

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

**ACADEMIC SCIENTIFIC
JOURNAL OF COMPUTER SCIENCE**

**№4
2025**

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



CENTRAL ASIAN ACADEMIC
RESEARCH CENTER



**ACADEMIC SCIENTIFIC
JOURNAL OF COMPUTER
SCIENCE**

4 (356)

OCTOBER – DECEMBER 2025

**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, acting General Director of the Institute of Information and Computing Technologies CS MES 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>

BIYASHEV Rustam Gakashevich, doctor of technical sciences, professor, Deputy Director of the Institute for Informatics and Management Problems, Head of the Information Security Laboratory (Kazakhstan), <https://www.scopus.com/authid/detail.uri?authorId=6603642864>, <https://www.webofscience.com/wos/author/record/3802016>

KAPALOVA Nursulu Aldazarovna, 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/>

БАС РЕДАКТОР:

МҮТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=6506682964>, <https://www.webofscience.com/wos/author/record/1423665>

РЕДАКЦИЯ АЛҚАСЫ:

ҚАЛИМОЛДАЕВ Максат 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>

БАЙҒҮНЧЕКОВ Жұмаділ Жанабайұлы, 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>

БИЯШЕВ Рустам Гакашевич, техника ғылымдарының докторы, профессор, Информатика және басқару мәселелері институты директорының орынбасары, Ақпараттық қауіпсіздік зертханасының меңгерушісі (Қазақстан), <https://www.scopus.com/authid/detail.uri?authorId=6603642864>, <https://www.webofscience.com/wos/author/record/3802016>

КАПАЛОВА Нұрсұлу Алдаржарқызы, 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, 2025

ГЛАВНЫЙ РЕДАКТОР:

МУТАНОВ Галимжаир Мутанович, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МНВО 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 Валдемар, 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>

ОТМАН Мохамед, 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>

БИЯШЕВ Рустам Гакашевич, доктор технических наук, профессор, заместитель директора Института проблем информатики и управления, заведующий лабораторией информационной безопасности (Казахстан), <https://www.scopus.com/authid/detail.uri?authorId=6603642864>, <https://www.webofscience.com/wos/author/record/3802016>

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 Александр 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 учет 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 КОКСНВО 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», 2025

CONTENTS

B. Assanova, Zh. Moldasheva, A.T. Kishubaeva Decision support system structure and blocks for selecting efficient delayed coking modes.....	11
Zh.T. Abildayeva, R.K. Uskenbayeva, G.S. Beketova, N.B. Konyrbaev, S.B. Seydazimov Multi-criterion optimization of advertising budget allocation in the agro-industrial complex based on NSGA-III algorithm.....	26
A.O. Aliyeva, B.S. Omarov, R.B. Abdrakhmanov, D.R. Sultan, A.B. Toktarova Neural network model for automatic detection of Kazakh-language hatespeech.....	40
O. Auyelbekov, E. Bostanov, S. Sapakova, L. Tukenova, A. Kozhagul Modeling and analysis of a generator with permanent and variable magnets.....	55
G. Autova, G. Nurtayeva, E. Zulfukharova, G. Yeleussizova, R. Zhumabekova Theoretical foundations of interdisciplinary integration of physics and computer science.....	73
A.Zh. Akhmetova, M.A. Kantureyeva, A.A. Abisheva, A. Aubakirova, A.A. Shekerbek Analysis of the social network user's environment.....	89
A.Sh. Barakova, K.S. Shadinova, A.S. Orynbaeva, G. Sugurzhanova Design of a model for protecting a website's authentication data and content based on blockchain technology.....	102
A.N. Zhidebayeva, G.U. Madaliyeva, B.O. Tastanbekova, S.S. Karzhaubekova, G.S. Shaimerdenova Deep neural network Conv-LSTM for ECG-based cardiac disorder identification.....	122
N.M. Zhunissov, A.B. Aben, A.B. Amanzholova The fraud detection model in text messages.....	138
A. Issakhov, A. Alzhanov, A. Akhmedov, A. Amanzholov, T. Murat Numerical simulation of thermohydrodynamics during heated water discharge into Lake Balkhash.....	152

Z. Kaderkeyeva, B. Razakhova, G. Bekmanova, A. Nazyrova, M. Zhasuzakova
Q-Bilim: an intelligent system for assessing learning outcomes based on competencies.....171

N. Karymsakova, A. Boltaboyeva, D. Turmakhanbet, M. Maulenbekov, T. Abdirova
Unsupervised learning for the identification of critical conditions in renewable energy production.....184

A.Kulakayeva, E.Daineko, B. Medetov, A. Nurlankyzy
Evaluation of the effectiveness of modern neural network architectures for VAD under low snr ratio conditions.....203

B. Orazbayev, A. Zhumadillayeva, K. Orazbayeva, R. Yessirkessinov, Zh. Tuleuov
Development of models of sulfur production processes based on artificial neural networks and simulation.....216

L. Rzayeva, A. Ryzhova, M. Zhaparkhanova, A. Myrzatay, Zh. Kozhakhmet
A new LSTM-based web application for automated password strength evaluation.....234

D. Sagidoldin, A. Zhetpisbayeva, B. Zhumazhanov, B. Zhumazhanov
Increasing the reliability of data transmission from small spacecraft using SDR equipment.....259

A.N. Seraly, A.D. Mekhtiyev, G.Z. Ziyatbekova, K.B. Begalieva, R.A. Mekhtiyev
Development of hardware for monitoring optical parameters.....274

A.A. Taurbekova, M.V. Markosyan
Development and implementation of a computational model of magmatic processes in the bowls of the Earth and on its surface.....288

K. Chezhimbayeva, A. Mukhamejanova, Y. Garmashova
Fuzzy-logic-based expert system for predicting QoS in 5G networks.....306

