recognized. To increase the stability of the model, one text can be divided into several lines, and regularization can be used to visualize the quality of the model (Kudo, 2018). Another approach is to use multilingual training mode to train models on a small number of examples without any training. This feature helps translate common multilingual scenes from annotated data sources without supervision (Johnson et al., 2017). The lack of a parallel corpus for translation from Kazakh to the glossary is also exacerbated by the lack of a single standard for the glossary.

There are several approaches that have proven beneficial when translating

sign language into text. First, let's look at the BART structure. It reduces noise through pre-training and effectively builds a generative model for text generation. The next model is the Transformer model, which is pre-trained using a chain-of-chains approach. This model provides reliable initialization because it encodes general linguistic patterns in its parameters (Lewis et al., 2020). mBART reduces noise similarly to the BART structure mentioned above. However, it extends pre-training to a multilingual environment, using pre-training to reduce noise. It has been shown that initializing the encoder and decoder significantly improves translation scenarios in resource-limited and domain-specific environments (Liu et al., 2020).

Text-to-text formulation in T5 allows wide range of NLP tasks including translation to be represented as conditional generation task within the same framework (Raffel et al., 2019). Further, mT5 adapted this paradigm to multilingual pretraining for more than 100 languages (Xue et al., 2021). Simultaneously, developed a lot of other multilingual systems, such as M2M-100, which lets to deal with non-English texts (Fan et al., 2020). Also tools are essential here: Marian provides an efficient and research oriented toolkit for NMT which reduces implementation costs and evaluation process (Junczys-Dowmunt et al., 2018).

Overall, these findings stimulate comparative analysis in pretrained multilingual Transformers. Differences of the models are in the objectives of pretraining, in language coverage or in capacity allocation, and it may impact on performance when the training data is limited as in Kazakh Sign Language.

Kazakh Sign Language (KSL) remains considerably under-resourced in comparison with more extensively studied sign languages, and only recently has the literature begun to report reusable parallel resources. Initial efforts have focused on constructing corpora that align Kazakh written text with KSL data, thereby establishing the basic prerequisites for systematic model training and evaluation (Yerimbetova et al., 2025). At the same time, much of the existing KSL-related research continues to concentrate on recognition-oriented tasks, such as alphabet or gesture classification. While these directions are valuable, they do not directly address the translation problem that lies at the core of many accessibility-oriented applications (Buribayev et al., 2025).

To date, there has been no controlled and reproducible comparison of multilingual encoder-decoder architectures for translation from Kazakh to gloss during training and decoding. To address this issue, this paper conducts a comparative analysis of five pre-trained Transformer models, supplemented by a targeted analysis of linguistically significant shortcomings, using a unified fine-tuning protocol and a common set of automated evaluation metrics.

In particular, there are still a number of unresolved issues in translating from Kazakh into the glossary. These include: (i) the limited and potentially inconsistent nature of glossary conventions, (ii) apparent sensitivity to tokenization choices and morphological variations, and (iii) uncertain behavior during transfer when multilingual pre-trained models are adapted to glossary notation that does not

correspond to the standard written language. The situation is further complicated by evaluation: recent metric studies show that automatic evaluations may only imperfectly correlate with semantic preservation, especially for short, template-based, or appearance-dependent outputs (Lee et al., 2023).

Overall, existing research still does not provide a controlled and reproducible comparison of multilingual encoder-decoder architectures for translation from Kazakh to glossary. To fill this gap, this paper conducts a comparative analysis of five pre-trained Transformer models, supplemented by a targeted analysis of linguistically significant shortcomings, using a common tuning protocol and a common set of automatic evaluation metrics.

Materials and methods. In this section, the method of translation from Kazakh language to KSL using Transformer architecture is described. The structure of Transformer architecture is illustrated in Figure 1.

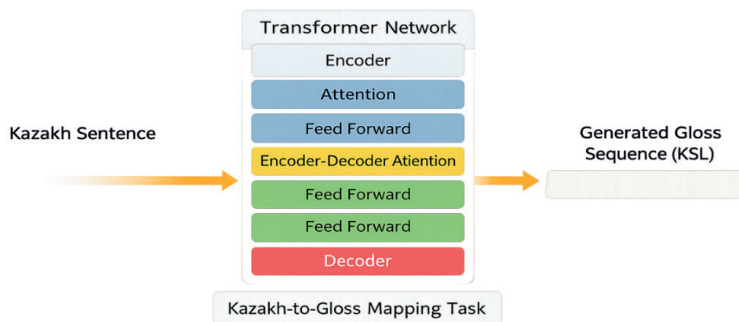


Figure 1 – Architecture of the Kazakh-to-Gloss translation model based on an encoder--decoder Transformer. The encoder processes the Kazakh input sequence using multi-head self-attention and feed-forward layers, producing contextual representations. The decoder generates the gloss sequence autoregressively using masked self-attention and encoder--decoder (cross) attention.

A. Problem Statement

Suppose that the input sentence is represented as a sequence of tokens

$$s = (s_1, s_2, \dots, s_N), \quad (1)$$

and the corresponding target gloss sequence as

$$g = (g_1, g_2, \dots, g_M) \quad (2)$$

The objective is to learn a model that estimates the conditional probability of the target sequence given the source

$$P(g|s) = \prod_{t=1}^M P(g_t | g_{<t} s) \quad (3)$$

where $\mathbf{g}_{<t}$ denotes the previously generated gloss tokens.

B. Model Architecture

Transformer architecture is an encoder–decoder model based on a self-attention mechanism. The encoder converts the input sentence into a sequence of contextual representations. The decoder outputs a sequence of tokens based on the input and output of the previous word. The internal components of the Transformer architecture are shown in Figure 1.

1) Token Embedding and Positional Encoding: To begin with, an embedding vector is created for each input token s_p , and information about the position is stored there.

$$\mathbf{Z}^{(0)} = \text{Emb}(\mathbf{s}) + \text{PosEnc}(1:N), \quad (4)$$

where $\mathbf{Z}^{(0)} \in \mathbb{R}^{N \times d}$ denotes the input to the encoder stack and d is the model dimensionality. Source and target sequences are padded or truncated to a maximum length of 128 tokens.

2) Encoder Layers: The encoder consists of multiple identical layers, each comprising multi-head self-attention followed by a position-wise feed-forward network. For a given layer input, attention projections are computed as

$$\mathbf{A} = \mathbf{Z}\mathbf{W}^{(a)}, \mathbf{B} = \mathbf{Z}\mathbf{W}^{(b)}, \mathbf{C} = \mathbf{Z}\mathbf{W}^{(c)}, \quad (5)$$

where $\mathbf{W}^{(a)}, \mathbf{W}^{(b)}, \mathbf{W}^{(c)}$ are learned projection matrices. Scaled dot-product attention is defined as

$$\text{SA}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \text{softmax}\left(\frac{\mathbf{A}\mathbf{B}^\top}{\sqrt{d_h}}\right) \mathbf{C}, \quad (6)$$

where d_h denotes the dimensionality of each attention head. Multi-head attention concatenates multiple such heads and applies a linear projection. Multi-head attention combines several attention mechanisms into a single representation.

Then feed-forward sublayer is applied to each word:

$$\text{FF}(\mathbf{u}) = \phi(\mathbf{u}\mathbf{W}_1 + \beta_1)\mathbf{W}_2 + \beta_2 \quad (7)$$

Here, $\phi(\cdot)$ nonlinear activation function. Residual connections and layer normalization are applied to all sublayers and then these sublayers are normalized.

3) Decoder and Cross-Attention. The decoder is similar in structure to the encoder, but it works with already generated glosses and translates them one by one. Let

$$\mathbf{R}^{(0)} = \text{Emb}(\mathbf{g}_{<t}) + \text{PosEnc}(1:t-1) \quad (8)$$

be the decoder input representation. Masked self-attention applies a causal mask $\mathbf{M}$:

$$\text{MSA} = \text{softmax}\left(\frac{\mathbf{AB}^\top + \mathbf{M}}{\sqrt{d_h}}\right) \mathbf{C} \quad (9)$$

Encoder--decoder attention conditions the decoder on encoder outputs $\mathbf{E}$:

$$\text{CA}(\mathbf{D}, \mathbf{E}) = \text{softmax}\left(\frac{\mathbf{DE}^\top}{\sqrt{d_h}}\right) \mathbf{E} \quad (10)$$

allowing the model to align gloss tokens with relevant source representations.

4) Output Projection: The decoder state $\mathbf{r}_t$ is mapped to the vocabulary distribution as

$$P(g_t | \mathbf{g}_{<t} s) = \text{softmax}(\mathbf{W}_o \mathbf{r}_t + \gamma), \quad (11)$$

where $\mathbf{W}_o$ and γ are learned parameters.

C. Optimization Objective

During training, the model compares predicted gloss with correct one and is penalized for each mismatch. As a result, the model learns to bring its predictions closer to the correct translation:

$$\mathcal{J} = - \sum_{t=1}^M \log P(g_t | \mathbf{g}_{<t} s) \quad (12)$$

Empty tokens (padding) are excluded from the process so they don't affect the training.

D. Training Configuration

All models are trained under the same conditions with same configurations. Hugging Face Transformers library is used. Experiments are conducted on an NVIDIA RTX 3080 GPU with 10 GB of memory. The learning rate is set to 3×10^{-5} , the batch size is 4, and each model is trained for five epochs. Mixed-precision training is employed in order to reduce memory consumption and improve computational efficiency.

E. Evaluation

Model performance is evaluated using BLEU, computed with SacreBLEU. The BLEU score is defined as

$$\text{BLEU} = \text{BP} \cdot \exp \left(\sum_{n=1}^N \alpha_n \log \pi_n \right), \quad (13)$$

where π_n denotes modified n -gram precision, α_n are uniform weights, and BP is the brevity penalty. After each epoch, the model is evaluated on test sets.

Results. This section reports both quantitative and qualitative results for the Kazakh to gloss translation task. Five Transformer-based models are evaluated under identical training conditions, and their performance is compared in terms of automatic evaluation metrics, training efficiency, and representative translation examples.

A. Overall Performance Comparison

Figure 2 presents a comparison of BLEU scores across the evaluated Transformer models, together with their corresponding training times. Among all models, mBART-large-50 achieves the highest BLEU score, followed closely by M2M100-418M. Models with fewer parameters, such as MarianMT and T5-base, exhibit lower BLEU scores but substantially reduced training time.

The results highlight a clear trade-off between translation quality and computational cost. Larger multilingual pretrained models benefit from richer cross-lingual representations, which translate into higher accuracy for gloss generation.

B. Quantitative Evaluation Metrics

Table 1 summarizes the BLEU, ROUGE-L, and chrF++ scores obtained by all evaluated models. The relative ordering of systems is consistent across metrics, with mBART-large-50 achieving the best overall performance, followed by M2M100-418M and mT5-base.

The nearly seven-point BLEU difference between mBART-large-50 and T5-base suggests that both model capacity and multilingual pretraining substantially influence the preservation of semantic content in gloss generation. ROUGE-L and chrF++ exhibit comparable patterns, indicating that higher BLEU scores are accompanied by improvements in both sequence-level alignment and character-level accuracy.

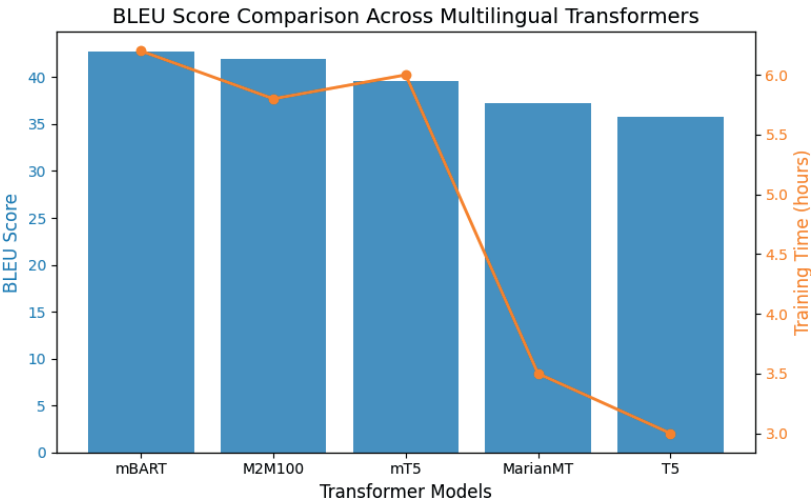


Figure 2 - Comparison of BLEU scores and training time across Transformer models for Kazakh-to-Gloss translation. Larger multilingual models achieve higher BLEU scores at the cost of increased training time.

Table 1 – Automatic evaluation results for Kazakh-to-Gloss translation

Model	BLEU	ROUGE-L	chrF++
mBART-large-50	42. 7	0. 78	0. 61
M2M100-418M	41. 9	0. 77	0. 60
mT5-base	39. 6	0. 75	0. 58
MarianMT	37. 2	0. 73	0. 55
T5-base	35. 8	0. 70	0. 53

scores correspond to improved sequence-level alignment and character-level accuracy.

C. Model Size and Training Efficiency

Table 2 reports the number of parameters and the observed training time for each model. While mBART-large-50 and mT5-base have similar parameter scales, mT5-base attains slightly lower BLEU scores, suggesting that differences in architecture and pretraining objectives influence downstream performance beyond raw model size alone.

Table 2 – Model size and training time comparison

Model	Parameters (M)	Training Time (hrs)
mBART-large-50	610	6. 2
M2M100-418M	418	5. 8
mT5-base	580	6. 0
MarianMT	300	3. 5
T5-base	220	3. 0

MarianMT and T5-base require roughly half the training time of mBART-large-50, but incur a reduction of approximately 5–7 BLEU points. Such a trade-off may be acceptable in settings where computational resources are constrained or rapid experimentation is prioritized.

D. Qualitative Analysis

Table 3 provides representative translation examples that illustrate common error patterns. Higher-performing models, including mBART and M2M100, tend to preserve temporal expressions, interrogative markers, and canonical gloss word order. By contrast, smaller models more frequently omit modifiers or question particles and show greater variability in word ordering.

Table 3 – Qualitative examples of Kazakh-to-Gloss generation with per-example BLEU (in parentheses)

Kazakh input	Reference gloss	mBART	M2M100	MarianMT	T5
Мен үйге ерте келемін.	МЕН ҮЙ ЕРТЕ КЕЛ	МЕН ҮЙ ЕРТЕ КЕЛ (100)	МЕН ҮЙ ЕРТЕ КЕЛ (100)	МЕН ҮЙ КЕЛ ЕРТЕ (54)	МЕН ҮЙ КЕЛ (41)
Біз дүкеннен нан сатып алдық.	БІЗ ДҮКЕН НАН САТЫП АЛ	БІЗ ДҮКЕН НАН САТЫП АЛ (100)	БІЗ ДҮКЕНДЕН НАН САТЫП АЛ(82)	БІЗ НАН АЛ ДҮКЕН (46)	БІЗ НАН АЛ (33)
Сен ертең университетке барасың ба?	СЕН ЕРТЕҢ УНИВЕРСИТЕТ БАР МА	СЕН ЕРТЕҢ УНИВЕР- СИТЕТ БАР МА (100)	СЕН ЕРТЕҢ УНИВЕР- СИТЕТКЕ БАР МА (78)	СЕН ЕРТЕҢ УНИВЕР- СИТЕТ БАР (52)	СЕН УНИВЕР- СИТЕТ БАР (38)

Discussions. These examples indicate that interrogative particles and temporal modifiers are especially prone to omission in lower-capacity models. Such errors are of particular importance in sign language contexts, where these elements are often realized through distinct facial expressions or specific signing patterns.

Overall, the experimental findings reveal a consistent relationship between model capacity, multilingual pretraining, and Kazakh-to-gloss translation quality. Larger models, including mBART-large-50 and M2M100, achieve higher BLEU, ROUGE-L, and chrF++ scores, reflecting improved preservation of both semantic content and gloss-level structure. This performance gain, however, is accompanied by increased training time and computational cost.

By contrast, smaller models such as MarianMT and T5-base converge more quickly and require substantially fewer resources, but they exhibit systematic omissions of temporal modifiers and interrogative markers in qualitative analyses. These shortcomings are particularly consequential for sign language processing, where such elements frequently correspond to grammatical distinctions and non-manual features.

The resulting trade-off between accuracy and efficiency suggests that model selection should be informed by application-specific constraints. High-precision

gloss generation pipelines are better served by larger multilingual models, whereas resource-limited scenarios may favor smaller architectures despite reduced semantic completeness.

Taken together, the reported results establish strong baseline performance for Kazakh-to-gloss translation and offer empirical guidance for the development of future Kazakh Sign Language systems, including downstream sign synthesis and avatar-based generation.

Conclusion. This paper presented a systematic investigation of Transformer-based models for Kazakh-to-gloss translation as a foundational step toward computational processing of Kazakh Sign Language. By formulating the task within a sequence-to-sequence framework and evaluating multiple multilingual pretrained architectures under a unified experimental protocol, we established clear performance baselines for this under-resourced setting.

The experimental results show that larger multilingual models, most notably mBART-large-50 and M2M100-418M, deliver superior performance across BLEU, ROUGE-L, and chrF++ metrics. In contrast, smaller models achieve shorter training times but exhibit reduced semantic completeness. Qualitative analyses further indicate that lower-capacity models tend to omit temporal modifiers and interrogative markers, elements that play an important linguistic role in sign language representations.

Taken together, these findings underscore a practical trade-off between translation quality and computational efficiency, suggesting that model choice should be informed by the requirements of the intended application. This study provides an initial benchmark for Kazakh-to-gloss translation and establishes a foundation for future work on end-to-end Kazakh Sign Language systems, including integration with sign synthesis components and evaluation involving native signers.

References

- Bragg D., Koller O., Bellard M., Berke L., Boudreault P., Braffort A., Caselli, N., Huenerfauth M., Kacorri H., Verhoef T., Vogler C., & Morris M.R. (2019) Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. — P. 16–31. <https://doi.org/10.1145/3308561.3353774> (in Eng.)
- Buribayev Z., Aouani M., Zhangabay Z., Yerkos A., Abdirazak Z., & Zhassuzak M. (2025) Enhancing Kazakh Sign Language recognition with BiLSTM using YOLO keypoints and optical flow. *Applied Sciences*, 15(10). — 5685 p. <https://doi.org/10.3390/app15105685> (in Eng.)
- Camgöz N. C., Hadfield S., Koller O., Ney H., & Bowden R. (2018) Neural sign language translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — P. 7784–7793. <https://doi.org/10.1109/CVPR.2018.00812> (in Eng.)
- Camgöz N.C., Koller O., Hadfield S., & Bowden R. (2020) Sign language transformers: Joint end-to-end sign language recognition and translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — P. 10020–10030. <https://doi.org/10.48550/arXiv.2003.13830> (in Eng.)
- Cheok M.J., Omar Z., & Jaward M.H. (2019) A review of hand gesture and sign language recognition techniques. *International Journal of Machine Learning and Cybernetics*, 10. — P. 131–

153. <https://doi.org/10.1007/s13042-017-0705-5> (in Eng.)

Dabre R., Chu C., & Kunchukuttan A. (2021) A survey of multilingual neural machine translation. *ACM Computing Surveys*, 53(5), Article 99. <https://doi.org/10.1145/3406095> (in Eng.)

Fan A., Bhosale S., Schwenk H., Ma Z., El-Kishky A., Goyal S., Baines M., Celebi O., Wenzek G., Chaudhary V., Goyal N., Birch T., Liptchinsky V., Edunov S., Grave E., Auli M., & Joulin A. (2020) Beyond English-centric multilingual machine translation. *arXiv*. <https://doi.org/10.48550/arXiv.2010.11125> (in Eng.)

Haddow, B., Kocmi, T., Birch, A., & Bogoychev, N. (2022). Survey of low-resource machine translation. *Computational Linguistics*, 48(3). — P. 673–732. https://doi.org/10.1162/coli_a_00446 (in Eng.)

Johnson M., Schuster M., Le Q.V., Krikun M., Wu Y., Chen Z., Thorat N., Viégas F., Wattenberg M., Corrado G., Hughes M., & Dean J. (2017) Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics*, 5. — P. 339–351. <https://doi.org/10.1162/tacla00065> (in Eng.)

Koller O., Zargaran O., Ney H., & Bowden R. (2018) Deep sign: Enabling robust statistical continuous sign language recognition via deep learning. *International Journal of Computer Vision*, 126. — P. 1311–1325 <https://doi.org/10.1007/s11263-018-1121-3> (in Eng.)

Kudo T. (2018) Subword regularization: Improving neural network translation models with multiple subword segmentations. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*. — P. 66–75. <https://doi.org/10.48550/arXiv.1804.10959> (in Eng.)

Lee J., Kim H., & Lee J. (2023) A survey of automatic metrics for machine translation evaluation. *Mathematics*, 11(4), 1006. <https://doi.org/10.3390/math11041006> (in Eng.)

Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed A., Levy O., Stoyanov V., & Zettlemoyer L. (2020) BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. — P. 7871–7880. <https://doi.org/10.18653/v1/2020.acl-main.703> (in Eng.)

Li H., & Meng J. (2022) Multi-view sign language recognition with deep learning: A review and new framework. *Applied Intelligence*, 52. — P. 14534–14560 <https://doi.org/10.1007/s10489-022-03407-5>

Li X., Zhang Y., & Wang Z. (2025) Gloss-free sign language translation based on fusion attention. *Applied Soft Computing*, 185, 113848. <https://doi.org/10.1016/j.asoc.2025.113848> (in Eng.)

Liang Z., Li H., & Chai J. (2023) Sign language translation: A survey of approaches and techniques. *Electronics*, 12(12), 2678. <https://doi.org/10.3390/electronics12122678> (in Eng.)

Lin K., Wang X., Zhu L., Sun K., Zhang B., & Yang Y. (2023) Gloss-free end-to-end sign language translation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*. — P. 12904–12916. <https://doi.org/10.48550/arXiv.2305.12876> (in Eng.)

Liu Y., Gu J., Goyal N., Li X., Edunov S., Ghazvininejad M., Lewis M., & Zettlemoyer L. (2020) Multilingual denoising pre-training for neural machine translation. *arXiv*. <https://doi.org/10.48550/arXiv.2001.08210> (in Eng.)

Papatsimouli M., Sarigiannidis P., & Fragulis G. (2023) A survey of advancements in real-time sign language translators: Integration with IoT technology. *Technologies*, 11(4), 83. <https://doi.org/10.3390/technologies11040083> (in Eng.)

Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., & Liu P.J. (2019) Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv*. <https://doi.org/10.48550/arXiv.1910.10683> (in Eng.)

Ranathunga S., Plank B., & El-Kishky A. (2023) Neural machine translation for low-resource languages: A survey. *ACM Transactions on Asian and Low-Resource Language Information Processing*. <https://doi.org/10.1145/3567592> (in Eng.)

Rastgoo R., Kiani K., & Escalera S. (2021) Sign language recognition: A deep survey. *Expert Systems with Applications*, 164, 113794. <https://doi.org/10.1016/j.eswa.2020.113794> (in Eng.)

Stoll S., Camgöz N., Hadfield S., & Bowden R. (2020) Text2Sign: Towards sign language

production using neural machine translation and generative adversarial networks. *International Journal of Computer Vision*, 128. — P. 891–908. <https://doi.org/10.1007/s11263-019-01281-2> (in Eng.)

Xue L., Constant N., Roberts A., Kale M., Al-Rfou R., Siddhant A., Barua A., & Raffel C. (2021) mT5: A massively multilingual pre-trained text-to-text transformer. *arXiv*. <https://doi.org/10.48550/arXiv.2010.11934> (in Eng.)

Yerimbetova A., Sakenov B., Sambetbayeva M., Daiyrbayeva E., Berzhanova U., & Othman M. (2025) Creating a Parallel Corpus for the Kazakh Sign Language and Learning. *Applied Sciences*, 15(5), 2808. <https://doi.org/10.3390/app15052808> (in Eng.)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 433–448

<https://doi.org/10.32014/2026.2518-1726.447>

IRSTI 50.37.15

UDC 336.77:338.43(477)

© Shangpeng Lei, Balakayeva G.*, 2026.

Al-Farabi Kazakh National University, Almaty, Kazakhstan.

E-mail: gulnardtsa@gmail.com

DUAL-BRANCH PHYSICAL INFORMATION NEURAL NETWORKS FOR DATA CENTER AIRFLOW VELOCITY AND THERMAL MODELING

Shangpeng Lei — Doctoral student of Computer Science, Department of Computer Science, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: leishangpeng@foxmail.com, <https://orcid.org/0009-0002-7156-4631>;

Balakayeva Gulnar — Doctor of Physical and Mathematical Sciences, Professor, Department of Computer Science, Al-Farabi Kazakh National University, Almaty, Kazakhstan,

E-mail: gulnardtsa@gmail.com, <https://orcid.org/0000-0001-9440-2171>.

Abstract. With the rapid development of large-scale artificial intelligence models, the construction of data centers has been gradually accelerating. Data centers are equipped with numerous servers and network devices, and energy consumption is expected to keep increasing in the future. To improve thermal management, data center energy modeling has become a new challenge. Traditional numerical simulations are time-consuming and poor transferable, while data-driven machine learning models provide a feasible solution for data center modeling. Physics-informed neural networks embed prior physical knowledge into the loss function as soft constraints, guiding the modeling process to follow both data patterns and physical laws. To decouple multi-physics field modeling in data centers, we propose a dual-branch physics-informed neural network (Dual-PINN). Meanwhile, we design an adaptive loss re-weighting mechanism on the loss function based on the logistic function (Sigmoid) nonlinear transformation to balance the convergence speed among different loss terms. In the experiments, to comprehensively evaluate the proposed Dual-PINN, we choose the standard PINN, standard neural network, and random forest for comparison. Finally, we used MAE, RMSE, and the number of trainable parameters to measure model performance in terms of accuracy and complexity. The results demonstrate that Dual-PINN achieves the highest prediction accuracy with the fewest trainable parameters, compared with standard PINN, Dual-PINN reduces the MAE and RMSE by 19.95% and 15.79% for

temperature prediction, 34.01% and 31.86% for velocity prediction, respectively. The Dual-PINN provides a novelty solution to physical fields modeling to optimize the operation of data center.

Keywords: Physics-informed neural networks, data center, thermal management; prediction

For citations: Shangpeng Lei, Balakayeva G. Dual-branch physical information neural networks for data center airflow velocity and thermal modeling. Academic Scientific Journal of Computer Science, 2026. — No.2. — P.433-448. DOI <https://doi.org/10.32014/2026.2518-1726.447>

© Шанпэн Лей, Балакаева Г.*, 2026.

Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

E-mail: gulnardsa@gmail.com.

ДЕРЕКТЕР ОРТАЛЫҒЫНЫҢ АУА АҒЫНЫНЫҢ ЖЫЛДАМДЫҒЫНА ЖӘНЕ ТЕРМИЯЛЫҚ МОДЕЛЬДЕУГЕ АРНАЛҒАН ЕКІ ТАРМАҚТЫ ФИЗИКАЛЫҚ АҚПАРАТТЫҚ НЕЙРОНДЫҚ ЖЕЛІЛЕР

Шанпен Лей — Әл-Фараби атындағы Қазақ ұлттық университетінің информатика кафедрасының информатика докторанты, Алматы, Қазақстан,

E-mail: leishangpeng@foxmail.com, <https://orcid.org/0009-0002-7156-4631>;

Балакаева Гүлнар — физика-математика ғылымдарының докторы, профессоры, әл-Фараби атындағы Қазақ ұлттық университетінің информатика кафедрасы, Алматы, Қазақстан,

E-mail: gulnardsa@gmail.com, <https://orcid.org/0000-0001-9440-2171>.

Аннотация. Ірі ауқымды жасанды интеллект модельдерінің қарқынды дамуына байланысты деректер орталықтарының құрылысы біртіндеп жеделдеуде. Деректер орталықтары көптеген серверлермен және желілік құрылғылармен жабдықталған, сондықтан олардың энергия тұтынуы болашақта одан әрі артады деп күтіледі. Жылулық басқаруды жақсарту мақсатында деректер орталықтарының энергетикалық моделін құру жаңа өзекті мәселе ретінде қарастырылуда. Дәстүрлі сандық модельдеу әдістері көп уақытты талап етеді және олардың тасымалдану қабілеті шектеулі, ал деректерге негізделген машиналық оқыту модельдері деректер орталықтарын модельдеу үшін тиімді шешім ұсынады. Физикалық ақпаратпен толықтырылған нейрондық желілер априорлық физикалық білімді шығын функциясына жұмсақ шектеулер түрінде енгізіп, модельдеу процесін деректер заңдылықтарына да, физикалық заңдарға да сәйкес бағыттайды. Деректер орталықтарындағы көпфизикалық өрістерді модельдеуді декомпозициялау үшін бұл жұмыста екі тармақты физикалық ақпаратпен толықтырылған нейрондық желі (Dual-PINN) ұсынылады. Сонымен қатар, әртүрлі шығын мүшелерінің жинақталу жылдамдығын теңестіру мақсатында логистикалық

функцияның (Sigmoid) бейсызық түрлендіруіне негізделген шығындарды бейімделмелі қайта салмақтау механизмі әзірленді. Эксперименттерде ұсынылған Dual-PINN моделін жан-жақты бағалау үшін салыстыру әдістері ретінде стандартты PINN, стандартты нейрондық желі және кездейсоқ орман тандалды. Соңында, модельдің дәлдігі мен күрделілігін бағалау үшін MAE, RMSE және оқытылатын параметрлер саны қолданылды. Нәтижелер Dual-PINN моделінің ең аз оқытылатын параметрлер санымен ең жоғары болжау дәлдігіне қол жеткізетінін көрсетеді. Стандартты PINN моделімен салыстырғанда, Dual-PINN температураны болжауда MAE және RMSE көрсеткіштерін тиісінше 19,95% және 15,79%-ға, ал жылдамдықты болжауда тиісінше 34,01% және 31,86%-ға төмендетеді. Dual-PINN деректер орталықтарының жұмысын оңтайландыру үшін физикалық өрістерді модельдеудің жаңа шешімін ұсынады.

Түйін сөздер: Физикаға негізделген нейрондық желілер, деректер орталығы, жылулық басқару, болжау

© Шанпэн Лэй, Балакаева Г.*, 2026.

Казахский национальный университет имени аль-Фараби,

Алматы, Казахстан.

E-mail: gulnardsa@gmail.com

ДВУХВЕТВЕВЫЕ ФИЗИЧЕСКИЕ ИНФОРМАЦИОННЫЕ НЕЙРОННЫЕ СЕТИ ДЛЯ МОДЕЛИРОВАНИЯ ВОЗДУШНЫХ ПОТОКОВ И ТЕПЛОВЫХ УСЛОВИЙ В ЦЕНТРАХ ОБРАБОТКИ ДАННЫХ

Шанпэн Лэй — докторант кафедры компьютерных наук, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,

E-mail: leishangpeng@foxmail.com, <https://orcid.org/0009-0002-7156-4631>;

Балакаева Гульнар — доктор физико-математических наук, профессор кафедры информатики, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан,

E-mail: gulnardsa@gmail.com, <https://orcid.org/0000-0001-9440-2171>.

Аннотация: *Актуальность.* Быстрое развитие крупномасштабных моделей искусственного интеллекта сопровождается ускоренным строительством и расширением центров обработки данных. Такие центры оснащены большим количеством серверов и сетевого оборудования, что приводит к росту энергопотребления и повышает требования к эффективности теплового управления. В этих условиях энергетическое и тепловое моделирование центров обработки данных становится актуальной научно-прикладной задачей. Традиционные численные методы моделирования требуют значительных вычислительных ресурсов и обладают ограниченной переносимостью, тогда как модели машинного обучения могут обеспечить более гибкое и эффективное решение. *Цель.* Разработать и оценить

двухветвевую физически информированную нейронную сеть Dual-PINN для моделирования воздушных потоков и тепловых условий в центрах обработки данных. *Методы.* В исследовании предложена двухветвевая физически информированная нейронная сеть Dual-PINN, предназначенная для декомпозиции моделирования мультифизических полей в центрах обработки данных. Физически информированные нейронные сети включают априорные физические знания в функцию потерь в виде мягких ограничений, что позволяет модели одновременно учитывать закономерности данных и физические законы. Дополнительно разработан адаптивный механизм перевзвешивания потерь, основанный на нелинейном преобразовании логистической функции Sigmoid, для балансировки скорости сходимости различных компонентов функции потерь. В экспериментальной части модель Dual-PINN сравнивалась со стандартной PINN, стандартной нейронной сетью и моделью случайного леса. Для оценки точности и сложности модели использовались метрики MAE, RMSE и количество обучаемых параметров. *Результаты и выводы.* Полученные результаты показывают, что Dual-PINN достигает наиболее высокой точности прогнозирования при наименьшем количестве обучаемых параметров. По сравнению со стандартной PINN предложенная модель снижает MAE и RMSE при прогнозировании температуры на 19,95% и 15,79% соответственно, а при прогнозировании скорости воздушного потока - на 34,01% и 31,86% соответственно. Результаты подтверждают эффективность двухветвевой архитектуры и адаптивного перевзвешивания потерь при моделировании связанных физических процессов в центрах обработки данных. Практическая значимость исследования заключается в возможности применения Dual-PINN для оптимизации теплового управления, повышения энергоэффективности и поддержки принятия инженерных решений при эксплуатации центров обработки данных.