МАЗМҰНЫ

Б.У. Асанова, Ж.Ж. Молдашева, А. Кишубаева Баяу кокстеу қондырғысы үшін тиімді жұмыс режимдерін таңдауға шешім қолдау жүйесі құрылымы.....	11
Ж.Т. Әбілдаева, Р.К. Ускенбаева, Г.С. Бекетова, Н.Б. Қоңырбаев, С.Б. Сейдазимов NSGA-III алгоритмі негізінде агроөнеркәсіптік кешендегі жарнамалық бюджетті бөлуді көп критериялы оңтайландыру.....	26
А.О. Әлиева, Б.С. Омаров, Р.Б. Абдрахманов, Д.Р. Султан, А.Б. Тоқтарова Қазақ тіліндегі дискриминацияны автоматты анықтауға арналған нейрондық желілік моделі.....	40
О. Әуелбеков, Е. Бостанов, С. Сапақова, Л. Түкенова, А. Қожағұл Тұрақты және айнымалы магниттері бар генераторды модельдеу және талдау.....	55
Г.М. Аутова, Г.К. Нуртаева, Ә.М. Зильбухарова, Г.С. Елеусизова, Р.Р. Жұмабекова Физика мен информатика пәндерінің пәнаралық интеграциясының теориялық негіздері.....	73
А.Ж. Ахметова, М.А. Кантуреева, А.А. Абишева, А. Аубакирова, А.А. Шекербек Әлеуметтік желі қолданушыларының ортасын талдау.....	89
А.Ш. Баракова, К.С. Шадинова, А.С. Орынбаева, Г. Сугуржанова Блокчейн технологиясы негізінде веб сайттың аутентификациялық деректері мен өнімін қорғау моделін құрастыру.....	102
А.Н. Жидебаева, Г.У. Мадалиева, Б.О. Тастанбекова, С.С. Қаржаубекова, Г.С. Шаймерденова Жүрек ауруларын анықтауда Conv-LSTM архитектурасына негізделген терең нейрондық желі.....	122
Н.М. Жунисов, А.Б. Абен, Ә.Б. Аманжолова Мәтіндік хабарламалардағы алаяқтықты анықтау моделі.....	138
А.А. Исахов, А. Альжанов, А. Ахмедов, А. Аманжолов, Т. Мурат Балқаш көліне жылы су ағызу кезіндегі термогидродинамиканы сандық модельдеу.....	152

З.К. Кадеркеева, Б.Ш. Разахова, Г.Т. Бекманова, А.Е. Назырова, М.Ж. Жасұзақова Q-Bilim: құзыреттерге негізделген оқу нәтижелерін бағалауға арналған интеллектуалды жүйе.....	171
Н. Карымсакова, А. Болтабоева, Д. Тұрмаханбет, М. Мауленбеков, Т. Абдирова Жанартылатын энергия өндірісіндегі критикалық режимдерді анықтауға арналған мұғалімсіз оқыту.....	184
А. Кулакаева, Е. Дайнеко, Б. Медетов, А. Нурланқызы Сигнал/шуыл қатынасы төмен жағдайларда заманауи нейрондық желілік VAD архитектураларының тиімділігін бағалау.....	203
Б. Оразбаев, А. Жумадиллаева, К. Оразбаева, Р. Есиркесинов, Ж. Тулеуов Күкірт өндіру процесстерінің модельдерін жасанды нейрондық желілер негізінде әзірлеу және модельдеу.....	216
Л. Рзаева, А. Рыжова, М. Жапарханова, А. Мырзатай, Ж. Кожамет, Құпиясөздің беріктігін автоматты бағалауға арналған LSTM негізіндегі жаңа веб-қосымша.....	234
Д.Т. Сагидолдин, А.Т. Жетписбаева, Б.Р. Жумажанов, Б.С. Жумажанов SDR жабдықтарын пайдалану арқылы, шағын ғарыш аппараттарынан деректерді берудің сенімділігін арттыру.....	259
А.Н. Сералы, А.Д. Мехтиев, Г.З. Зиятбекова, К.Б. Бегалиева, Р.А. Мехтиев Оптикалық параметрлерді бақылауға арналған аппараттық құрылғыны әзірлеу.....	274
А.А. Таурбекова, М.В. Маркосян Жер көзіндегі және оның бетіндегі магматтық процестердің есептік моделін әзірлеу және енгізу.....	288
К.С. Чежимбаева, А. Мухамеджанова, Ю. Гармашова Айқын емес логика негізінде 5G желілеріндегі QoS болжау expertтік жүйесі.....	306

СОДЕРЖАНИЕ

Б.У. Асанова, Ж.Ж. Молдашева, А. Кишубаева Структура и функциональные блоки системы поддержки решений для выбора режимов замедленного коксования.....	11
Ж.Т. Абилдаева, Р.К. Ускенбаева, Г.С. Бекетова, Н.Б. Конырбаев, С.Б. Сейдазимов Многокритериальная оптимизация распределения рекламного бюджета в апк на основе алгоритма NSGA-III.....	26
А.О. Алиева, Б.С. Омаров, Р.Б. Абдрахманов, Д.Р. Султан, А.Б. Токтарова Нейросетевая модель для автоматического обнаружения дискриминации в казахском языке.....	40
О. Ауельбеков, Е. Бостанов, С. Сапакова, Л. Туkenова, А. Кожугул Моделирование и анализ генератора с постоянными и переменными магнитами.....	55
Г.М. Аутова, Г.К. Нуртаева, Э.М. Зулбухарова, Г.С. Елеусизова, Р.Р. Жумабекова Теоретические основы междисциплинарной интеграции физики и информатики.....	73
А.Ж. Ахметова, М.А. Кантуреева, А.А. Абишева, А. Аубакирова, А.А. Шекербек Анализ окружения ползователей социальной сети.....	89
А.Ш. Баракова, К.С. Шадинова, А.С. Орынбаева, Г. Сугуржанова Разработка модели защиты аутентификационных данных и контента веб-сайта на основе технологии блокчейн.....	102
А.Н. Жидебаева, Г.У. Мадалиева, Б.О. Тастанбекова, С.С. Каржаубекова, Г.С. Шаймерденова Глубокая нейронная сеть на основе архитектуры Conv-LSTM для выявления сердечных заболеваний.....	122
Н.М. Жунисов, А.Б. Абен, А.Б. Аманжолова Модель обнаружения мошенничества в текстовых сообщениях.....	138
А.А. Исahов, А. Альжанов, А. Ахмедов, А. Аманжолов, Т. Мурат Численное моделирование термогидродинамики при сбросе подогретых вод в озеро Балхаш.....	152

З.К. Кадеркеева, Б.Ш. Разахова, Г.Т. Бекманова, А.Е. Назырова, М.Ж. Жасузакова Q-Bilim: интеллектуальная система оценки результатов обучения на основе компетенций.....	171
Н. Карымсакова, А. Болтабоева, Д. Тұрмаханбет, М. Мауленбеков, Т. Абдирова Обучение без учителя для выявления критических режимов в производстве возобновляемой энергии.....	184
А. Кулакаева, Е. Дайнеко, Б. Медетов, А. Нурланкызы Оценка эффективности современных нейросетевых архитектур VAD при низком отношении сигнал/шум.....	203
Б. Оразбаев, А. Жумадиллаева, К. Оразбаева, Р. Есиркесинов, Ж. Тулеуов Разработка моделей процессов производства серы на основе искусственных нейронных сетей и моделирование.....	216
Л. Рзаева, А. Рыжова, М. Жапарханова, А. Мырзатай, Ж. Кожамет Новое веб-приложение на основе LSTM для автоматизированной оценки надежности паролей.....	234
Д.Т. Сагидолдин, А.Т. Жетписбаева, Б.Р. Жумажанов, Б.С. Жумажанов Повышение надёжности передачи данных с малых космических аппаратов с применением SDR оборудования.....	259
А.Н. Сералы, А.Д. Мехтиев, Г.З. Зиятбекова, К.Б. Бегалиева, Р.А. Мехтиев Разработка аппаратного средства для контроля оптических параметров.....	274
А.А. Таурбекова, М.В. Маркосян, Н.Т. Карымсакова Разработка и реализация вычислительной модели магматических процессов в недрах земли и на её поверхности.....	288
К.С. Чежимбаева, А. Мухамеджанова, Ю. Гармашова Экспертная система прогнозирования QoS в 5G-сетях на основе нечеткой логики.....	306

<https://doi.org/10.32014/2025.2518-1726.382>

IRSTI 27.25.19

UDC 004.89; 004.912

© **A.O. Aliyeva¹, B.S. Omarov², R.B. Abdrakhmanov³, D.R. Sultan⁴,
A.B. Toktarova^{5*}, 2025.**

¹Khoja Akhmet Yassawi International Kazakh – Turkish University,
Turkistan, Kazakhstan;

²International information technology university, Almaty, Kazakhstan;

³International University of Tourism and Hospitality, Turkestan, Kazakhstan;

⁴Narxoz University, Almaty, Kazakhstan;

⁵Mukhtar Auezov South Kazakhstan University, Shymkent, Kazakhstan.

E-mail: toktar.aigerim@list.ru

NEURAL NETWORK MODEL FOR AUTOMATIC DETECTION OF KAZAKH-LANGUAGE HATE SPEECH

Aliyeva Akbayan — Candidate of Pedagogical Sciences, department of Computer engineering, Khoja Akhmet Yassawi International Kazakh – Turkish University, Turkistan, Kazakhstan,
E-mail: aliyevaakbayan@mail.ru, <https://orcid.org/0009-0006-6932-7302>;

Omarov Batyrkhan — PhD, associate professor of the Department of Information System, International information technology university, Almaty, Kazakhstan,
E-mail: aliyevaakbayan@mail.ru, <https://orcid.org/0009-0006-6932-7302>;

Abdrakhmanov Rustam — Cand.of Techn.Sc., Ass.Prof., International University Of Tourism And Hospitality Turkestan, Kazakhstan,
E-mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>;

Sultan Danyar — PhD, Department of Information System, Narxoz university, Almaty, Kazakhstan,
E-mail: sultan.daniyar96@gmail.com, <https://orcid.org/0000-0002-1611-1923>;

Toktarova Aigerim — PhD, department of Information communication technologies, M. Auezov South Kazakhstan university, Shymkent, Kazakhstan,
E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>.

Abstract. This article describes how to use modern machine and deep learning algorithms to automatically detect offensive language in Kazakh-language web content. The relevance of the research work is aimed at protecting the digital environment and reducing discriminatory language in social media and user platforms. To achieve this goal, methods for data collection, pre-processing, classification model creation and optimization were developed and tested. The article describes the collection of data from social networks (Instagram, TikTok, YouTube, Facebook), news sites and forums with Kazakh-language user-generated content as the object of research. As a practical significance of the research work, data pre-processing includes noise removal, tokenization, lemmatization and stop

word removal. As a result of the presented work, the annotated corpus classified communications as “offensive language words” or “neutral opinions”. The authors use a hybrid deep learning architecture to create an efficient and flexible system for automatically detecting discrimination on the Kazakh-language Internet. CNN, LSTM, Bi-LSTM and Transformer models and their integration with the attention process are studied to consider local and contextual text features. The result of the article is aimed at training and validating a model that can recognize offensive and negative words in comments in online content in real time. This approach to social media moderation platforms is aimed at improving user safety, removing dangerous comments, and laying the foundation for research into natural language processing in the Kazakh language. Multilingual platforms can use the proposed method to create intelligent systems for automatic moderation and long-term protection of the digital environment.

Key words: collecting data, online content, hate speech, neural network, deep learning, detection

Acknowledgments

This study was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant No. AP23488900-Automatic detection of cyberbullying among young people in social networks using artificial intelligence)

© А.О. Әлиева¹, Б.С. Омаров², Р.Б. Абдрахманов³, Д.Р. Султан⁴,
А.Б. Тоқтарова^{5*}, 2025.

¹Қожа Ахмет Яссауи атындағы Халықаралық Қазақ – Түрік университеті,
Түркістан, Қазақстан;

²Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан;

³Халықаралық туризм және меймандостық университеті,
Түркістан, Қазақстан;

⁴Naqhoz университеті, Алматы, Қазақстан;

⁵М. Әуезов атындағы Оңтүстік Қазақстан университеті,
Шымкент, Қазақстан.

E-mail: toktar.aigerim@list.ru

ҚАЗАҚ ТІЛІНДЕГІ ДИСКРИМИНАЦИЯНЫ АВТОМАТТЫ АНЫҚТАУҒА АРНАЛҒАН НЕЙРОНДЫҚ ЖЕЛІЛІК МОДЕЛІ

Әлиева Ақбаян — педагогика ғылымдарының кандидаты, Компьютерлік инженерия кафедрасының доценті, Қожа Ахмет Яссауи атындағы Халықаралық Қазақ – Түрік университеті, Түркістан, Қазақстан,

E-mail: aliyevaakbayan@mail.ru, <https://orcid.org/0009-0006-6932-7302>;

Омаров Батырхан — PhD, қауымдастырылған профессор, Ақпараттық жүйелер кафедрасының доценті, Халықаралық Ақпараттық технологиялар университеті, Алматы, Қазақстан,

E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

Абдрахманов Рустам — т.ғ.к., доцент, Халықаралық туризм және меймандостық университеті Түркістан, Қазақстан,

E-mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>;

Султан Данияр — PhD, Ақпараттық жүйелер кафедрасының доценті, Naғхоз университетінің Алматы, Қазақстан,

E-mail: sultan.daniyar96@gmail.com, <https://orcid.org/0000-0002-1611-1923>;

Токтарова Айгерім — PhD, Ақпараттық коммуникациялық технологиялар кафедрасының аға оқытушысы, М.Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан,

E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>.