Ключевые слова: физически информированные нейронные сети, Dual-PINN, PINN, центр обработки данных, воздушные потоки, тепловое управление, прогнозирование, MAE, RMSE, Sigmoid

Introduction. With the development of artificial intelligence (AI) and big data technologies, the requirements for data storage and computation have been continuously increasing. To ensure the efficient and reliable utilization of data resources, centralized data centers have become critical infrastructure in the digital revolution. Data centers are equipped with a large number of servers and network devices, which generate considerable heat during operation, imposing strict requirements on the control of the building environment (Lu et al., 2026) the energy consumption of AI tasks has experienced exponential growth, and how to schedule arriving AI tasks in a low-carbon manner is worth investigating for data centers. However, due to the computationally intensive and resource-demanding properties of AI tasks, current deferral-based scheduling methods cannot efficiently fit a large AI task (e.g., training a large language model. Therefore, to guarantee the reliable

operation of data centers, a significant amount of electricity is consumed by cooling systems to maintain the data center temperature within an optimal range. Currently, data centers worldwide consume approximately 460 TWh of electricity annually, accounting for about 2% of global electricity consumption (Liu et al., 2026). Given that emerging large AI models with billions of parameters require massive data resources during training, the power consumption of data centers will continue to rise in the future.

To optimize the design and operation of data centers, effective measures are necessary to be taken to reduce energy consumption while ensuring the reliable performance of IT equipment. Accurate and efficient airflow and thermal modeling techniques provide a fundamental basis for decision. Although experimental proofs can provide accurate and detailed data for modeling, they are limited by high material and labor costs. Importantly, data centers host numerous cloud service systems involving critical fields, making experimental investigations highly impractical. Traditionally, data center modeling relies highly on physical simulations, such as the simplified lumped-capacitance models (Rasku et al., 2024) and computational fluid dynamics (CFD) simulations (Lin et al., 2023).
 id: "ITEM-1", issue: "3", issued: {"date-parts": [{"2023"}]}, page: "571-590", publisher: "IEEE", title: "Thermal modeling and thermal-aware energy saving methods for cloud data centers: A review", type: "article-journal", volume: "9", uris: [{"http://www.mendeley.com/documents/?uuid=164bd5c1-e06a-47fc-9147-6438a140b541"}]}, mendeley: {"formattedCitation": "(Lin et al., 2023. The simplified lumped-capacitance models are computationally fast and low-cost. However, with the increasing requirements of automatic control and the spatial layout, the thermal coupling has become more complex. Simplified lumped-capacitance models cannot accurately describe the spatiotemporal variations of temperature and airflow. While CFD simulations have long been the dominant method for thermal behavior modeling in data centers. By setting the geometry and boundary conditions, CFD discretizes the target domain into grids and computes the airflow and temperature distribution. Nevertheless, the detailed parameter input and condition setup result in a high knowledgeable barrier. Meanwhile, with the construction of large-scale data centers and multi-timescale control strategies, CFD faces challenges in terms of computational cost and generalization ability."}

With the development of machine learning (ML) techniques, data-driven models have emerged as an efficient alternative. By extracting the relationship between input and output variables from massive observation data, parameterized ML models are constructed to characterize the spatiotemporal variations of environmental parameters in data centers. However, the performance of data-driven models relies on the reliability and size of the training dataset, which limits their application scenarios and compatibility. Physics-informed machine learning (PIML) embeds prior knowledge or physical theorems into the ML model as soft constraints, alleviating the heavy dependence of data-driven models on training data. As a representative method of PIML, physics-informed neural networks (PINNs) (Lawal

et al., 2022) take the residuals of physical constraints as a regularization loss term, guiding neural networks to learn data patterns while complying with physical laws.

Data center modeling has been mainly focused on customized physical field prediction for specific applications (LOI et al., 2026) Heating, Ventilation, and Air Conditioning (HVAC, and accelerated computation using PINNs in physical simulations (Deng et al., 2026) machine learning (ML. However, the inconsistency in magnitude and significance between the data residual loss and the physical constraint loss leads to convergence difficulties in ML models. Several studies have proposed methods to address convergence issues, such as optimizing the loss function (Wang et al., 2025) providing a powerful approach for solving forward and inverse PDE problems. During training, physical loss calculation relies on predefined spatiotemporal collocation points. However, when solving equations with steep gradients or singularities, conventional fixed or randomly distributed training points often fail to capture critical solution structures, reducing PINN's prediction accuracy. Inspired by firefly phototaxis, this paper proposes a bio-inspired dynamic training point movement strategy named firefly adaptive collocation point movement (FAM and adjusting the neural network architecture (Zheng et al., 2026) we develop interface-gated physics-informed neural networks (IG-PINNs. Nevertheless, PINNs for multi-physics field prediction in data centers still lack sufficient attention.

In this paper, we propose a dual-branch PINN (Dual-PINN) for predicting the temperature and velocity distribution in data centers. To decouple the physical fields, we employ a dual-branch architecture. Meanwhile, to align the convergence pace of losses between the two branches, we apply the logistic function for normalization mapping. Finally, we conduct experiments to evaluation the performance of the model. Our contributions are as follows:

- Architecture design. We design a novel Dual-PINN to decouple temperature and velocity into branches for prediction and alleviate convergence conflicts.

- Loss function optimization. We reconstructed the loss function and adjusted the weights of loss terms using the logistic function.

- Application Expansion. We successfully use a novel Dual-PINN for the data centers airflow velocity and thermal modeling.

Literary review.

Physics-informed neural networks.

The application of PINNs in data centers focuses on the prediction and analysis of physical fields. As a variant of PINNs, Physics-Consistent Neural Networks (PCNNs) are applied to building thermal modeling for the first time, regarding physical information and data labels as distinct modules (Di Natale et al., 2022). To improving the generalization of PCNNs, the adaptive physical consistency neural network (APCNNs) framework replaces the traditional preset system with the Softplus activation function, reducing the parameter trial-and-error cost for temperature prediction (D. Chen et al., 2025). Furthermore, knowledge graphs are combined with PINNs, raising the interpretability and response of PINNs, to achieve

accurate prediction of the airflow field in data centers (LOI et al., 2026). In terms of optimization control, the combination of PINNs and model predictive control (MPC) effectively improves the robustness and generalization ability of thermal control (D. Chen et al., 2026). With the digital transformation of data centers, the cost of data acquisition has decreased. A parameterized PINN framework is developed, and the hybrid modeling combining physical constraints and labeled data effectively improves the prediction accuracy (Hashemi et al., 2025).

Dual-branch neural networks.

Dual-branch neural network architectures are commonly applied to solve strongly coupled problems in fields such as image recognition (Long et al., 2025) and industrial control (Zhao and Li, 2024). Information transmission in a single neural network suffers from mutual interference, whereas a dual-branch architecture enables the decoupling of coupled variables. In the building energy sector, dual-branch neural networks have been successfully applied to the energy consumption prediction of building chilled water systems (Chen et al., 2025) and the anomaly detection of building energy (Tian et al., 2026) equipment status, and human factors, posing significant challenges to the effective identification of anomaly patterns within these data. To address these challenges, a multi-task learning model based on a dual-channel graph attention network (GAT, demonstrating their applicability to complex building energy systems. Dual-branch neural networks for data center modeling have mainly focused on attention mechanisms, for resource allocation (Yu et al., 2026) scheduling, energy optimization, and sustainability. However, existing models face challenges in long-term cyclic predictions due to nonlinear and irregular fluctuations in resource data. To address this issue, we propose STL(C-TS or anomaly detection (Sun et al., 2025) this paper introduces a bidirectional cross-attention LSTM-Informer with uncertainty-aware multi-task learning framework (BiCA-LI via attention adjustment. However, research that combines dual-branch neural networks with physics-informed mechanisms remains blank.

Multiphysics modeling of data centers.

Building energy consumption is affected by multiple physical factors such as building layout and airflow distribution. Although the ultimate goal is to reduce energy consumption, multi-physics field analysis helps evaluate the reliability of results and analyze the spatiotemporal distribution of energy consumption. The research on multi-physics thermal management in data centers mainly focuses on liquid-cooled data centers. The Transformer-GRU model combines the self-attention mechanism with the gated recurrent unit (GRU) to achieve self-attention focus on the time series of data center liquid cooling (Ma et al., 2025). Although liquid cooling represents the future trend of data centers, the energy consumption of traditional air-cooled data centers cannot be ignored. The concept of an energy map is introduced to decouple the cooling process of data centers and analyze thermal energy changes (Li et al., 2025). Reinforcement learning has been used to comprehensively consider environmental parameters and operational data of data centers, thereby optimizing energy-efficient control (Suresh et al., 2026) the

framework learns cooperative control strategies that eliminate redundant cooling. The agents' learning is guided by a novel physics-informed reward function that divides the server's thermal headroom into distinct operational zones, adding penalties to mitigate fan vibrations while dynamically balancing energy efficiency and thermal safety. To validate generalization, the MADDPG algorithm is trained in a simulation environment and subsequently deployed on experimental mock-up servers. A total of five configurations and power maps are used for validation. Each fan agent relies solely on local temperatures of its state space, while the centralized critic receives the global state of the server during training to penalize redundant cooling actions. The MADDPG controller reduced fan energy consumption by an average of 31.4 % compared to a conventional fan-table controller, while maintaining all component temperatures below their critical thresholds. The results also revealed that performance is highly dependent on server layout, with energy savings ranging from 43.8 % in centrally-located CPU configurations to 20.5 % when CPUs are at the chassis extremes, highlighting the importance of hardware-aware control policies.”,”author”:[{“dropping-particle”:"",“family”:”Suresh”,“given”:”Pratheek”,“non-dropping-particle”:"",“parse-names”:false,”suffix”:""},{“dropping-particle”:"",“family”:”Chang”,“given”:”Kai-Wei”,“non-dropping-particle”:"",“parse-names”:false,”suffix”:""},{“dropping-particle”:"",“family”:”-Lua”,“given”:”Kim Boon”,“non-dropping-particle”:"",“parse-names”:false,”suffix”:""},{“dropping-particle”:"",“family”:”Wang”,“given”:”Chi-Chuan”,“non-dropping-particle”:"",“parse-names”:false,”suffix”:""}],“container-title”:”Applied Energy”,“id”:"ITEM-1”,“issued”:{“date-parts”:[["2026"]]},“page”:"127168”,“title”:"Deep reinforcement learning for energy-efficient thermal management in 2U air-cooled server systems”,“type”:"article-journal”,“volume”:"404"},“uris”:[“http://www.mendeley.com/documents/?uuid=08e37a79-99db-4a6b-b4f1-a88a0b4c9174”]],“mendeley”:{“formattedCitation”:(Suresh et al., 2026. However, existing data-driven modeling methods for data centers have not fully considered the physical characteristics of thermal management.

Methods.

Preliminary.

The theoretical foundation of PINNs is the universal function approximation theorem, which states that neural networks have the ability to approximate any continuous function, providing a theoretical basis for solving partial differential equations (PDEs) using neural networks (Hornik et al., 1990). Standard PINN (SPINN) formulates a loss function that includes physical residuals, transforming PDE solving into an optimization problem of neural network parameters. Specifically, the approximate solution output by the neural networks for a physical problem is computed based on the loss function:

$$\mathcal{L}_r = \frac{1}{N_r} \sum_{k=1}^{N_r} [\mathcal{F}(\hat{u}_k) - f(X_k)] \quad (1)$$

$$\mathcal{L}_b = \frac{1}{N_b} \sum_{k=1}^{N_b} [\mathcal{G}(\hat{u}_k) - g(X_k)] \quad (2)$$

$$\mathcal{L}_{data} = \frac{1}{N_{data}} \sum_{k=1}^{N_r} [\hat{u}_k - u_k] \quad (3)$$

$$\mathcal{L}_t = \mathcal{L}_r + \mathcal{L}_b + \mathcal{L}_{data} \quad (4)$$

where the physics loss $\mathcal{L}_r$ ensures physical consistency, the boundary loss $\mathcal{L}_b$ enforces the boundary conditions, and the data loss $\mathcal{L}_{data}$ is used to fit the observed data. N_r , N_b and N_{data} denote the numbers of sampling points for each term, respectively; u_k represents the observed data, and $\hat{u}_k$ represents the predicted solution; $\mathcal{F}$ and $\mathcal{G}$ denote the operator in interior and boundary domain. PINN achieves a deep integration of physical laws and data-driven learning through the cyclic process of neural network forward propagation, residual calculation via automatic differentiation, loss function construction, and backpropagation optimization.

Dual-branch physical information neural networks

As shown in Figure 1, we improve the SPINN architecture. Dual-PINN adopt the same input as the SPINN and use a multi-layer perceptron (MLP) structure.

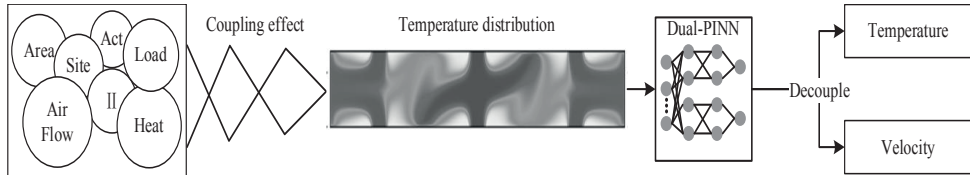


Figure 2 – Coupling mechanism of data center thermal management

Data center thermal management involves multiple coupled factors, such as computing load, operating behavior, and airflow distribution. In Figure 2, we use a dual-branch PINN to decouple thermal management into airflow velocity and temperature, which facilitates the analysis of the physical laws inside the data center. To alleviate the convergence conflict between the two branches during training, the branches share the input layer of the network and employ a shared hidden layer. To avoid additional redundancy caused by the dual-branch structure, the number of hidden neurons in each branch is set to half of that in the shared hidden layer.

Reconstruction of loss function.

We employ a hybrid modeling approach, where the model input is the coordinate position, and the outputs are the predicted values of temperature and velocity. In

the training phase, labeled data are used, and the residuals between the predicted and observed values of temperature and velocity are adopted to compute the data losses and . The physics-consistent loss uses the Pearson correlation coefficient to measure the correlation between temperature and velocity. In the airflow thermal management of data centers, to improve heat dissipation efficiency, positions with higher flow velocity in the cold aisle usually correspond to lower temperatures. Therefore, we use the Pearson correlation coefficient to quantify the negative correlation between temperature and velocity, so as to characterize the underlying physical information. We improve the loss function by introducing an adaptive loss re-weighting mechanism. is transformed by the Sigmoid function and then multiplied by to guide the gradient descent of temperature and velocity. Notably, the Sigmoid function yields a positive output even when the input is zero, which prevents the model from neglecting the velocity gradient. As an additional regularization term, imposes a constraint on the loss function and guides the model to converge while considering the physically consistent negative correlation between temperature and velocity. Accordingly, our final loss function is given as follows:

$$\mathcal{L}_u = \frac{1}{N} \sum_{k=1}^N [\widehat{u}_k - u_k] \quad (5)$$

$$\mathcal{L}_s = \frac{1}{N} \sum_{k=1}^N [\widehat{u}_s - u_s] \quad (6)$$

$$\mathcal{L}_p = \frac{\sum_{k=1}^{N_r} [\widehat{u}_k - \bar{u}_k][\widehat{u}_k - \bar{u}_k]}{\sqrt{\sum_{k=1}^{N_r} [\widehat{u}_k - u_k]^2} \sqrt{\sum_{k=1}^{N_r} [\widehat{s}_k - s_k]^2}} \quad (7)$$

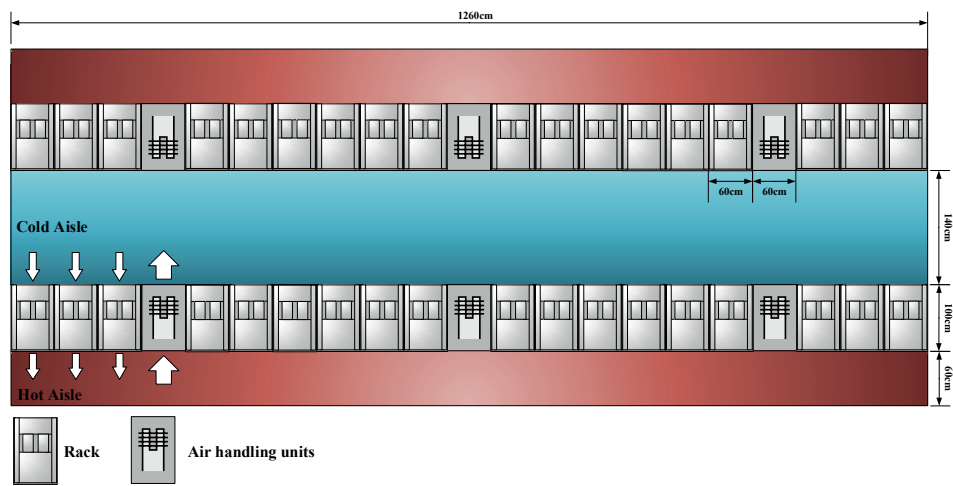
$$w = \text{sigmoid}(\mathcal{L}_u) \quad (8)$$

$$\mathcal{L}_t = w\mathcal{L}_s + \mathcal{L}_p \quad (9)$$

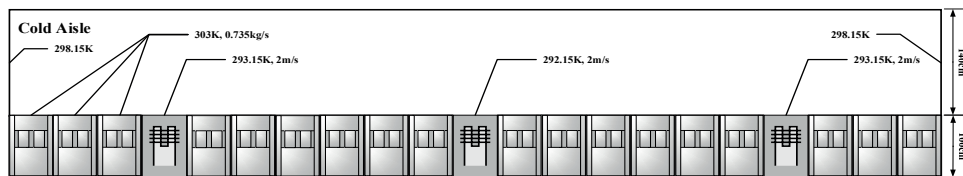
Experiments

Problem description.