Аннотация. Бұл мақала қазақ тіліндегі веб-контенттегі бейәдеп сөздерді автоматты түрде анықтау үшін заманауи машина және терең оқыту алгоритмдерін қолдану тәсілдерін сипаттайды. Ғылыми зерттеудің өзектілігі цифрлық ортаны қорғау және әлеуметтік медиа мен пайдаланушылық платформаларда дискриминациялық тілдің таралуын азайту қажеттілігімен анықталады. Осы мақсатқа жету үшін деректерді жинау, алдын ала өңдеу, классификация моделін құру және оңтайландыру әдістері әзірленіп, сынақтан өткізілді. Мақалада зерттеу нысаны ретінде әлеуметтік желілерден (Instagram, TikTok, YouTube, Facebook), жаңалықтар сайттарынан және қазақтілді қолданушылар құрастырған контенті бар форумдардан алынған деректерді жинау үдерісі сипатталады. Зерттеудің практикалық маңыздылығы деректерді алдын ала өңдеу кезеңінің — шуды жою, токенизация, лемматизация және тоқтау сөздерді алып тастау сияқты қадамдарды қамтуымен түсіндіріледі. Зерттеу нәтижесінде аннотацияланған корпус коммуникацияларды «бейәдеп сөз» немесе «бейтарап пікір» деп жіктей алды. Авторлар қазақтілді интернеттегі дискриминацияны автоматты түрде анықтаудың тиімді әрі икемді жүйесін құру үшін гибриді терең оқыту архитектурасын қолданады. Жергілікті және мәтінмәндік ерекшеліктерді ескеруге мүмкіндік беретін CNN, LSTM, Bi-LSTM және Transformer үлгілері, сондай-ақ назар аудару механизмін интеграциялау тәсілдері талданады. Мақаланың нәтижесі онлайн-контенттегі пікірлерден бейәдеп және жағымсыз сөздерді нақты уақытта тани алатын модельді оқыту және валидациялау болды. Әлеуметтік желілерді модерациялау платформалары үшін ұсынылған әдіс қолданушылардың қауіпсіздігін арттыруға, қауіпті пікірлерді жоюға және қазақ тіліндегі табиғи тілді өңдеу бойынша болашақ зерттеулерге негіз қалайды. Көптілді платформалар ұсынылған тәсілді автоматты модерацияға арналған интеллектуалды жүйелерді дамыту және цифрлық ортаны ұзақ мерзімді қорғау үшін пайдалана алады.

Түйін сөздер: деректерді жинау, желілік контент, дискриминациялық сөздер, нейрондық желі, терең оқыту, анықтау

© А.О. Алиева¹, Б.С. Омаров¹, Р.Б. Абдрахманов³, Д.Р. Султан⁴,
А.Б. Токтарова^{5*}, 2025.

¹Международный казахско-турецкий университет имени Ходжи Ахмета Яссави, Туркестан, Казахстан;

²Международный университет информационных технологий, Алматы, Казахстан;

³Международный университет туризма и гостеприимства, Туркестан, Казахстан;

⁴Университет Нархоз, Алматы, Казахстан;

⁵Южно-Казахстанский университет имени М. Ауэзова, Шымкент, Казахстан.

E-mail: toktar.aigerim@list.ru

НЕЙРОСЕТЕВАЯ МОДЕЛЬ ДЛЯ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ ДИСКРИМИНАЦИИ В КАЗАХСКОМ ЯЗЫКЕ

Алиева Акбаян — кандидат педагогических наук, доцент кафедры Компьютерной инженерии, Международный казахско-турецкий университет имени Ходжи Ахмета Яссави, Туркестан, Казахстан,

E-mail: aliyevaakbayan@mail.ru, <https://orcid.org/0009-0006-6932-7302>;

Омаров Батырхан — PhD, доцент кафедры Информационные системы, Международный университет информационных технологий, Алматы, Казахстан,

E-mail: batyahan@gmail.com, <https://orcid.org/0000-0002-8341-7113>;

Абдрахманов Рустам — кандидат технических наук, доцент, Международный университет туризма и гостеприимства, Туркестан, Казахстан,

E-mail: abdrakhmanov.rustam@iuth.edu.kz, <https://orcid.org/0000-0003-0385-0684>;

Султан Данияр — PhD, доцент кафедры Информационные системы, Университет Нархоз, Алматы, Казахстан,

E-mail: sultan.daniyar96@gmail.com, <https://orcid.org/0000-0002-1611-1923>;

Токтарова Айгерим — PhD, кафедра Информационно-коммуникационных технологий, Южно-Казахстанский университет им. М. Ауэзова, Шымкент, Казахстан,

E-mail: toktar.aigerim@list.ru, <https://orcid.org/0000-0002-6265-9236>.

Аннотация. В данной статье рассматривается использование современных алгоритмов машинного и глубокого обучения для автоматического выявления ненормативной лексики в казахоязычном веб-контенте. Актуальность исследования заключается в необходимости защиты цифровой среды и снижении уровня дискриминационной лексики в социальных сетях и на пользовательских платформах. Для достижения этой цели были разработаны и апробированы методы сбора, предварительной обработки данных, а также построения и оптимизации моделей классификации. В статье приводится описание сбор данных из социальных сетей (Instagram, TikTok, YouTube, Facebook), новостных сайтов и форумов, объектом исследования которых является казахоязычный пользовательский контент. Практическая значимость исследования заключается в том, что предварительная обработка

данных включает в себя удаление шума, токенизацию, лемматизацию и удаление стоп-слов. В результате представленной работы аннотированный корпус позволил классифицировать сообщения как «ненормативную лексику» или «нейтральные мнения». Авторы используют гибридную архитектуру глубокого обучения для создания эффективной и гибкой системы автоматического выявления дискриминации в казахскоязычном интернете. Изучаются модели CNN, LSTM, Bi-LSTM и Transformer, а также их интеграция с процессом внимания для учёта локальных и контекстных особенностей текста. Результатом исследования стало обучение и валидация модели, способной распознавать оскорбительные и негативные выражения в комментариях к онлайн-контенту в режиме реального времени. Предложенный подход ориентирован на повышение безопасности пользователей, удаление опасных комментариев и создание базы для дальнейших исследований в области обработки естественного языка на казахском языке. Предложенная методология может быть адаптирована многоязычными платформами для разработки интеллектуальных систем автоматической модерации и долгосрочной защиты цифровой среды.

Ключевые слова: сбор данных, онлайн-контент, дискриминационные слова, нейронная сеть, глубокое обучение, обнаружение

Introduction. The advancement of digital technology and the pervasive utilization of social networks, instant messaging applications, and open communication platforms have substantially augmented the volume of user-generated information disseminated on the Internet. Simultaneously, there is a rise in the volume of messages featuring toxic language, insulting terminology, and various manifestations of hate speech. Such phenomena provide a significant risk to the psychological well-being of users and to the establishment of a secure and conducive digital environment. The necessity of creating efficient techniques for the automatic identification and filtration of such content arises from the imperative to safeguard vulnerable populations, particularly children and teenagers, while also ensuring a high standard of information security on the Internet. Recent years have witnessed notable advancements in the scientific literature regarding automated text analysis and the identification of undesirable content. Machine learning and deep neural network methodologies have been established and are currently employed, encompassing models based on convolutional and recurrent architectures (CNN, RNN, LSTM), in addition to transformers and their variants.

These methodologies are effectively employed to address the issues of classifying toxic communications, detecting cyberbullying, and forecasting manifestations of violent behavior in both English-language and multilingual online environments (Omarov, 2025). Furthermore, the deployment of such systems on platforms for automated moderation of social networks, online forums, and news portals can markedly diminish the velocity of harmful information propagation. Nonetheless,

despite advancements in this domain, the Kazakh-language section of the Internet remains inadequately researched. The scarcity of comprehensive labeled corpora, inadequate representation of the Kazakh language in global research, and the unique morphological characteristics of the language provide significant challenges in developing effective and precise models (Azhibekova, 2025).

Developing and training algorithms necessitates intricate tasks, including data collection from many sources, preprocessing, annotation, and the creation of a specific corpus. Creating tailored systems that accurately recognize toxic language and symptoms of hate speech in the local digital environment is contingent upon a high-quality and representative corpus. Consequently, undertaking research in this domain is essential for establishing a secure and culturally attuned Internet environment in Kazakhstan.

Related work. Recent years have seen a proliferation of research focused on the automatic identification of toxic language and hate speech, predominantly within the English-speaking domain of the Internet. International research emphasizes the advancement and comparative evaluation of machine learning algorithms, encompassing traditional classification techniques and contemporary deep learning methodologies.

Initially, such issues were addressed by conventional text processing techniques and statistical models. Specifically, Support Vector Machine (SVM), Logistic Regression, Naive Bayes, and Random Forest were employed, yielding appropriate outcomes with meticulously chosen features. Nevertheless, these strategies were inadequate when addressing substantial quantities of unstructured texts and when attempting to discern the concealed semantics of hurtful remarks.

With the advancement of deep learning technologies, recurrent neural networks (RNN), long short-term memory networks (LSTM) and its variants, together with convolutional neural networks (CNN), have been extensively employed to extract contextual and local information from text. Recent research illustrates the superior efficacy of hybrid models that include LSTM and CNN, alongside the implementation of attention mechanisms and transformers (BERT, RoBERTa), which facilitate the consideration of both local and global dependencies within the text.

Significant emphasis in global literature is placed on the development of specialized corpora for hate speech detection tasks. Huang proposed a multilingual Twitter corpus to examine demographic bias in the identification of harmful information (Huang, 2020). Sachdeva utilized Rasch measurement theory to evaluate the reliability of annotations and discern variations in the perception of message toxicity (Sachdeva, 2024). Fanton developed a corpus of counter-narratives to counteract hate speech employing the Human-in-the-Loop methodology (Fanton, 2022).

Notwithstanding advancements in international research, the Kazakh-language segment remains inadequately explored. The scarcity of open corpora and the

limited availability of annotated data hinder the creation of precise models. This underscores the necessity of developing localized corpora and tailoring contemporary deep learning architectures to the morphological characteristics of the Kazakh language (Mussiraliyeva, 2021).

Materials and Methods. This study employs a comprehensive methodology for the automatic identification of hate speech in Kazakh-language online content, encompassing multiple sequential phases: data collection, pre-processing, annotation, corpus formation, and the development and evaluation of deep learning models (Bolatbek, 2024).

The process of constructing a data corpus is crucial for developing a system that automatically detects hate speech in Kazakh-language online content, as the efficacy of subsequent training of machine and deep learning models is directly contingent upon the quality and representativeness of the gathered data.

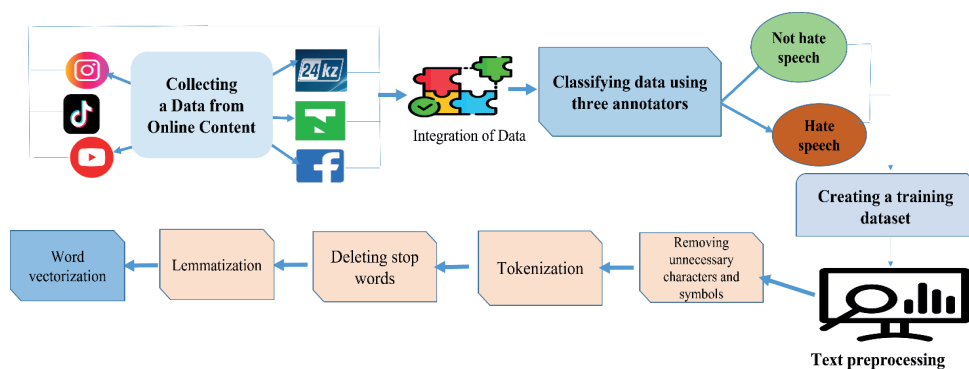


Figure 1 - Data collection and preprocessing framework

A variety of sources were utilized to construct the corpus, ensuring extensive representation of user speech (Figure 1). The primary data sources were:

a) Social networks such as Instagram, TikTok, YouTube, and Facebook offer a plethora of casual content accompanied by emotive and spontaneous commentary, where instances of hate speech are prevalent.

b) News portals and online media such as Tengrinews, Informburo, and Massaget feature comments on news that reflect socially significant issues and include instances of violent language.

c) Internet forums and specialized platforms, where in-depth conversations occur and covert instances of foul language may be present.

Two methodologies were employed to automate the collecting:

a) Utilization of Python scripts employing the BeautifulSoup and Selenium libraries for the direct processing of web pages to extract comments and messages from open sources.

b) Utilizing official social media APIs (YouTube Data API, Facebook Graph API) enabled the collection of substantial data in a structured format (Haleem,2021).

To guarantee the representativeness of the corpus, the sample encompassed texts that illustrated both informal user discourse (colloquial language, emotive expressions) and public discourses (socially and politically pertinent subjects). At this step, initial filtering was conducted, which involved eliminating duplicates, discarding unnecessary messages (such as advertisements), and identifying the language to keep solely Kazakh-language content.

Following the establishment of the primary corpus, manual annotation was conducted with the involvement of three voluntary annotators: a researcher, a psychologist, and a Kazakh language educator. Each text was categorized as either "hate speech" or "non-hate speech". This method guaranteed a significant degree of dependability and consistency in the annotation.

Consequently, a systematic and representative corpus of Kazakh-language internet content was established, serving as the foundation for later preprocessing, model creation, and training of classification algorithms (Ojo, 2022).