In data center thermal management, the horizontal supply of cold air can force the hot air to be exhausted laterally, thereby reducing the influence of vertical buoyancy on the cooling performance. Thus, the airflow circulation can be described by a two-dimensional model(Hashemi et al., 2025)requiring accurate airflow and temperature predictions. This study pioneers the application of physics-informed neural networks (PINNs, which is suitable for evaluating the performance of the PINNs framework.



As shown in Figure 3, this study adopts the layout from case in the literature(Hashemi et al., 2025)requiring accurate airflow and temperature predictions. This study pioneers the application of physics-informed neural networks (PINNs. The data center is equipped with two rows of 36 server racks and six air handling units (AHUs). Air in the hot aisle enters the cold aisle through the AHUs, and the air in the cold aisle passes through the racks into the hot aisle, realizing air circulation. The entire data center is symmetrically distributed.



In this case, heat exchange mainly occurs in the cold aisle. Taking the cold aisle as the research object, we impose the boundary conditions as shown in Figure 4. As the cold aisle is symmetrically distributed, we present the boundary conditions for the lower half. To meet practical requirements, the boundary conditions at both ends are set to ambient temperature. The supply air temperature of the air handling units in the middle region is lower than that on both sides. The boundary condition for the air handling units is velocity inflow, while that for the server racks is mass flow rate.

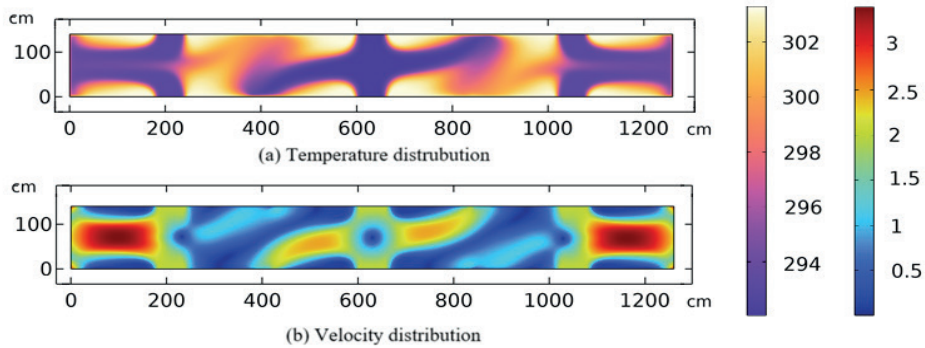


Figure 5 – Temperature and velocity numerical simulation results

Notably, we modified the boundary conditions to be more realistic, as shown in Figure 4, which differs from the original literature. All raw data were recalculated using the scientific numerical simulation software COMSOL 6.3, which has been widely adopted for thermal management simulations. The simulation results are shown in Figure 5.

Dataset preparation.

The simulation dataset has 6049 data points, including 228 boundary points and 5821 internal points. Since the length of data center is 1260 cm and symmetric about its central axis, we select the first 80% of the horizontal distance as the training set (i.e., points with $x < 1008$ cm) and the remaining 20% as the test set. The training set contains 4839 points, and the test set contains 1210 points.

Experimental settings.

We compare four models: the proposed Dual-PINN, the SPINN, standard neural network (SNN), and random forest (RF). SPINN was used to evaluate the effectiveness of the dual-branch architecture, and SNN was adopted to verify the role of the adaptive physics-informed loss. RF, which is widely used in prediction tasks, represents the baseline performance of traditional machine learning models. The learning rate for all neural network models was set to 0.01, with 4 hidden layers and 30 neurons per layer, except for the branch neurons of the Dual-PINN, which were set to 15. For the RF model, the number of decision trees was set to 100 and the maximum depth to 10. Notably, SPINN adopts the traditional single network structure with loss function as shown in Eq. (9). However, the SNN does not employ any physical constraints with loss function as shown in Eq. (10).

$$\mathcal{L}_t = \mathcal{L}_s + \mathcal{L}_u \quad (10)$$

RMSE and MAE are used as evaluation metrics to quantify the prediction performance in Eq. (11) and Eq. (12). Furthermore, to comprehensively evaluate the parameter characteristics of each model, we statistically analyzed the number of trainable parameters for different neural network models.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (12)$$

Results. As shown in Figure 6, Dual-PINN and PINN converge faster than SNN, indicating that the embedding of physical information is effective. Meanwhile, Dual-PINN converges slower than SPINN in the early stage, which is attributed to the need for coordinating the convergence pace of the two branches. Although no significant difference in loss variation is observed between SPINN and Dual-PINN, Dual-PINN achieves better prediction performance and exhibits stronger generalization ability, as shown in Table 1. The final results demonstrate that Dual-PINN achieves the best performance in both temperature and velocity prediction with the fewest trainable parameters. Compared with SPINN, Dual-PINN reduces the MAE and RMSE by 19.95% and 15.79% for temperature prediction, and by 34.01% and 31.86% for velocity prediction, respectively, demonstrating the significant effectiveness of the dual-branch structure. The improvement of performance is even more pronounced compared with SNN, which showing a MAE decrease in temperature by 24.51% and in velocity by 25.89%. The prediction performance of RF is between that of SNN and SPINN. In particular, although the architecture of Dual-PINN is more complex, we reduce the number of neurons in each branch. As a result, the number of trainable parameters in Dual-PINN is reduced by 16.32% compared with other neural network models with the same number of hidden layers.

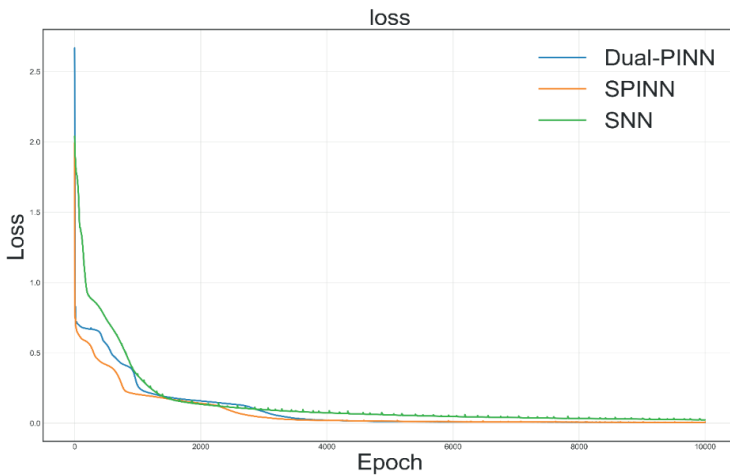


Figure 6 – Loss with epochs of different neural networks

Table 1 – Evaluation metrics of different models. Here ‘–’ denotes that the trainable parameters cannot be counted as the same as neural networks models.

Models	MAE		RMSE		Trainable parameters
	Temperature	Velocity	Temperature	Velocity	
Dual-PINN	3.6793	0.9455	4.5589	1.1703	2462
SPINN	4.5962	1.4326	5.4140	1.7175	2942
SNN	4.8742	1.3050	6.1515	1.5737	2942
RF	4.6052	1.2675	5.1360	1.4995	–

Figure 7 shows the prediction error distributions of different models. It can be observed that all models reasonably capture the characteristics of temperature distribution and airflow velocity. In the temperature prediction error distribution, SNN yields the largest error, with the boundary error reaching 14.4. Random Forest exhibits the most uniform error distribution. Dual-PINN and SPINN achieve a good balance between error distribution and error range. In particular, as shown in Figure 7(a), Dual-PINN provides superior prediction performance at the airflow intersection in the middle of the aisle compared with other models. For airflow velocity prediction, SPINN has the largest error, while Dual-PINN performs the best. Meanwhile, from the error distribution, Dual-PINN has the smallest high-error region. Furthermore, the velocity error distribution is consistent with the temperature error distribution for all models except Dual-PINN. Benefiting from its dual-branch structure, Dual-PINN avoids introducing erroneous guidance from temperature prediction into velocity prediction, demonstrating the effectiveness of the proposed decoupling strategy.

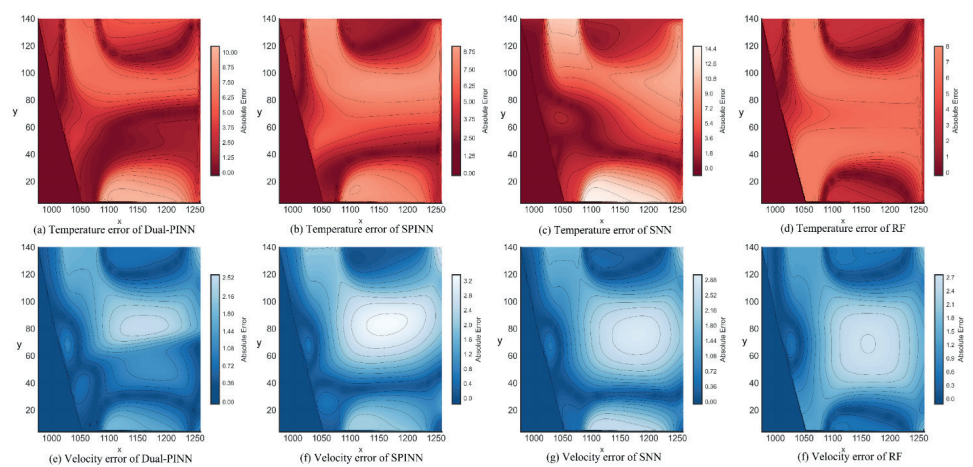


Figure 7 – Temperature and velocity errors distribution of different models

Discussion. PINNs provide a novel solution for data center modelling. The performance of data-driven PINNs models can be significantly enhanced. Under the

guidance of physical information, the modelling error of data centers is reduced, and the error distribution becomes more uniform, which provides valuable references for thermal management of data center. In addition, architecture optimization of PINNs should be emphasized in multi-physics modelling of data centers. Multi-branch architectures can better decouple physical quantities and improve modelling performance. However, PINNs still face challenges. For instance, in temperature prediction, the maximum error of Dual-PINN is not the smallest, even though its error distribution is more uniform. In the future, how to optimize the convergence performance of branch-structured PINNs and further improve prediction accuracy deserves more attention. Furthermore, we plan to further optimize architecture of Dual-PINN to balance the number of branches and multi-physics field prediction. Meanwhile, transient modeling of data centers in three-dimensional scenarios will also be considered to meet the requirements of complex energy operation control.

Conclusion. This paper presents a Dual-PINN for modeling temperature and airflow velocity in data center. We adopt a hybrid modeling approach and embed the Pearson correlation coefficient between predicted temperature and airflow velocity into the loss function as a soft physical constraint. Meanwhile, an adaptive loss reweighting mechanism is designed, in which the temperature residual is nonlinearly transformed via the Sigmoid function and then multiplied by the velocity residual. In the experiments, we compare the proposed method with SPINN, SNN, and RF. The results demonstrate that Dual-PINN achieves the best prediction performance. This study explores the application of Dual-PINN in data centers, providing a reference for the use of PINNs in modeling within the data center energy domain.

References

- Chen A., Wu M., Chen C., Chen C., Huang Y., and Lv X. (2025) DNTB: Dual-branch network model based on transformer and Bi-LSTM for energy consumption prediction in building chiller systems. *PLoS One*, 20(10). — P. e0330187. (in Eng.).
- Deng P., Xu Q., Zhu X., and Dong W. (2026) CFD-based data-driven machine learning temperature prediction: case study on a 180-rack server room within a large-scale data center. *International Journal of Thermal Sciences*, 221. — P. 110465. <https://doi.org/https://doi.org/10.1016/j.ijthermalsci.2025.110465> (in Eng.).
- Hashemi Z., Gholampour M., Su H.-H., Chiang C.-Y., Chang H.-H., and Wang C.-C. (2025) Hybrid physics-informed neural networks for airflow and thermal modeling in data center environments. *International Communications in Heat and Mass Transfer*, 168. — P. 109511. <https://doi.org/https://doi.org/10.1016/j.icheatmasstransfer.2025.109511> (in Eng.).
- Hornik K., Stinchcombe M., and White H. (1990) Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural networks*, 3(5). — P. 551–560. (in Eng.).
- Lawal Z.K., Yassin H., Lai D.T.C., and Che Idris A. (2022) Physics-informed neural network (PINN) evolution and beyond: A systematic literature review and bibliometric analysis. *Big Data and Cognitive Computing*, 6(4). — P. 140. (in Eng.).
- Li N., Li H., Duan K., and Tao W.-Q. (2025) Evaluation of the cooling effectiveness of air-cooled data centers by energy diagram. *Applied Energy*, 382. — P. 125215. <https://doi.org/https://doi.org/10.1016/j.apenergy.2024.125215> (in Eng.).

Lin J., Lin W., Huang H., Lin W., and Li K. (2023) Thermal modeling and thermal-aware energy saving methods for cloud data centers: A review. *IEEE Transactions on Sustainable Computing*, 9(3). — P. 571–590. (in Eng.).

Liu L., Liao X., Li Y., and Zhang S. (2026) Multi-Time Scale Optimization Scheduling of Data Center Considering Workload Shift and Refrigeration Regulation. *Energy Engineering*, 123(2). <https://doi.org/https://doi.org/10.32604/ee.2025.072631> (in Eng.).

LOI K., Song C., Chao H., Chen L., and Guo A. (2026) High temporal-resolution HVAC control in GPU-centric data centers via reinforcement learning and a knowledge-embedded physics-informed neural network. *Journal of Building Engineering*,. — P. 115499. <https://doi.org/https://doi.org/10.1016/j.jobe.2026.115499> (in Eng.).

Long J., Liu Y., and Ren Y. (2025) Semi-supervised medical image segmentation with dual-branch mixup-decoupling confidence training. *Medical Engineering & Physics*, 136. — P. 104285. <https://doi.org/https://doi.org/10.1016/j.medengphy.2025.104285> (in Eng.).

Lu N., Yao R., Wang Z., Yan Y., and Wang Y. (2026) Low-carbon and QoS-aware operation of data centers by AI task splitting and allocation. *Applied Energy*, 405. — P. 127132. <https://doi.org/https://doi.org/10.1016/j.apenergy.2025.127132> (in Eng.).

Ma S., Liu Y., Liu Y., Wang J., Fang Q., and Huang Y. (2025) Artificial intelligence-enabled predictive energy saving planning of liquid cooling system for data centers. *Advanced Engineering Informatics*, 65. — P. 103283. <https://doi.org/https://doi.org/10.1016/j.aei.2025.103283> (in Eng.).

Rasku T., Simson R., and Kiviluoma J. (2024) Sensitivity of a Lumped-Capacitance Building Thermal Modelling Approach for Energy-Market-Scale Flexibility Studies. *Buildings*. — P. 1614. <https://doi.org/10.3390/buildings14061614> (in Eng.).

Sun Z., Zhou Y., Gong Z., Wen C., Cai Z., and Zeng X. (2025) BiCA-LI: A Cross-Attention Multi-Task Deep Learning Model for Time Series Forecasting and Anomaly Detection in IDC Equipment. *Applied Sciences*. — P. 7168. <https://doi.org/10.3390/app15137168> (in Eng.).

Suresh P., Chang K.-W., Lua K.B., and Wang C.-C. (2026) Deep reinforcement learning for energy-efficient thermal management in 2U air-cooled server systems. *Applied Energy*, 404. — P. 127168. <https://doi.org/https://doi.org/10.1016/j.apenergy.2025.127168> (in Eng.).

Tian C., Liu Q., Tian C., Ma X., and Kong W. (2026) An innovative model integrating dual-channel graph attention network and multi-task learning for building energy consumption anomaly detection. *Building and Environment*, 287. — P. 113790. <https://doi.org/https://doi.org/10.1016/j.buildenv.2025.113790> (in Eng.).

Wang Y., Qiu X., Pei Q., Wang J., Zhang P., and Bai X. (2025) FAMAW-PINN: A physics-informed neural network integrating adaptive loss weighting with firefly-inspired adaptive point movement. *Journal of Computational Physics*, 542. — P. 114363. <https://doi.org/https://doi.org/10.1016/j.jcp.2025.114363> (in Eng.).

Yu X., Wang X., Kulevome D.K.B., Wang H., Khokhar S.-U.-D., and Li N. (2026) Multi-Scale temporal modeling via STL decomposition and dual-Transformer for resource prediction. *Future Generation Computer Systems*, 178. — P. 108323. <https://doi.org/https://doi.org/10.1016/j.future.2025.108323> (in Eng.).

Zhao J., and Li N. (2024) DFDAN: A Dual-Branch Feature Decoupling-Driven Auxiliary Network for Enhanced Fault Diagnosis Safety. *IEEE Transactions on Instrumentation and Measurement*, 74. — P. 1–10. (in Eng.).

Zheng J., Huang Y., and Yi N. (2026) IG-PINNs: Interface-gated physics-informed neural networks for solving elliptic interface problems. *Journal of Computational Physics*, 547. — P. 114540. <https://doi.org/https://doi.org/10.1016/j.jcp.2025.114540> (in Eng.).