The concluding phase of the study entailed the use and evaluation of deep learning techniques for the automatic classification of Kazakh-language internet content as either including or not comprising hate speech. The study initially encompassed traditional deep learning architectures, including LSTM, Bi-LSTM, CNN, and MLP, facilitating an evaluation of their efficacy in addressing the agglutinative structure of the Kazakh language and revealing the limitations of each model (Zhu, 2019).

The preliminary stage results indicated that employing each architecture alone yields adequate accuracy; nonetheless, it is insufficient for detecting contextually obscured insults and latent manifestations of hate speech (Guo, 2025). A hybrid model was devised and executed, integrating the advantages of multiple deep learning techniques.

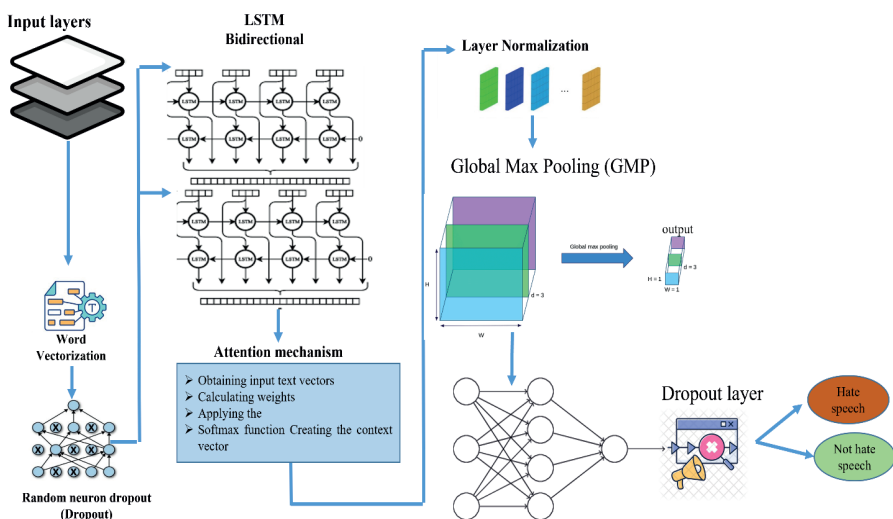


Figure 2 - Framework for the development and training of models

The hybrid architecture's structure, seen in the diagram, comprises the subsequent stages (Figure 2):

Word vectorization is an important step in preparing text data for processing by deep learning algorithms. After text preprocessing, including tokenization, lemmatization, and stop word removal, each word is converted into a fixed-dimensional numerical vector representation that reflects its meaning and contextual relationships.

$$W_i \rightarrow \bar{v}_i = (v_{i1}, v_{i2}, \dots, v_{id}) \quad (1)$$

Where W_i – input words, \bar{v}_i – It is a vector representation of dimension d , comprising numerical values that encapsulate the semantic attributes of the word.

The Word2Vec and FastText models are employed to generate such vectors. The Word2Vec model generates vector representations derived from word contexts, assigning like vectors to words that appear in comparable situations. FastText enhances this methodology by incorporating subwords and morphological components, which is particularly significant for the Kazakh language due to its agglutinative nature (Chen, 2024).

Vectorization assigns a numerical representation to each lexeme, enabling input into neural networks like Bi-LSTM or CNN for enhanced feature extraction and categorization of messages into the categories of "hate speech" and "non-hate speech".

During the extraction of contextual data, bidirectional LSTM layers are employed, enabling the model to concurrently assess the text sequence from both left to right and right to left (Zhang, 2018). This method guarantees that both preceding and subsequent words are considered for each text element, which is particularly crucial for the Kazakh language due to its agglutinative morphology and flexible word order. Consequently, deep context vectors are generated, representing semantic dependencies inside a phrase and latent semantic relationships among words.

$$h_t = [h_t^{fwd}; h_t^{bwd}] \quad (2)$$

Where, h_t^{fwd} – concealed state of advancement,

h_t^{bwd} – hidden state of the opposite direction,

h_t – shared context vector for position t

The utilization of Bi-LSTM facilitates the successful extraction of long-term dependencies and consideration of sentence context when categorizing text as "hate speech" or "non-hate speech." The integration of these features with subsequent convolutional or fully connected layers enhances the model's accuracy while processing real data from social networks and news sources.

The subsequent phase in text processing is employing a Convolutional Neural Network (CNN) to extract localized features and distinctive patterns

within text input. In the realm of hate speech analysis, CNN excels in detecting n-gram dependencies that may signify the existence of hostile or toxic words. A convolutional layer systematically "slides" over a matrix of word embeddings, calculating a filter response for each segment of text. Each filter of fixed width k encapsulates the local context of k successive words (Naser,2025).

Consequently, a feature map is generated for each filter, indicating the frequency of a specific pattern's occurrence in the text.

$$c_i = f(W * x_{i:i+k+1} + b) \quad (3)$$

$x_{i:i+k+1}$ –vector representation of a word sequence of length k

W - filter weight matrix

b –bias

$f(*)$ – activation function

c_i – the value of the attribute at position i

Subsequent to the use of convolutional filters, Global Max Pooling is employed to identify the most salient features over the full sentence length. This enables us to diminish the dimensionality of the feature space and concentrate on critical signs of hate speech, like combinations of harsh language, aggressive expressions, or distinct markers of toxicity.

Employing a CNN at this juncture equips the model to identify localized patterns of harmful discourse that may remain obscured when solely examining the overarching context. In the hybrid architecture, these features are integrated with context vectors derived from Bi-LSTM, enhancing the precision of message categorization (Sharifi, 2025).

The concluding phase of the model's functionality entails the amalgamation of two categories of features:

- a) Contextual characteristics derived from Bi-LSTM, which consider the semantic relationships of words inside a phrase;
- b) Local characteristics derived from CNN, which represent distinctive n-gram patterns and localized indicators of abusive language.

The features are amalgamated into a singular vector, which is subsequently input into a fully linked layer. This layer categorizes messages utilizing the Softmax function, which computes the likelihood of the text being classified as either "Hate speech" (contains hate speech) or "Not hate speech" (does not contain it).

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (4)$$

Where, z_i – output of the model for class i ,

K – number of classes

$P(y_i)$ - likelihood of a message being classified as belonging to class i

The research covered all important processes for automatically detecting hate

speech in Kazakh-language internet content, from data collection and annotation to hybrid deep learning model creation and evaluation. A sample corpus from social networks, news portals, and forums was preprocessed to remove stop words, tokenize, lemmatize, and filter unnecessary content. Three volunteer professionals annotated corpus. After converting the data into vector representations using Word2Vec and FastText models, Kazakh semantic and morphological properties were numerically represented. A hybrid architecture using Bi-LSTM for contextual dependency capture and CNN for local pattern recognition extracted features and built a classifier. The incorporated features were fed into a fully connected layer using a Softmax function to determine whether a message was "Hate speech" or "Not hate speech".

Results. Experiments were conducted on three data corpora—Kazakh, Russian, and English (Hate Speech and Offensive Language Dataset)—to assess the effectiveness of the proposed model. The model was evaluated against traditional machine learning algorithms (Logistic Regression, Random Forest, Decision Tree, Naïve Bayes, K-Nearest Neighbors, SVM) as well as contemporary deep learning architectures (LSTM, BiLSTM, CNN, RNN, MLP, Transformer).

Table 1. Evaluation of the efficacy in Kazakh, Russian, and English

Type of database	Method	Algorithm	Accuracy	Precision	Recall	F1-score	AUC-ROC		
Database in Kazakh	Deep learning	Proposed model	92.1%	93%	92.5%	92.9%	95%		
		LSTM	91.0%	92.5%	92.8%	92.5%	92%		
		BiLSTM	91.3%	92.6%	93%	92.7%	92%		
		CNN	89.2%	89%	89%	88.7%	90%		
		RNN	89.7%	89.5%	90%	89.9%	94%		
	Machine learning	LR	87.3%	85.2%	86.2%	85.1%	78%		
		RF	85.6%	83.9%	83.1%	83.7%	92%		
		DT	87.4%	83.2%	86.3%	85.1%	80%		
		NB	60.2%	52.4%	58.5%	64.2%	65%		
		KNN	85.1%	85.4%	82.2%	85.6%	77%		
		SVM	86.2%	85.3%	83.7%	85.8%	78%		
		Database in Russian (Russian Language Toxic Comments)	Deep learning	Ұсынылған модель	91.5%	93%	92%	92.4%	94%
				BiLSTM	91%	92%	92%	92%	93%
CNN	88.7%			89%	89%	88.9%	90%		
LSTM	90%			91%	91.5%	91.2%	93%		
RNN	89%			89%	91%	90%	92%		
Machine learning	LR		87.3%	85.6%	87.4%	60.2%	85.1%		
	RF		85.2%	83.9%	83.2%	52.4%	85.4%		
	DT		86.2%	83.1%	86.3%	58.5%	82.2%		
	NB		85.1%	83.7%	85.1%	64.2%	85.6%		
	KNN		75%	90%	76%	68%	77%		
SVM	87.3%	85.6%	87.4%	60.2%	85.1%				

Type of database	Method	Algorithm	Accuracy	Precision	Recall	F1-score	AUC-ROC
Database in English (<i>hate-speech-and-offensive-language</i>)	Deep learning	Recommended model	95%	96%	96.5%	96%	97%
		LSTM	93%	92.5%	94%	95%	95%
		MLP	91.5%	92%	93%	92.5%	94%
		CNN	90%	92%	91%	91.5%	93%
		RNN	91%	93%	92%	92%	92%
	Machine learning	LR	89.3%	90%	89.5%	89.3%	89%
		RF	88.6%	89%	88%	88.5%	89%
		DT	87.4%	83.2%	86.3%	85.1%	80%
		NB	86.2%	82.4%	85.5%	84.5%	85%
		KNN	87.1%	85.4%	87.2%	86.4%	87%
		SVM	87.2%	87.3%	86.7%	85.8%	88%

The hybrid model applied to the Kazakh corpus achieved an Accuracy of 92.1%, Precision of 93%, Recall of 92.5%, F1-score of 92.9%, and AUC-ROC of 95%, surpassing all traditional and individual neural architectures. Among the base models, BiLSTM demonstrated the highest performance; however, the incorporation of CNN and attention mechanisms into the hybrid model resulted in a substantial enhancement in performance, particularly in the F1-score metric, which indicates the equilibrium between recall and accuracy.

The Russian-language corpus yielded the following indicators: Accuracy 91.5%, Precision 93%, Recall 92%, F1-score 92.4%, and AUC-ROC 94%. These results further validate the stability and versatility of the proposed architecture within a morphologically rich language context.

The English-language corpus (Hate Speech and Offensive Language Dataset) yielded notable results: Accuracy 95%, Precision 96%, Recall 96.5%, F1-score 96%, and AUC-ROC of 97%. This results from the greater volume and improved balance of English-language data, facilitating more effective model training.

The comparative results indicate that the proposed model offers:

- a) Notable enhancement in all primary metrics relative to baseline models;
- b) Significant robustness to class imbalance, particularly relevant for Kazakh-language and Russian-language datasets.

The potential for generalization across various languages and domains, encompassing social networks and open platforms, is noteworthy.

The proposed architecture serves as an optimal solution for the automated detection of hate speech in multilingual online environments, achieving a balance among accuracy, recall, and overall performance.

The developed model's practical applicability is demonstrated through the implementation of the web interface "Kazakh Hate Speech Detection," which analyzes text messages to classify them as either "Hate speech" or "Not hate speech" in real time.

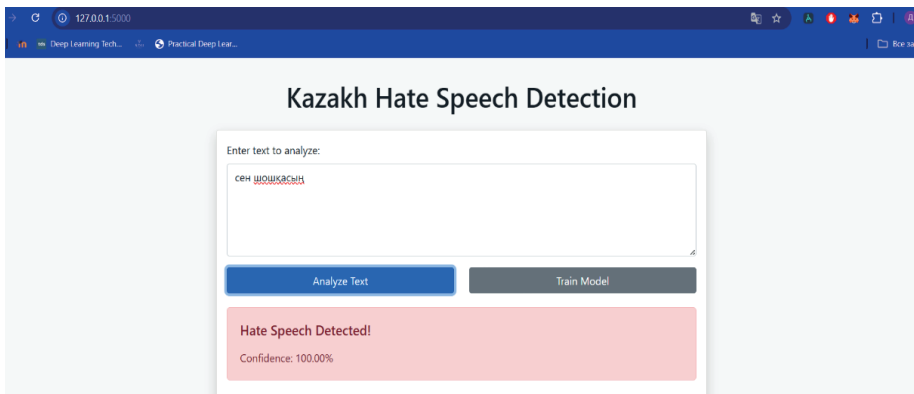


Figure 3 - Web interface of hate speech detection (hate speech)

Figure 3 illustrates an instance of inputting a message with offensive language ("sen shoshqasyn"), which the system accurately categorizes as "Hate speech" with a certainty of 100%.