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

ISSN 1991-346X

Volume 2.

Number 358 (2026). 449–469

<https://doi.org/10.32014/2026.2518-1726.448>

IRSTI 28.23.15

UDC 004.8:004.932

© Shynzhigit B.B.¹, Balabekova M.O.^{1*}, Amangeldy T.T.¹, Malik G.J.²,
Balgimbekova U.B.², 2026.

¹M. Auezov South-Kazakhstan Research University, Shymkent, Kazakhstan;

²Central Asian Innovation University, Shymkent, Kazakhstan.

E-mail: dadyiba@mail.ru

AUTOMATIC BRICK DEFECTS DETECTION BY USING A CNN-BASED DEEP LEARNING MODEL

Shynzhigit Bekmurat — doctoral student, M. Auezov South-Kazakhstan Research University, Shymkent, Kazakhstan,

E-mail: bekmuratsinzigit@gmail.com, <https://orcid.org/0009-0001-2970-9967>;

Balabekova Madina — Candidate of Technical Sciences, M. Auezov South-Kazakhstan Research University, Shymkent, Kazakhstan,

E-mail: dadyiba@mail.ru, <https://orcid.org/0000-0002-5558-3472>;

Amangeldy Togzhan — doctoral student, M. Auezov South-Kazakhstan Research University, Shymkent, Kazakhstan,

E-mail: togzhan.amangeldi1@mail.ru, <https://orcid.org/0009-0001-5081-9466>;

Malik Gulshat — Senior Lecturer, Central Asian Innovation University, Shymkent, Kazakhstan, E-mail: gulshat_malikova@bk.ru, <https://orcid.org/0009-0002-4519-6548>;

Balgimbekova Uldar — Senior Lecturer, Central Asian Innovation University, Shymkent, Kazakhstan,

E-mail: gulshat_malikova@bk.ru, <https://orcid.org/0009-0002-4519-6548>.

Abstract. The proposed research paper states that early and accurate detection of surface defects in the brick manufacturing industry is crucial for maintaining product quality, reducing material waste, and improving production efficiency. Traditional manual inspection methods are often time-consuming and rely on human expertise. Therefore, this study proposes an intelligent brick defect detection approach based on deep learning methods, especially convolutional neural networks (CNNs), for automated quality assessment. The proposed system is designed to detect and classify various types of brick defects, such as cracks, surface defects, edge defects, and non-defective samples, using image-based analysis. A publicly available brick defect dataset was used to train and evaluate the model. The CNN architecture consists of multiple convolutional and max pooling layers for feature extraction, followed by fully connected layers, and a Softmax classifier for final prediction. The model was trained and tested using supervised learning methods

to achieve accurate classification performance. Experimental results show that the proposed model achieved high performance on all evaluation metrics. The obtained Precision, Accuracy, Recall and F1 score values exceeded 93%, which indicates the effectiveness of the proposed approach. In particular, the model showed high performance in detecting crack defects and successfully distinguished between defective and non-defective brick surfaces. Comparative analysis also showed that deep learning-based approaches outperform traditional machine learning methods in terms of accuracy and automation capabilities. The results of this study demonstrate the potential of CNN-based intelligent systems for real-time industrial defect detection and automated quality control. The proposed approach can contribute to improving production reliability, reducing human intervention, and supporting the development of smart manufacturing environments in the building materials industry.

Key words: brick production, defect detection, automation, CNN, deep learning algorithms, computer vision, quality control

For citations: Shynzhigit B.B., Balabekova M.O., Amangeldy T.T., Malik G.J., Balgimbekova U.B. Automatic brick defects detection by using a CNN-based deep learning model. Academic Scientific Journal of Computer Science, 2026. — No.2. — P. 449–469. DOI <https://doi.org/10.32014/2026.2518-1726.448>

© Шынжігіт Ш.Б.¹, Балабекова М.О.^{1*}, Амангелді Т.Т.¹, Мәлік Г.Ж.²,
Балгимбекова У.Б.², 2026.

¹М. Әуезов атындағы Оңтүстік Қазақстан ғылыми-зерттеу университеті,
Шымкент, Қазақстан,

²Орталық Азия инновациялық университеті, Шымкент, Қазақстан,
E-mail: dadyiba@mail.ru

КІРПІШ АҚАУЛАРЫН АВТОМАТТЫ АНЫҚТАУДА CNN НЕГІЗІНДЕГІ ТЕРЕҢ ОҚЫТУ МОДЕЛІН ПАЙДАЛАНУ

Шынжігіт Бекмұрат — докторант, М. Әуезов атындағы Оңтүстік Қазақстан ғылыми-зерттеу университеті, Шымкент, Қазақстан,
E-mail: bekmuratsinzigit@gmail.com, <https://orcid.org/0009-0001-2970-9967>;

Балабекова Мадина — техника ғылымдарының кандидаты, М. Әуезов атындағы Оңтүстік Қазақстан ғылыми-зерттеу университеті, Шымкент, Қазақстан,
E-mail: dadyiba@mail.ru, <https://orcid.org/0000-0002-5558-3472>;

Амангелді Тоғжан — докторант, М. Әуезов атындағы Оңтүстік Қазақстан ғылыми-зерттеу университеті, Шымкент, Қазақстан,
E-mail: togzhan.amangeldi1@mail.ru, <https://orcid.org/0009-0001-5081-9466>;

Мәлік Гүлшат — аға оқытушы, Орталық Азия инновациялық университеті, Шымкент, Қазақстан,
E-mail: gulshat_malikova@bk.ru, <https://orcid.org/0009-0002-4519-6548>;

Балгимбекова Улдар — аға оқытушы, Орталық Азия инновациялық университеті, Шымкент, Қазақстан,
E-mail: gulshat_malikova@bk.ru, <https://orcid.org/0009-0002-4519-6548>.

Аннотация. Ұсынылып отырған ғылыми зерттеу жұмысында кірпіш өндірісі өнеркәсібінде беткі ақауларды ерте және дәл анықтау өнім сапасын сақтау, материал қалдықтарын азайту және өндіріс тиімділігін арттыру үшін өте маңызды екендігі баяндалған. Дәстүрлі қолмен тексеру әдістері көбінесе уақытты қажет етеді және адами тәжірибеге тәуелді. Сондықтан, бұл зерттеу терең оқыту әдістеріне, әсіресе автоматтандырылған сапаны бағалау үшін конволюциялық нейрондық желілерге (CNN) негізделген интеллектуалды кірпіш ақауларын анықтау тәсілін ұсынады. Ұсынылған жүйе кескінге негізделген талдауды қолдана отырып, жарықтар, беткі ақаулар, жиек ақаулары және ақаусыз үлгілер сияқты кірпіш ақауларының әртүрлі түрлерін анықтау және жіктеу үшін жасалған. Модельді оқыту және бағалау үшін жалпыға қолжетімді кірпіш ақаулары туралы деректер жиынтығы пайдаланылды. CNN архитектурасы ерекшеліктерді алу үшін бірнеше конволюциялық және максималды пул қабаттарынан, содан кейін толығымен қосылған қабаттардан және соңғы болжау үшін Softmax классификаторынан тұрады. Дәл жіктеу өнімділігіне қол жеткізу үшін модель бақыланатын оқыту әдістерін қолдана отырып оқытылды және сынақтан өтті. Эксперименттік нәтижелер ұсынылған модельдің барлық бағалау көрсеткіштері бойынша жоғары өнімділікке қол жеткізгенін көрсетеді. Алынған Precision, Accuracy, Recall және F1 балл мәндері 93%-дан асты, бұл ұсынылған тәсілдің тиімділігін көрсетеді. Атап айтқанда, модель жарықшақ ақауларын анықтауда жоғары өнімділік көрсетті және ақаулы және ақаусыз кірпіш беттерін сәтті ажыратты. Салыстырмалы талдау сонымен қатар терең оқытуға негізделген тәсілдер дәлдік пен автоматтандыру мүмкіндігі тұрғысынан дәстүрлі машиналық оқыту әдістерінен асып түсетінін көрсетті. Бұл зерттеудің нәтижелері CNN негізіндегі интеллектуалды жүйелердің нақты уақыт режимінде өнеркәсіптік ақауларды анықтау және автоматтандырылған сапаны бақылау үшін әлеуетін көрсетеді. Ұсынылған тәсіл өндіріс сенімділігін арттыруға, адамның араласуын азайтуға және құрылыс материалдары өнеркәсібінде ақылды өндірістік ортаны дамытуды қолдауға ықпал ете алады.

Түйін сөздер: кірпіш өндірісі, ақауларды анықтау, автоматтандыру, CNN, терең оқыту алгоритмдері, компьютерлік көру, сапаны бақылау

© Шынжыгит Ш.Б. ¹, Балабекова М.О. ^{1*}, Амангелды Т.Т. ¹,
Малик Г.Ж. ², Балгимбекова У.Б. ², 2026.

¹Южно-Казахстанский исследовательский университет имени
М.О. Ауэзова, Шымкент, Казахстан;

²Центрально-Азиатский инновационный университет, Шымкент, Казахстан.

*E-mail: dadyiba@mail.ru

**ИСПОЛЬЗОВАНИЕ МОДЕЛИ ГЛУБОКОГО ОБУЧЕНИЯ НА
ОСНОВЕ CNN ДЛЯ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ
ДЕФЕКТОВ КИРПИЧНОЙ КЛАДКИ**

Шынжыгит Бекмурат — докторант, Южно-Казахстанский исследовательский университет имени М.О. Ауэзова, Шымкент, Казахстан,

E-mail: bekmuratsinzigit@gmail.com, <https://orcid.org/0009-0001-2970-9967>;

Балабекова Мадина — кандидат технических наук, Южно-Казахстанский исследовательский университет имени М.О. Ауэзова, Шымкент, Казахстан,

E-mail: dadyiba@mail.ru, <https://orcid.org/0000-0002-5558-3472>;

Амангелды Тогжан — докторант, Южно-Казахстанский исследовательский университет имени М.О. Ауэзова, Шымкент, Казахстан,

E-mail: togzhan.amangeldi1@mail.ru, <https://orcid.org/0009-0001-5081-9466>;

Малик Гулшат — старший преподаватель, Центрально-Азиатский инновационный университет, Шымкент, Казахстан,

E-mail: gulshat_malikova@bk.ru, <https://orcid.org/0009-0002-4519-6548>;

Балгимбекова Улдар — старший преподаватель, Центрально-Азиатский инновационный университет, Шымкент, Казахстан,

E-mail: gulshat_malikova@bk.ru, <https://orcid.org/0009-0002-4519-6548>.

Аннотация. *Актуальность.* Раннее и точное обнаружение поверхностных дефектов в кирпичной промышленности имеет важное значение для поддержания качества продукции, сокращения отходов материалов и повышения эффективности производства. Традиционные методы ручной проверки требуют значительных временных затрат и зависят от квалификации специалиста, что ограничивает их применимость в условиях автоматизированного промышленного контроля. В связи с этим актуальным является применение методов глубокого обучения и компьютерного зрения для автоматического выявления дефектов кирпичной кладки и оценки качества строительных материалов. *Цель.* Разработать и оценить модель глубокого обучения на основе сверточной нейронной сети CNN для автоматического обнаружения и классификации дефектов кирпичной кладки по изображениям. *Методы.* В исследовании предложен интеллектуальный подход к обнаружению дефектов кирпича, основанный на применении сверточных нейронных сетей. Система предназначена для выявления и классификации различных типов дефектов, включая трещины, поверхностные повреждения, дефекты краев и образцы без дефектов. Для обучения и оценки модели использован общедоступный набор данных о дефектах кирпича. Архитектура CNN включает несколько сверточных слоев и слоев максимального пулинга для извлечения признаков, за которыми следуют полносвязанные слои и классификатор Softmax для итогового прогнозирования. Обучение и тестирование модели выполнялись с применением методов контролируемого обучения. *Результаты и выводы.* Экспериментальные результаты показали, что предложенная модель демонстрирует высокую производительность по основным оценочным метрикам. Значения точности, правильности, полноты и F1-меры превысили 93%, что подтверждает эффективность предложенного подхода. Модель особенно успешно выявляла дефекты в виде трещин и различала дефектные и недефектные поверхности кирпича. Сравнительный анализ показал, что подходы на основе глубокого обучения превосходят

традиционные методы машинного обучения по точности классификации и уровню автоматизации. Полученные результаты демонстрируют потенциал интеллектуальных систем на основе CNN для обнаружения дефектов в реальном времени и автоматизированного контроля качества в промышленности строительных материалов. Практическая значимость исследования заключается в возможности применения предложенного подхода для повышения надежности производства, сокращения ручного контроля и развития интеллектуальных производственных сред.

Ключевые слова: производство кирпича, обнаружение дефектов, кирпичная кладка, автоматизация, CNN, глубокое обучение, компьютерное зрение, контроль качества, Softmax, F1-мера.

Кіріспе. Кірпіш - қандайда бір ғимараттың тұрғызылуына қолданылатын маңызды құрылыс материалы болып табылады. Кірпіш тығыздығы мен беректілік сапасы, тусі, қолданылған шикізат құрамы, қолжетімді бағасы және әр түрлі ерекшеліктеріне байланысты құрылыстар мен инфрақұрылымдардың белгілі бір стандарттарға сай болуын қамтамасыз ететін маңыздылығы көп материал. Кірпіштің өндірілу процесін өндірістің тізбектелген жүйесі ретінде алғанда сапасы, консистенциясы және тиімділік жағын алып өзара қатысы бар маңызды бірнеше кезеңді сипаттауға болады. Өндіріс тұрғысын заманына сай дамуына байланысты материалдары мен түрлерінің өзгеруі ғана емес, дайындалу кезеңдерін автоматтандыру технологияларын қолданып басқару және басқа да оңтайлы процестерді қарастыру мүмкіндігі бар (Beskopylny, 2023).

Бірінші кезеңі - шикізаттың дайындалуы, бұл кезеңде саз бен минералдық қоспалардың бастапқы материалдарды біртектілік құрамда болу үшін өңделуі. Бөлшектерінің мөлшері және ылғалдылық сапасын біркелкі ұстауға ұсақтап дайындау, іріктеп кейін оны араластырып біріктіру сияқты жұмыстарды жүзеге асырады. Осы кезең барысында шикізатты біркелкі тұрақты деңгейге жеткізуге назар аудару қажет, қоспалардың өзіндік қасиеттерінің салдарынан кейінгі көзделген процесте немесе соңындағы өнімнің сапасы мен жарамдылығына ықпал етеді. Процесті автоматтандырудағы ғылыми зерттеу жұмыстарында шикізаттардың өзгергіштігі мен консистенциясы өнімді пайдалануы және сапа деңгейін арттыруда негізі себеп болады деп көрсетілген.

Екінші кезеңі – осы кезеңге дейін дайын материалдарды қалыптауға жабдықтарды қолданып стандартқа сай кірпіштәң геометриясын пішіндейді. Осы кезеңнің барысында өнімдердің негізгі қасиеттерінің сақталуын және өлшем сипаттамасын анықтап нақтылайды. Бұл процесті автоматтандыруда қалыптаудың дәлдігін күшейтіп, өндірістің санын біркелкі сақтап немес жүйелі түрде арттырады. Процестегі негізгі мақсат қалыптаудың параметрлерін бақылап, шикізат көлемін дәл енгізуде және дайын өнімнің сапасын бақылауда автоматтандыру маңызды рөл атқарады.

Үшінші кезеңі – бұл қалыптаудан кейінгі кезеңде кірпіш бақыланып ылғалдықты кептіруде болады. Осы процессте деформациясы мен жарықтары болмауын қадағалап термиялық өңдеуге жіберіледі. Өндірісті автоматтандыруда өнімнің кептірілу жағдайы, температурасы мен ылғалдылығы, ылғалдықтың бірте-бірте булануып реттеп тұрақтылығы қамтамасыз етіп, қол жұмысын араластыруды тоқтатады.

Төртінші кезеңі - кірпішті өндірудегі маңызды процестің бірі күйдіру болады. Осы процесті автоматтандыруда кірпішті 900°C-тан 1100°C-қа дейін температура аралығында куйдіріп, өнімді берік және төзімді екенін анықтап жүйелі түрде жұмыс атқаруын қамтамасыз етеді.

Сонғы кезеңі - өнімдерді дайындап өңдеу, күйдіріліп болған кірпішті салқындатып, тексеріп және сақтауға жөнелту кезеңі. Бұл процесті автоматтандырғанда кірпіш сапасын бақылап, беріктігін, өлшем дәлдігі мен сыртқы пішін тұтастығын тексетуді жолын реттеп, қателіктерді жояды. Сондай-ақ, өндірісте жұмыс күшін азайтып қауіпсіздікті арттырып қана қоймай, өндірістің өнімділігін арттырады (Wang, 2019).

Көрсетілген зертеулер мен тәжірибелік ұсыныстарды қолдай отырып кірпіш өндірудің процесін автоматтандыруда тиімділігін алсақ, өндіріс кезеңдерін ретімен біріктіріп, басқару жүйесін оңтайлы түрде іске асырып, қателіктерді болдырмай және өнімнің сапасы мен өндірістің көзделген шарттарын толыққанды қамтамасыз етеді.

Әдеби шолу. Өндіріс процесін автоматтандыру барысында тиімділік артады сондай-ақ, бұл өндірісте бұрын сонды болмаған құрылғылармен жұмыс жүргізуде пайдалануының қиындықтары және экономикалық келеңсіздіктер болуы мүмкін.

а) Инвестицияның жоғарылығы

Кірпіш өндірісімен айналысатын автоматтандарылған зауыттарды жүйелі түрде іске қосу ауқымды соммадағы жаңа жабдықтар мен арнайы басқаратын құрылғыларды және құралдарды орнатуды талап етеді. Бұл өндірісте бастапқы шығынның көптігінен шағын бизнестер қаражаттың кемшілігінен құлдырау қаупі бар.

б) Техникалардың күрделілігі мен оларға қызмет көрсету

Өндіріс процесінің өзара байланысының күрделілігінен өндірістің қаржы шығынына әкелетін тоқтауларсыз бір қалыпты жұмыс істеуін қамтамасыз ету үшін тұрақты түрде арнайы мамандардың қызмет көрсетуін қажет етеді. Зауыттағы электр қуатының тұрақты болуы ең маңызды және бастапқы шарт, электр қуатының ауыспалылығы техникалардың жұмыс істеуіне әсері көп және электр шығыны ұлғайады. Сондай-ақ, өнімнің дизайнын тек арнайы мамандарға жүгіну арқылы ғана өзгертіп қосалқы бөлшектерінің шығыны болады.

с) Техникалық ақаулар мен істен шығуы

Автоматтандырылған жүйедегі ретімен жұмыс атқаратын жабдықтардың біреуі бұзылса оны қайта іске қосқанға дейін қалған кезеңдері тоқтату қаупі бар (Kataev, 2025).

Кірпіш өндірісін автоматтандырғанда тұрақты және тиімді, бірақ техниканың күрделілігіне байланысты шығынның көптігі мен тұрақты имараттарға тәуелді болу сияқты проблемалар болады.

Кірпіш өндірісін қол күшімен басқарудағы кемшіліктер. Өндірісте қол күшіне жүгіну шығынды азайту мақсатында әдеттегі жағдай болып қалыптасқан, алайда бұл әдісте автоматтандыру жүйесіне қарағанда бірнеше өзіндік кемшіліктері бар.

1. Өнімділіктің төмендігі. Өндіріс процесінің баяулығы, жұмыс уақытының шектеулігі кндірісті шектейді

2. Сапаның сай келмеуі. Қолдан жасалған өнімдер адамдардың өнірістегі біліктілігіне сай түрлі өзгешеліктері мен өзгерістері болуы мүмкін, сондықтан сапа сәйкестігі төмен болады.

3. Жұмысшы күштеріне тәуелділігі және көрінбейтін шығындар. Өндірісте жұмысшылар санының аздығы, олардың жұмыс барысында шаршауы, тамақтану мен табиғи қажеттілік уақыттары тиімділікке кері әсерін тигізу мүмкін және шикізаттарды араластыру мен өңдеу істерінде дәлдікті сақтау қиындығынан материал шығыны артады

4. Жұмысшылардың денсаулығы мен қауіпсіздігі. Жұмысшылардың шаңда, зиян қоспалардың ортасында, ыстық цехтардың ішінде жүруі және ауыр салмақты материалдарды көтеруі жұмысшылардың денсаулығы мен қауіпсіздігіне кепіл бермейді.

5. Процесті басқарудағы қиыншылықтар. Өндіріс кезеңдерінде температура, ылғалдылық және кептіру уақытының бірқалыптылығын сақтау қиындығы өнім сапасын төмендету мүмкін.

6. Өнімнің саны мен сапасы туралы деректердің дәл болмауы. Өндірісті қолмен басқару өнімділікті талдауда, саны мен сапасы жөнінде нақты ақпарат болмайды және әр кезең процестеріндегі қателіктер мен кемшіліктерді бақылап оларды жоюға деген жұмыстың қиындығы болады

Кірпіш өндірісі автоматтандыру жүйесіне қарағанда қолмен басқарған оңай, әрі бастауы аз шығынды талап етеді, бірақ белгілі уақыт аралығын есептегенде өнімнің саны мен сапасының төмендігі жағынан тиімсіз. Өндірісті автоматтандыру жүйесіне көшіру өнімнің жалпы саны мен сапасының арттыруына ықпал етеді және өндірістің қауіпсіздігі мен ыңғайлылығын қамтамасыз етеді (Dhali, 2025).

Автоматтандырылған кірпіш өндірісіндегі жасанды интеллект ақауларын анықтау алгоритмдері. Автоматтандырылған кірпіш өндірісінің реттілік кезеңдерінде болған ақаулар мен кемшілік орындарын сенсорлық деректер мен машиналық оқыту жүйелерін қолданып тез табу үшін жасанды интеллект алгоритмдері процессте негізгі қызмет атқарады. AS-IS моделі (қазіргі жағыдайда) - кірпішті өндіру процесінің қазіргі кезде болып жатқан жағдайын көрсетеді, автоматтандыру немесе оңтайландыру әдістерін енгізгенге дейін өндірістік әрекеттердің іс жүзінде қалай орындалатынын сипаттау жүйесі (Cumbajin, 2023). Бұл модель қолданыстағы жұмыс процестерін жан-жақты

түсінуге мүмкіндік бреді және жүйедегі тиімсіздіктерді, кедергілерді және өзгергіштік көздерін анықтауға негіз болады. пҚазіргі өндірістік ортада процесс шикізатты дайындаудан, материалды тасымалдаудан, қалыптаудан, кептіруден, күйдіруден және соңғы өнімді өңдеуден тұратын тізбекті құрылымды ұстанады (Cumbajin, 2023). Бұл кезеңдер негізінен қолмен немесе жартылай автоматтандырылған операциялар арқылы орындалады, олардың арасындағы интеграция мен үйлестіру шектеулі болып келеді.

Кесте 1 — жасанды интеллект әдістерінің кірпіштің ақауларын анықтаудағы ғылыми зерттеу жұмыстарына шолу.

Автор-лар	Санаттары	Әдіс Алгоритм	Сипаттама	Артықшылықтары	Шектеулер
Zhou, 2025	CNN негізіндегі әдістер	CNN	Кірпіш суреттерін жарықтарға, деформацияларға, беткі ақауларға талдау	Жоғары дәлдік, нақты уақыт режимінде анықтау	Үлкен белгіленген деректер жиынтығын қажет етеді
Karimi, 2025	Кескінді өңдеу + жасанды интеллект	Computer vision	Жарылған жерлерді, сегменттердің ақауларын, жарылғаннан соңғы пішінді анықтайды	Жақсырақ мүмкіндіктерді алу	Жарықтандыру жағдайларына сезімтал
Nazir, 2020	Классикалық ML	SVM, Random Forest, KNN	Кірпіштерді ақаулы немесе ақаусыз деп жіктейді	Қарапайым іске асыру	Терең оқытуға қарағанда дәлдігі төмен
Ji, 2021	Нысанды анықтау	YOLO, R-CNN, SSD	Суреттердегі ақауларды анықтайды және локализациялайды	Жылдам анықтау (YOLO), дәл локализация	Есептеу құны жоғары
Salman, 2025	Аномалияны анықтау	Autocoder, GAN	Бақыланатын/жартылай бақылаулы оқытуды қолдана отырып, ауытқуларды анықтайды	Шектеулі белгіленген деректермен жұмыс істейді	Күрделі оқыту процесі
Şermet, 2024	Сенсорға негізделген жасанды интеллект	IOT+ жасанды интеллект	Температураны, дірілді, ылғалдылықты бақылайды	Нақты уақыт режимінде бақылау	Қосымша жабдықты қажет етеді

Katsi-giannis, 2025	Шешім қабылдау жүйелері	Автомат-тандырылған жүйелер	Ақаулы кірпіштерді алып тастайды, параметрлерді реттейді	Адамның араласуын азайтады	Жүйелік интеграцияның күрделілігі
Krishnan, 2025	Болжамдық аналитика	Болжау модельдері	Ақауларды болжайды және өндірісті оңтайландырады	Тиімділікті жақсартады	Тарихи деректер қажет
Çavuş, 2025	Артықшылықтары	-	Жоғары жылдамдық, дәлдік, автоматтандыру	Адамның қателігін азайтады	-
(Cum-bajin, 2023	Қиындықтар	-	Үлкен деректер жиынтығы, құны, жарықтандыруға тәуелділік	-	Іске асырудың күрделілігі

Процесті модельдеу тұрғысынан ағымдағы жұмыс процесін бизнес-процестің моделі мен нотациясын (BPMN) пайдаланып көрсетуге болады, мұнда әрбір кезең сапа мәселелерінен туындайтын кейде кері байланыс циклдарымен сызықтық тізбекте байланысты. Процесс ішіндегі шешім қабылдау көбінесе автоматтандырылған жүйелерге емес, адам операторларына байланысты, бұл орындауда сәйкессіздіктерді тудырады.

Мысалы ретінде алатын болсақ, кептіруден немесе күйдіруден кейін анықталған ақауы бар өнімдер қайта өңдеуді қажет етуі мүмкін, бұл қосымша кідірістерге және ресурстарды пайдаланудың артуына әкелетіні байқалды.

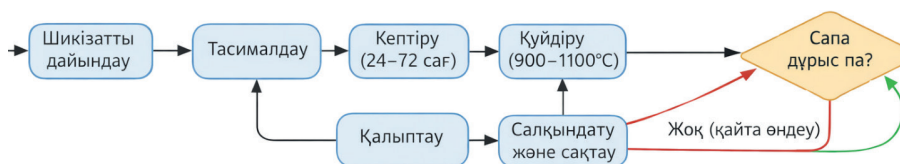
AS-IS процесінің уақыттық талдауы өндірістің әртүрлі кезеңдерінде айтарлықтай өзгергіштікті көрсетеді. Шикізатты дайындау және қалыптау әдетте салыстырмалы түрде қысқа мерзімді қажет етеді, ал кептіру және күйдіру айтарлықтай көп уақытты алады, көбінесе қоршаған орта жағдайлары мен пайдалану тәжірибесімен байланысты бірнеше сағаттан бірнеше күнге дейін созылады. Бұл кезеңдер өндіріс циклінде негізгі кедергілер болып табылады, жалпы өнімділікті шектейді (Krishnan, 2025). Сонымен қатар, кезеңдер арасында қолмен өңдеу қосымша кідірістерге әкеледі және процестің тиімділігін төмендетеді. Өндірісті автоматтандыру зерттеулерінде атап өтілгендей, өңдеу уақытының өзгергіштігі қолмен жасалатын операциялармен және өндіріс кезеңдерінің арасындағы синхрондаудың болмауымен тығыз байланысты, бұл жалпы жүйенің жұмысына кері әсер етеді.

Зерттеу әдістемесі. Ғылыми зерттеу жұмысында зерттеліп жатқан процесстерді егжей-тегжейлі талдау бірнеше маңызды мәселелерді көрсетеді.

Негізгі мәселелердің бірі - операторға тәуелді әрекеттерден туындаған жоғары процестің өзгергіштігі, бұл өндіріс уақытына да, өнім сапасына да

әсер етеді. Бір жерге орталықтандырылған басқарылатын жүйенің болмауы әртүрлі кезеңдер арасындағы нашар орындалатын интеграцияға әкеледі, бұл істі жүргізу барысындағы тиімсіздікке және үйлестірудің болмауына әкеліп соғады.

Сонымен қатар, кептіру және күйдіру кезеңдері ұзақ өңдеу уақыты кезеңінде және қоршаған ортадағы жағдайларға сезімталдығына байланысты айтарлықтай кегдергілерді де тудырады. Өнімнің сапасын бақылау да шектеуге әкеледі (Сурет 1), себебі тексеру процедуралары көбінесе қолмен орындалады, бұл адами қателіктер мен сәйкессіз бағалау ықтималдығын арттырады. Сондай – ақ, кептіру және күйдіру кезінде энергия тұтыну оңтайлы басқарылмайды, бұл пайдалану шығындарының артуына әкеледі.

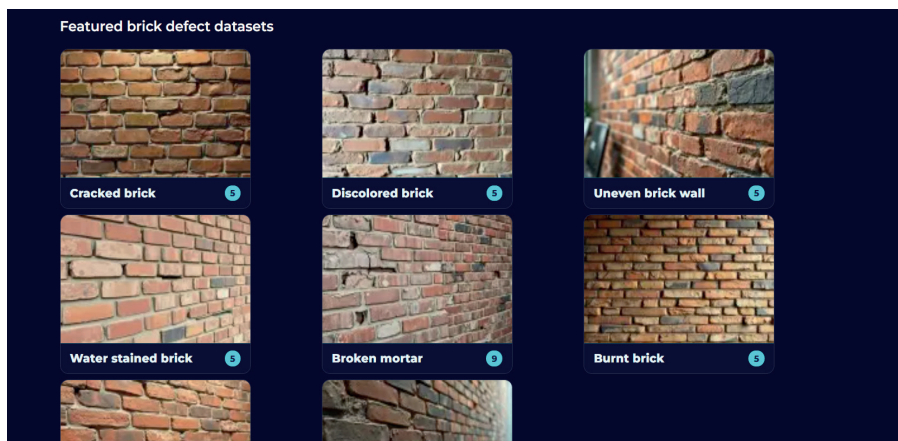


Сурет 1 — Кірпіш өндіріу процесі және сапаны тексеру процесінің жұмыс процесі

Тағы бір маңызды мәселе - жұмыс орнының қауіпсіздігі, себебі жоғары температуралы ортада қолмен жұмыс істеу жұмысшыларды ықтимал қауіптерге ұшыратады. Автоматтандыру саласындағы зерттеулер мұндай жағдайларда адамның қатысуын азайту қауіпсіздік пен пайдалану сенімділігін айтарлықтай жақсарту алатынын атап көрсетеді.

1. Деректер қоры

Бұл зерттеуде кірпіш өнімдеріндегі ақауларды автоматтандырылған түрде анықтау үшін <https://images.cv/category/brick-defect> платформасынан алынған ашық бастапқы кодты деректер жиынтығы пайдаланылды.



Сурет 2 — санаттарға бөлінген кірпіш ақаулары бар деректер жинағы

Деректер жиынтығы компьютерлік көру қолданбалары үшін арнайы жасалған және кірпіш материалдарында байқалған әртүрлі беттік ақауларды бейнелейтін әртүрлі кескіндер жиынтығынан тұрады. Бұл кескіндер жарықтар, жиектердің зақымдануы, беттік кедір-бұдырлар және басқа да құрылымдық ауытқулар сияқты әртүрлі ақау үлгілерін, сондай-ақ анықтамалық деректер ретінде пайдаланылатын ақаусыз үлгілерді көрсетеді.

Ақаулы және ақаусыз үлгілердің қолжетімділігі жіктеу, анықтау және сапаны бағалау тапсырмалары үшін машиналық оқыту модельдерін әзірлеуге және бағалауға мүмкіндік береді. Деректер жиынтығының құрылымы және оған сәйкес ақау санаттары Кесте 2- де қорытындыланған.

Кесте 2 — Кірпіш ақаулары бар деректер қорының жиынтығы

№	Категория атауы	Сипаттамасы	Ақау түрі	Машиналық оқытудағы рөлі
1	Crack (Жарық)	Кірпіш бетінде пайда болатын сызықтық немесе терең жарықтар	Құрылымдық ақау	Негізгі класс, defect detection үшін маңызды
2	Broken edges (Сынған бұрыштар)	Кірпіштің шеттері мен бұрыштарының сынуы	Механикалық зақым	Classification және detection үшін
3	Surface defects (Беткі ақаулар)	Дақтар, түс өзгеруі, қабыршақтану	Визуалды ақау	Image classification үшін қолданылады
4	Deformation (Деформация)	Кірпіш формасының өзгеруі, қисайуы	Геометриялық ақау	Shape analysis және detection
5	Porosity / Internal defects (Ішкі ақаулар)	Кеуектілік немесе ішкі құрылым бұзылыстары	Жасырын ақау	Advanced ML модельдерінде
6	Normal (Ақаусыз)	Сапалы, стандартқа сай кірпіш	Ақау жоқ	Базалық класс (baseline)

Сонымен қатар, деректер жиынтығы ұсынылған модельдердің жалпылау мүмкіндігін арттыратын өндірістік жағдайларды нақты түрде көрсетеді.

2. Кірпіш ақауларын CNN арқылы анықтау

Ұсынылып отырғанғылыми зерттеуде кірпіш кескіндеріндегі ақауларды автоматтандырылған түрде анықтау үшін конволюциялық нейрондық желіге (CNN) негізделген тәсіл қолданылды. Ұсынылған модель деректерді алдын ала өңдеуден, ерекшеліктерді алудан, модельді оқытудан және бағалау кезеңдерінен тұрады.

а) мәліметтер қорын дайындау кезеңі:

Деректер қорындағы барлық кескіндердің өлшемі $H \times W$ бекітілген өлшемге өзгертілді және $[0,1]$ диапазонына дейін қалыпқа келтірілді:

$$x' = \frac{x}{255} \quad (1)$$

мұндағы x бастапқы пиксель мәнін, ал x' - қалыпқа келтірілген мән.

Кіріс деректегі кескіндер 8 биттік форматта ұсынылғандықтан, пиксель

қарқындылығының мәндері 0-ден 255-ке дейін болады. Сондықтан, нормалау пиксель мәндерін 255-ке бөлу арқылы $[0,1]$ диапазонына дейін масштабтау арқылы орындалады. Жалпылауды жақсарту үшін деректерді көбейту әдістері, мысалы, айналдыру, аудару және масштабтау қолданылды.

б) *Конволюциялық нейрондық желі CNN*

CNN моделі бірнеше конволюциялық қабаттардан, пулдық қабаттардан және толығымен байланысты қабаттардан тұрады. Конволюция операциясы келесідей анықталады:

$$y(i, j) = \sum_m \sum_n x(i + m, j + n) * w(m, n) + b \quad (2)$$

мұнда:

x — кіріс кескіні;

w — конволюция ядросы;

b — ығысу;

$y(i, j)$ — шығыс функциясының картасы;

Активация ReLU функциясын пайдаланып қолданылады:

$$f(x) = \max(0, x) \quad (3)$$

Пулдеу кеңістіктік өлшемдерді азайту үшін қолданылады:

$$y = \max(x) \quad (4)$$

с) *Fully connected layer және классификация*

Функцияларды шығарғаннан кейін тегістеу мүмкіндік карталары толығымен қосылған қабаттарға беріледі. Соңғы жіктеу Softmax функциясы арқылы орындалады:

$$L = - \sum_{i=1}^n y_i \log(\widehat{y}_i) \quad (5)$$

y_i — шынайы белгі

$\widehat{y}_i$ — болжамды белгі

д) *Оңтайландыру*

Модель параметрлері Adam оптимизаторы арқылы оңтайландырылған:

$$\theta = \theta - \alpha \nabla l(\theta) \quad (6)$$

Мұндағы:

θ — моделдің параметрлері

α — оқу бағамы

Бұл кедергілер тиімділікті төмендетеді, өндіріс уақытын жоғарылатады және өнім сапасына кері әсер етеді, осылайша процесті оңтайландыру және төменде ұсынатын ТО-ВЕ (автоматтандыру процессі) моделінде автоматтандырылған шешімдерді енгізу қажеттілігін көрсетеді.

е) *Ұсынылған модельдің дәлдігін тексеру*

Алгоритмдерді бағалау үшін ең жиі қолданылатын көрсеткіштер:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Дәлдік (Precision): Дәлдік - шынайы оң нәтижелердің (TP) болжамды оң нәтижелердің жалпы санына қатынасы, оған шынайы оң нәтижелер (TP) және жалған оң нәтижелер (FP) кіреді. Бұл модель оң деп жіктеген нысандардың қаншасы шын мәнінде оң нәтиже беретінін көрсетеді. Дәлдік жалған оң нәтижелер қымбат немесе жағымсыз болған сценарийлерде маңызды.

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

Еске түсіру (Recall): Еске түсіру - бұл шынайы оң нәтижелердің (TP) шынайы оң нәтижелердің жалпы санына қатынасы, оған шынайы оң нәтижелер (TP) және жалған теріс нәтижелер (FN) кіреді. Ол модельдің деректер жиынындағы барлық тиісті жағдайларды дұрыс анықтау мүмкіндігін өлшейді. Еске түсіру әсіресе медициналық диагностика сияқты жалған теріс нәтижелер маңыздырақ болатын жағдайларда маңызды.

$$F1 = \frac{2 \text{ precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (9)$$

F1 ұпайы: F1 ұпайы дәлдік пен еске түсірудің гармоникалық орташасы болып табылады, бұл екеуінің арасында теңгерімді өлшемді қамтамасыз етеді. Бұл әсіресе дәлдік пен еске түсіруді теңестіру қажет болған кезде пайдалы және көбінесе сыныптың таралуы теңгерімсіз болған кезде қолданылады.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (10)$$

Дәлдік көрсеткіші (Accuracy). - бұл дұрыс болжанған жағдайлардың (шынайы оң және шынайы теріс) деректер жиынындағы жағдайлардың жалпы санына қатынасы. Дәлдікті түсіну оңай болғанымен, теріс мысалдардың саны оң мысалдардан әлдеқайда көп болатын теңгерімсіз деректер жиынтықтарында ол жаңылыстыруы мүмкін.

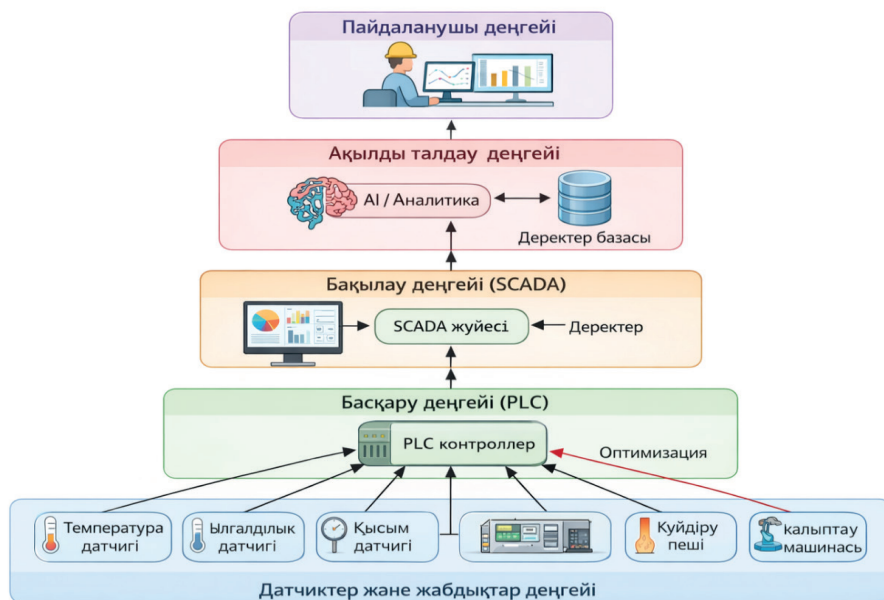
AS-IS моделі (қазіргі жағыдайда) қазіргі кірпіш өндірісі процесі - процестің өзгергіштігімен және өндіріс кезеңдерінде интеграцияның жеткіліксіздігімен сипатталатынын көрсетеді.

TO-BE моделі (автоматтандыру процессі). TO-BE (автоматтандыру процессі) моделінде қазіргі AS-IS моделінде анықталған шектеулерді еңсеру үшін әзірленген кірпіш өндіру процесінің оңтайландырылған және автоматтандырылған нұсқасын білдіреді. Бұл модель өндіріс тиімділігін арттыруға, процестердің өзгермелілігін азайтуға, өнім сапасын арттыруға және заманауи автоматтандыру технологияларын біріктіру арқылы қауіпсіз және сенімді жұмыс жағдайларын қамтамасыз етуге бағытталған.

Жақсартылған процесте шикізатты дайындаудың бастапқы кезеңі материал құрамын, бөлшектердің мөлшерін бөлуді және ылғалдылықты реттейтін датчиктермен жабдықталған автоматтандырылған жүйелерді қолдану арқылы орындалады. Бұл дәйекті және біртекті қоспаны қамтамасыз етеді, әдетте қолмен өңдеуден туындайтын өзгермелілікті айтарлықтай азайтады. Қалыптаудың келесі кезеңі кірпіштің пішіні мен өлшемдерін дәл бақылауға мүмкіндік беретін бағдарламаланатын қалыптау жабдығы арқылы жүзеге асырылады, осылайша өнімнің біркелкілігін жақсартады және ақауларды азайтады.

Кептіру кезеңі температура мен ылғалдылық үздіксіз бақыланатын және автоматты түрде реттелетін бақыланатын кептіру камераларын енгізу арқылы жақсартады. Бұл бұзылу немесе деформация сияқты құрылымдық ақауларды болдырмай, ылғалды бірте-бірте және біркелкі кетіруге мүмкіндік береді. Күйдіру кезеңі алдын ала анықталған температура профилдері мен сенсорлардан кері байланыс негізінде жұмыс істейтін, тұрақты термиялық өңдеуді және жақсартылған энергия тиімділігін қамтамасыз ететін ақылды пеш жүйелері одан әрі оңтайландырылған. Мұндай автоматтандырылған басқару әсіресе жоғары температурадағы өнеркәсіптік процестерде өте маңызды, тіпті шамалы ауытқулар өнім сапасына айтарлықтай әсер етуі мүмкін.

Соңғы кезеңде дайын өнімді тексеру, сұрыптау және сақтау үшін автоматтандырылған өңдеу және сапаны бақылау жүйелері енгізіледі. Жетілдірілген зондтау технологиялары ақауларды нақты уақытта анықтауға мүмкіндік береді, сәйкес келмейтін өнімдерді автоматты түрде жоюға немесе қайта өңдеуге қайта бағыттауға мүмкіндік береді. Бұл қолмен тексеруге тәуелділікті айтарлықтай азайтады және өндіріс жүйесінің жалпы сенімділігін арттырады. Оған қоса, барлық кезеңдерді бірыңғай басқару жүйесіне біріктіру операцияларды синхрондауға, бос тұру уақытын қысқартуға және өткізу қабілеттілігін арттыруға мүмкіндік береді.



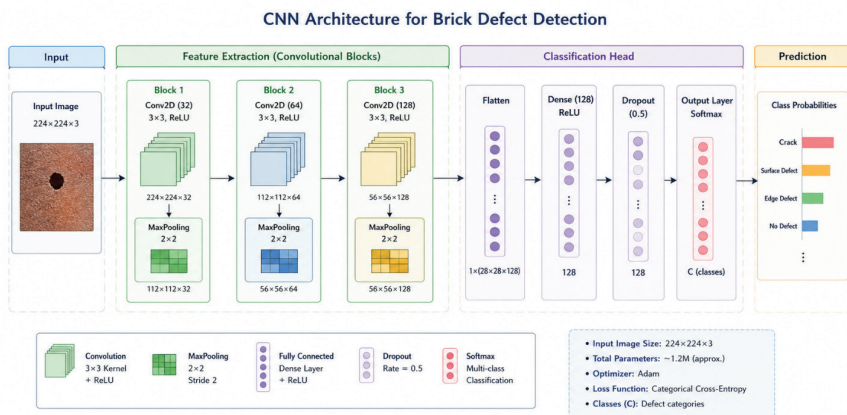
Сурет 3 — Ақылды автоматтандырылған кірпіш өндірісі жүйесінің архитектурасы

Дегенмен, ТО-ВЕ (*автоматтандыру процесі*) модельерлі дәстүрлі кірпіш өндіру процесін деректерге негізделген және кірпіш өндіріс жүйесіне айналдырады. Автоматтандыруды енгізу операциялық тиімділік пен өнімнің тұрақтылығын жақсартып қана қоймайды, сонымен қатар қауіпті орталарға адамның қатысуын азайту арқылы энергия тұтынуды азайтады және жұмыс орнындағы қауіпсіздікті арттырады. Нәтижесінде ұсынылған модель кірпіш өндірісін жаңғырту және өнеркәсіптік көрсеткіштердің жоғары деңгейіне жету үшін тұрақты және ауқымды шешімді ұсынады.

Зерттеу нәтижесі. Бұл бөлімде ұсынылған CNN негізіндегі кірпіш ақауларын анықтау моделінен алынған эксперименттік нәтижелер ұсынылған. Модельдің өнімділігі дәлдік, дәлдік, еске түсіру және F1 ұпайын қоса алғанда, стандартты жіктеу көрсеткіштерін қолдана отырып бағаланды. Сонымен қатар, ұсынылған тәсілдің әртүрлі кірпіш бетінің ақауларын анықтаудағы тиімділігін бағалау үшін деректер жиынтығы мен модель архитектурасын визуалды талдау жүргізілді. Ұсынылған конволюциялық нейрондық желі ақаулы және ақаусыз кірпіш үлгілерін қамтитын белгіленген кірпіш ақаулары деректер жиынтығын пайдаланып оқытылды. Модель жарықтар, беткі зақымдану, жиек ақаулары және ақаусыз кірпіштер сияқты ақау санаттарын автоматты түрде анықтау үшін жасалған.

Эксперименттік нәтижелер ұсынылған модельдің барлық бағаланған санаттар бойынша жоғары жіктеу көрсеткіштеріне қол жеткізгенін көрсетеді. Модель кірпіш кескіндерінен маңызды визуалды ерекшеліктерді тиімді түрде алып, ақаулы үлгілерді қалыпты кескіндерден сәтті ажыратты. Сонымен

қатар, алынған нәтижелер кірпіш өндірісі ортасында автоматтандырылған сапаны тексеру үшін терең оқыту әдістерінің жарамдылығын растайды.



Сурет 4 — Кірпіш ақауларын анықтауға арналған конволюциялық нейрондық желі архитектурасы

Ұсынылған конволюционды нейрондық желі (CNN) архитектурасы кірпіш беттеріндегі ақауларды автоматты түрде анықтауға және жіктеуге арналған. Модель $224 \times 224 \times 3$ өлшеміндегі кіріс кескіндерін өңдейді және үш негізгі кезеңнен тұрады: мүмкіндікті алу, жіктеу және болжау.

Функцияларды шығару кезеңінде кіріс кескіннің иерархиялық көріністерін біртіндеп үйрену үшін үш дәйекті конволюционды блоктар қолданылады. Әрбір блокта 3×3 ядросы бар конволюциялық қабат, одан кейін түзетілген сызықтық бірлік (ReLU) белсендіру функциясы бар. Бірінші блок 32 мүмкіндік картасын шығарады, екіншісі 64-ке дейін, ал үшіншісі 128 мүмкіндік картасына дейін кеңейтеді. Әрбір конволюциялық қабат 2 қадаммен максимум 2×2 біріктіру операциясымен жалғасады, бұл ең сәйкес мүмкіндіктерді сақтай отырып, кеңістік өлшемдерін азайтады. Бұл иерархиялық құрылым модельге төмен деңгейлі үлгілерді (мысалы, жиектер мен текстуралар) және ақаулармен байланысты жоғары деңгейлі семантикалық мүмкіндіктерді түсіруге мүмкіндік береді.

Алынған мүмкіндік карталары содан кейін жіктеу басына беріледі, онда олар бір өлшемді векторға тегістеледі. 128 нейроны бар толық қосылған (тығыз) қабат және ReLU белсендіру күрделі мүмкіндіктердің өзара әрекеттесуін үйрену үшін қолданылады. Шамадан тыс орнатуды азайту және жалпылауды жақсарту үшін 0,5 жылдамдығымен түсіру қабаты енгізіледі. Соңында, шығыс қабаты мақсатты сыныптар бойынша ықтималдық үлестірімдерін жасау үшін Softmax белсендіру функциясын пайдаланады.

Болжау сатысында үлгі алдын ала анықталған ақау санаттарына сәйкес келетін сынып ықтималдықтарын шығарады, соның ішінде Crack, Surface

Defect, Edge Defect және No Defect. Соңғы болжам ретінде ең жоғары ықтималдығы бар сынып таңдалады.

Модель Adam оптимизаторы және категориялық кросс-энтропия жоғалту функциясы арқылы оқытылады. Параметрлердің жалпы саны шамамен 1,2 миллионды құрайды, бұл модель күрделілігі мен есептеу тиімділігі арасындағы теңгерімді қамтамасыз етеді.

Тұтастай алғанда, бұл архитектура нақты әлем сценарийлерінде кірпіш қауаларын дәл және сенімді анықтауға мүмкіндік беретін сенімді жіктеу механизмімен конвульсиялық мүмкіндіктерді шығаруды тиімді біріктіреді.

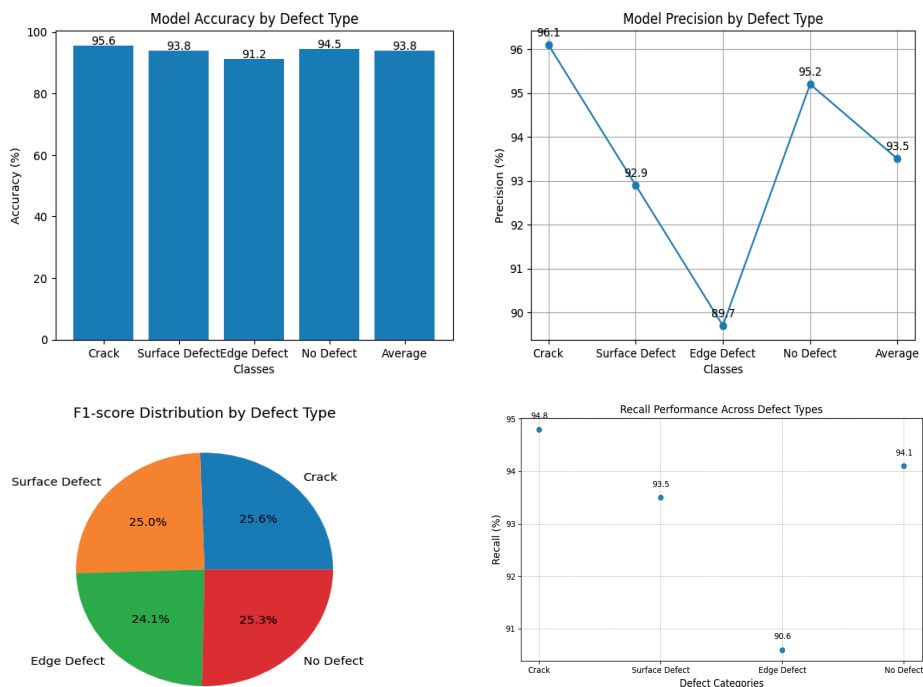
Ұсынылған үлгінің өнімділігін бағалау. Ұсынылған CNN негізіндегі модельдің өнімділігі төрт қау категориясы бойынша дәлдік, дәлдік, еске түсіру және F1 ұпайын қоса алғанда стандартты жіктеу метрикасын пайдалана отырып бағаланды: Жарық, бет қауы, жиек қауы және қау жоқ.

Кесте 3 — ұсынылған модель өнімділігі

Metric	Crack	Surface Defect	Edge Defect	No Defect	Average
Accuracy (%)	95.6	93.8	91.2	94.5	93.8
Precision (%)	96.1	92.9	89.7	95.2	93.5
Recall (%)	94.8	93.5	90.6	94.1	93.2
F1-score (%)	95.4	93.2	90.1	94.6	93.3

Кестеде келтірілген нәтижелерге сәйкес модель барлық санаттар бойынша тұрақты жоғары өнімділікті көрсетеді. Ең жоғары дәлдікке (95,6%) Crack сыныбына қол жеткізілді, бұл модельдің көрнекті құрылымдық қауаларды тиімді анықтайтынын көрсетеді. Сол сияқты, No Defect класы да жақсы өнімділікті (94,5%) көрсетеді, бұл модельдің қауасыз үлгілерді дұрыс анықтау мүмкіндігін растайды. Дегенмен, Edge Defect класы үшін (91,2%) сәл төмен дәлдік байқалады, бұл бұл санаттың нәзік функциялардың вариацияларына байланысты салыстырмалы түрде қиынырақ екенін көрсетеді. Дәлдік мәндері модельдің болжаудағы сенімділігін одан әрі көрсетеді. Ең жоғары дәлдік Crack үшін алынған (96,1%), бұл класс үшін төмен жалған оң көрсеткішті көрсетеді. Керісінше, Edge Defect сыныбы ең төменгі дәлдікті (89,7%) көрсетеді, бұл басқа қау түрлерімен кездейсоқ қате жіктелуді білдіреді.

Ескерту мәндері модельдің нақты қауаларды анықтаудағы сезімталдығын көрсетеді. Модель барлық сыныптар бойынша күшті еске түсіруге қол жеткізеді, Crack үшін ең жоғары мән (94,8%). Edge Defect (90,6%) үшін салыстырмалы түрде төменірек еске түсіру осы санаттағы кейбір қау даналарының толық түсірілмеуі мүмкін екенін көрсетеді. Дәлдік пен еске түсіруді теңестіретін F1 ұпайы модельдің жалпы беріктігін растайды. Crack сыныбы қайтадан ең жоғары F1 ұпайына қол жеткізеді (95,4%), ал Edge Defect ең төменгі көрсеткішті (90,1%) жазады. Осыған қарамастан, барлық F1 ұпайлары тұрақты және сенімді өнімділікті көрсете отырып, 90% жоғары болып қалады.



Сурет 5 — Ұсынылған үлгінің өнімділігін салыстыру

Суретте ұсынылған үлгінің төрт ақау категориясы бойынша салыстырмалы өнімділігі көрсетілген: Жарық, бет ақауы, жиек ақауы және ақау жоқ, төрт бағалау метрикасы — дәлдік, дәлдік, еске түсіру және F1 ұпайы қолданылады.

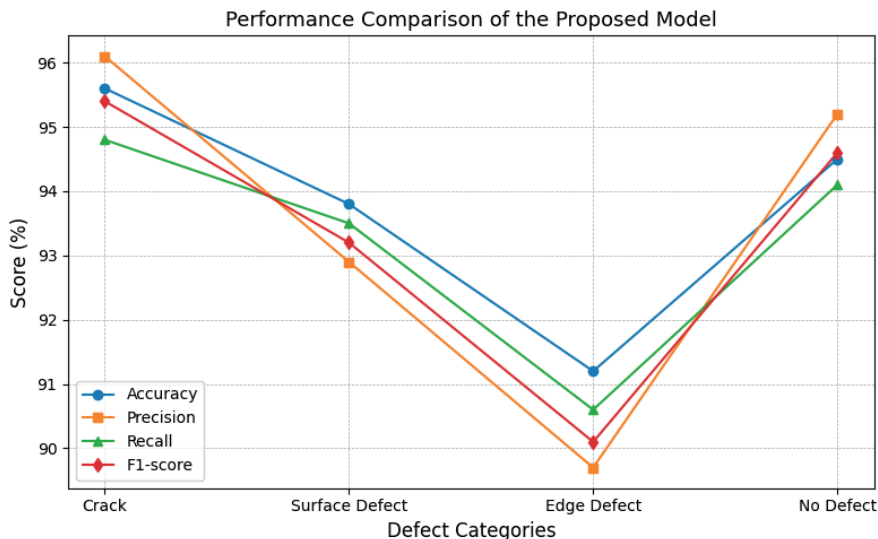
Жалпы алғанда, үлгі барлық санаттар бойынша дәйекті түрде жоғары өнімділікті көрсетеді, барлық метрикалық мәндер 89%-дан асады, бұл күшті жіктеу мүмкіндігін көрсетеді. Ең жақсы өнімділік Crack класында байқалады, мұнда дәлдік ең жоғары мәнге (96,1%) жетеді, одан кейін дәлдік (95,6%), F1 ұпайы (95,4%) және қайта шақыру (94,8%) болады. Бұл модель ең аз жалған позитивті және жалған теріс мәндері бар жарықшақпен байланысты ақауларды анықтауда жоғары тиімді екенін көрсетеді.

Беттік ақау санаты үшін барлық көрсеткіштер салыстырмалы түрде тұрақты болып қалады, шамамен 92,9% және 93,8% аралығында. Бұл Crack класынан сәл төмен болса да, теңдестірілген өнімділікті көрсетеді, мүмкін күрделі беттік вариацияларға байланысты.

Edge ақау класы барлық санаттар арасында ең төмен өнімділікті көрсетеді, дәлдік 89,7%-ға және F1 ұпайы 90,1%-ға дейін төмендейді. Сол сияқты, еске түсіру (90,6%) және дәлдік (91,2%) салыстырмалы түрде төмен. Бұл тенденция шеткі ақауларды, мүмкін, нәзік немесе анық емес сипаттамалардың себебінен ажырату қиынырақ екенін көрсетеді.

Керісінше, «Ақау жоқ» санаты өнімділіктің күшті қалпына келуін көрсетеді,

дәлдік (95,2%), F1 ұпайы (94,6%), дәлдік (94,5%) және қайта шақыру (94,1%) жоғары деңгейлерге оралады. Бұл модель ақаулы және ақаусыз үлгілерді тиімді ажырататынын көрсетеді.



Сурет 6 — Ұсынылған CNN моделінің ақау санаттары бойынша өнімділігін салыстыру

Графикалық тенденциялар барлық көрсеткіштер бойынша Edge ақауы санаты үшін өнімділіктің айтарлықтай төмендеуін көрсетеді, содан кейін ақау жоқ сыныбы үшін айтарлықтай өсу. Төрт метрикалық қисықтың санаттар бойынша салыстырмалы жақын туралануы үлгінің дәлдік пен еске түсіру арасындағы жақсы тепе-теңдікті сақтайтынын көрсетеді, нәтижесінде тұрақты F1 ұпайлары алынады.

Талқылау. Эксперимент нәтижелері ұсынылған CNN негізіндегі модель барлық ақаулар санаттарында күшті және дәйекті өнімділікке қол жеткізетінін көрсетеді. Дәлдік, дәлдік, еске түсіру және F1 баллының жоғары мәндері (барлығы 90%-дан жоғары) модельдің кірпіш ақауларын автоматтандырылған анықтау үшін сенімді және берік екенін көрсетеді. Нәтижелердегі негізгі бақылау барлық бағалау көрсеткіштері ең жоғары мәндерге жететін Crack ақауларын анықтаудағы жоғары өнімділік болып табылады. Бұны сызаттардың айқын көрнекі сипаттамаларына жатқызуға болады, мысалы, өткір жиектер және анық құрылымдық үзілістер, олар ерекшеліктерді алу кезінде конволюционды кабаттармен тиімді түрде түсіріледі. Нәтижесінде, модель басқа беттік текстуралардан жарықтар үлгілерін оңай ажырата алады.

Модель сонымен қатар ақаусыз кірпіштерді дұрыс жіктеу қабілетін көрсете отырып, ақаусыз үлгілерді анықтауда жақсы жұмыс істейді. Бұл нақты әлемдегі қолданбаларда өте маңызды, өйткені жалған позитивтерді азайту бұзылмаған материалдардың қате қабылданбауын қамтамасыз етеді.

Керісінше, Edge Defect санатында байқалған салыстырмалы түрде төмен өнімділік бұл сыныптың күрделірек болып қала беретінін көрсетеді. Жиектер ақаулары көбінесе басқа ақау түрлерінің ерекшеліктерімен немесе тіпті қалыпты текстуралармен қабаттасуы мүмкін нәзік немесе тұрақты емес үлгілерді көрсетеді. Бұл анық еместік қате жіктеуге әкелуі мүмкін, әсіресе дәлдік пен F1 ұпайына әсер етеді. Сол сияқты, беттік ақаулардың нәтижелері тұрақты болып қалса да, олар Crack нәтижелерінен сәл төмен, бұл беттік деңгейдегі бұзушылықтарды айырудың орташа күрделілігін көрсетеді.

Жалпы жоғары өнімділікке қарамастан, бірнеше шектеулерді атап өту керек. Біріншіден, деректер жиынтығының өлшемі мен әртүрлілігі модельдің жалпылау қабілетіне әсер етуі мүмкін. Екіншіден, ерекшеленбейтін белгілері бар кейбір ақау түрлері неғұрлым жетілдірілген мүмкіндіктерді шығару әдістерін қажет етуі мүмкін. Деректерді көбейту, назар аудару механизмдері немесе тереңірек архитектуралар сияқты әдістерді қосу өнімділікті одан әрі жақсартуы мүмкін.

Қорытынды. Ұсынылған зерттеу жұмысында кірпіш ақауларын автоматтандырылған түрде анықтау және сапаны тексеру үшін CNN негізіндегі терең оқыту моделін ұсынды. Өзірленген жүйе кескінге негізделген талдауды қолдана отырып, жарықтар, беткі ақаулар, жиек ақаулары және ақаусыз үлгілер сияқты әртүрлі ақау санаттарын сәтті анықтады.

Эксперименттік нәтижелер барлық санаттар бойынша дәлдік, дәлдік, еске түсіру және F1 ұпай мәндері 90%-дан асатын жоғары жіктеу өнімділігін көрсетті. Зерттеу нәтижелері терең оқыту әдістерінің өнеркәсіптік сапаны бақылау жүйелерінің дәлдігі мен сенімділігін тиімді түрде жақсарту алатынын растайды.

Сонымен қатар, жасанды интеллект пен автоматтандыру технологияларының интеграциясы қолмен тексеру жұмыстарын азайтып, адами қателіктерді азайтып, ақылды өндіріс процестерін қолдай алады. Жалпы алғанда, ұсынылған тәсіл автоматтандырылған кірпіш өндірісінде және ақылды өнеркәсіптік мониторинг жүйелерінде практикалық қолдану аясын кеңейтеді.

References

- Beskopylny A.N., Shcherban' E.M., Stel'makh S.A., Mailyan L.R., Meskhi B., Razveeva I., ... & Onore G. (2023) Discovery and classification of defects on facing brick specimens using a convolutional neural network. *Applied Sciences*, — №19. — P. 693 – 708 DOI: <https://doi.org/10.3390/app13095413> (in Eng.).
- Wang N., Zhao X., Zhao P., Zhang Y., Zou Z., & Ou J. (2019) Automatic damage detection of historic masonry buildings based on mobile deep learning. *Automation in Construction*, — № 103. — P. 53-66 DOI: [10.1016/j.autcon.2019.03.003](https://doi.org/10.1016/j.autcon.2019.03.003) (in Eng.).
- Kataev M.Y., & Bulysheva L.A. (2025) Computer vision-based automated defect detection in ceramic bricks. *Systems Research and Behavioral Science*, — № 42(4). — P.1131-1141 DOI: <https://doi.org/10.1002/sres.3040> (in Eng.).
- Dhali S., Islam M.H., Barmon S., & Ghosh A. (2023, December) An automatic brick grading system using convolutional neural network: Bangladesh perspective. In *International Conference on*

Advanced Computing and Intelligent Technologies Singapore: Springer Nature Singapore. — P. 135-150 DOI: https://doi.org/10.1007/978-981-97-1961-7_9 (in Eng.).

Cumbajin E., Rodrigues N., Costa P., Miragaia R., Frazão, L., Costa, N., ... & Pereira, A. (2023). A real-time automated defect detection system for ceramic pieces manufacturing process based on computer vision with deep learning. *Sensors*, — № 24(1). — P. 232-245 DOI: <https://doi.org/10.3390/s24010232> (in Eng.).

Cumbajin E., Rodrigues N., Costa P., Miragaia R., Frazão L., Costa N., ... & Pereira A. (2023) A systematic review on deep learning with CNNs applied to surface defect detection. *Journal of Imaging*, — № 9(10). — P. 193-219 DOI: <https://doi.org/10.3390/jimaging9100193> (in Eng.).

Zhou X., & Tiong R.L.K. (2025) Defects inspection system for building facades using drones and deep learning method. *Expert Systems with Applications*. — P. 129 145 DOI: 10.1016/j.eswa.2025.129715 (in Eng.).

Karimi N., Mishra M., & Lourenço P.B. (2025) Automated surface crack detection in historical constructions with various materials using deep learning-based YOLO network. *International Journal of Architectural Heritage*, — № 19(5). — P.581-597 DOI: <https://doi.org/10.1080/15583058.2024.2376177> (in Eng.).

Nazir U., Mian U.K., Sohail M.U., Taj M., & Uppal M. (2020) Kiln-net: A gated neural network for detection of brick kilns in South Asia. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, — №13. — P. 3251-3262 (in Eng.).

Ji A., Xue X., Dou Y., & Wang Y. (2021) Automatic detection of brick pavement defects using convolutional neural network. In *ICCREM 2021*. — P. 255-263 DOI: <https://doi.org/10.1061/9780784483848.029> (in Eng.).

Salman M. (2025, January) Computer Vision-based CNN Model Automated Defect Detection in Concrete Through Cellphone Camera. In *Conference: 2nd Olympiad in Engineering Science (OES 2025)* At: University of Stavanger, Oslo, Norway (in Eng.).

Şermet F., & Pacal I. (2024) Deep learning approaches for autonomous crack detection in concrete wall, brick deck and pavement. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, —№ 15(2). — P. 503-513 DOI: <https://doi.org/10.24012/dumf.1450640> (in Eng.).

Katsigiannis S., Seyedzadeh S., Agapiou A., & Ramzan N. (2023) Deep learning for crack detection on masonry façades using limited data and transfer learning. *Journal of Building Engineering*, — № 76. — P. 107 -115 DOI: 10.1016/j.jobe.2023.107105 (in Eng.).

Krishnan S.S.R., Karuppan M.N., Khadidos A.O., Khadidos A.O., Selvarajan S., Tandon S., & Balusamy B. (2025) Comparative analysis of deep learning models for crack detection in buildings. *Scientific Reports*, — №15(1). — P. 2125 -2146 DOI: <https://doi.org/10.1038/s41598-025-85983-3> (in Eng.).

Çavuş Ö., & Özkar M. (2025) Deep-Learning-Based Surface Crack Detection with a Mobile Robot for Preventive Conservation of Plastered Timber-Framed Historical Buildings. *International Journal of Architectural Heritage*. — P. 1-26 DOI: <https://doi.org/10.1080/15583058.2025.2595147> (in Eng.).

Publication Ethics and Publication Malpractice in the journals of the Central Asian Academic Research Center LLP

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the journals of the Central Asian Academic Research Center LLP implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The Central Asian Academic Research Center LLP follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the Central Asian Academic Research Center LLP.

The Editorial Board of the Central Asian Academic Research Center LLP will monitor and safeguard publishing ethics.

Requirements for articles design for publication in the journal are available on the websites:

www.nauka-nanrk.kz

<http://physics-mathematics.kz/index.php/en/archive>

ISSN2518-1726 (Online),

ISSN 1991-346X (Print)

Managing Editor: *A.Shormakova*

Editors: *D.S. Alenov, T. Apendiev*

Computer layout: *G.D. Zhadyranova*

Signed for print: June 15, 2026

Format: 70×90 1/16. 26.5 printed sheets. Order No. 2.