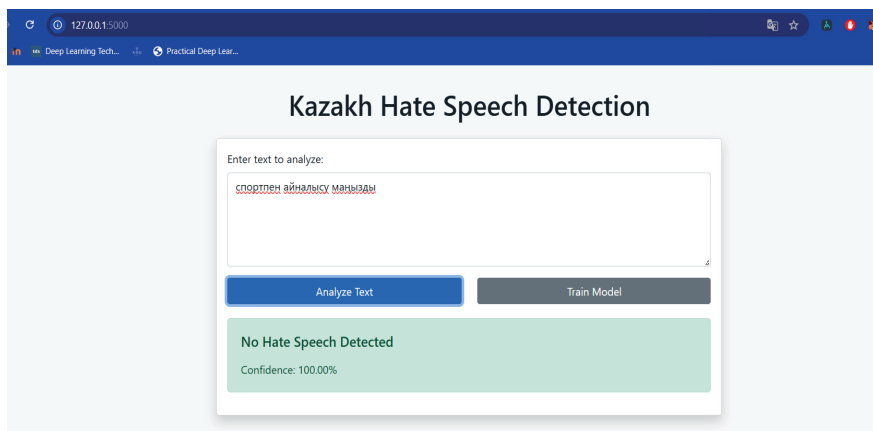


Figure 4 - Web interface of hate speech detection (not hate speech)

Figure 4 illustrates the processing of a neutral statement ("sportpen ainalysu mangyzdy"), which the model classifies as "Not hate speech" with a correspondingly high probability. The results indicate that the proposed model exhibits high accuracy rates on test samples and possesses practical applicability for integration into automatic online content moderation systems.

The authors developed and tested a hybrid deep learning system to automatically detect hate speech in Kazakh-language web content. The suggested model has good accuracy, F1-score, and AUC-ROC values of 92%, 92%, and 95%, respectively, suggesting its stability and capacity to identify offensive from neutral phrases.

Comparing classical machine learning algorithms (LR, RF, NB, KNN, SVM) and deep learning models (CNN, LSTM, Bi-LSTM, RNN), the hybrid architecture

achieves the best accuracy, completeness, and classification balance.

Real-time text message analysis was demonstrated by "Kazakh Hate Speech Detection," an online application. Empirical testing showed that the system can detect offensive and neutral messages, proving its applicability for online moderation.

The results show that the proposed approach can create a safe digital environment in Kazakh-language Internet.

Discussion. The study found hybrid deep learning architectures effective for automatically detecting hate speech in Kazakh-language web content. Comparative investigation indicated that combined Bi-LSTM and convolutional CNN layers enhance contextual data retrieved by the recurrent network and local patterns recognized by convolutional filters. This method detects overt and covert verbal aggressiveness more accurately than typical machine learning models. High Accuracy, F1-score, and AUC-ROC values on three language corpora (Kazakh, Russian, English) demonstrate the architecture's adaptability to multilingual data and class imbalance resistance. The study's "Kazakh Hate Speech Detection" web interface showed the model's real-time message analysis potential. This tool can be used in social network and forum moderating systems and digital security projects.

Despite the outcomes, certain issues must be addressed. First, the training data corpus does not contain all Kazakh-language Internet sources and stylistic traits. Second, the current architecture ignores complicated interphrase connections and discourse history, which could improve hidden aggressiveness detection. Expanding the corpus to include emoji and slang language and combining attention processes and transformer architectures to increase model interpretability are promising research areas.

Conclusion. The research proposes using new deep learning approaches to automatically detect hate speech in Kazakh-language web content. After tokenization, lemmatization, and stop word removal, a representative corpus of texts was constructed and annotated. A hybrid model incorporating Bi-LSTM and CNN was implemented on this corpus to achieve good classification accuracy and data imbalance resistance.

The system was tested in real time and shown to be effective for automatic content filtering, information security, and toxicity reduction on Kazakh-language Internet. Development prospects include increasing the corpus, incorporating multimodal data, and adding adaptive training methods to develop a universal tool for online communication monitoring and analysis.

References

Huang H.Y. et al. (2020) miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. *Nucleic acids research*. — T. 48. — №. D1. — P. D148-D154 (in English).

Omarov Batyrkhan, Rustam Abdrakhmanov, and Aigerim Toktarova (2025) "Glove-embedded attention bilstm networks for enhanced multiclassification of tweets in cyberbullying detection on online content." *Scientific Journal of Astana IT University*: P.55-70 (in English).

Azhibekova Zh.Zh, et al. (2024) "Bidirectional long short-term memory in hate speech detection

problem on networks." Bulletin of Abai KazNPU. Series of Physical and Mathematical sciences 87.3: 123-131(in English).

Sachdeva R. et al. (2024) Localizing and mitigating errors in long-form question answering. arXiv preprint arXiv:2407.11930 (in English).

Tekiroglu S.S. et al. (2022) Using pre-trained language models for producing counter narratives against hate speech: a comparative study. arXiv preprint arXiv:2204.01440 (in English).

Haleem M.S., Castaldo R., Pagliara S.M., Petretta M., Salvatore M., Franzese M., & Pecchia L. (2021) Time adaptive ECG driven cardiovascular disease detector. Biomedical Signal Processing and Control. — Vol.70. — P.19-26 (in English).

Zhu F., Ye F., Fu Y., Liu Q., & Shen B. (2019) Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network. Scientific reports. — Vol. 9(1). — P. 6734 -6748 (in English).

Chen J., Huang S., Zhang Y., Chang Q., ZhangY., Li D., ... & Liang H. (2024) Congenital heart disease detection by pediatric electrocardiogram based deep learning integrated with human concepts. Nature Communications. — Vol.15(1). — P. 976-985 (in English).

Naseer A., Khan M.M., Arif F., Iqbal W., Ahmad A., & Ahmad I. (2025) An improved hybrid model for cardiovascular disease detection using machine learning in IoT. Expert Systems. —Vol. 42(1). — e13520 (in English).

Sharifi S., & Donyadadi A. (2025) Detection and Diagnosis of Congenital Heart Disease from Chest X-Rays with Deep Learning Models. International Journal of Applied Data Science in Engineering and Health. — Vol. 1(1). — P. 1-9 (in English).

Bolatbek M. et al. Kazakh Language Dataset for Hate Speech Detection on Social Media Text. 2024 IEEE 9th International Conference on Computational Intelligence and Applications (ICCIA). – IEEE, 2024. — P. 94-98 (in English).

Mussiraliyeva S. et al. Bigram based deep neural network for extremism detection in online user generated contents in the kazakh language. International Conference on Computational Collective Intelligence. – Cham. Springer International Publishing, 2021. — P. 559-570 (in English).

Ojo O.E. et al. Automatic hate speech detection using deep neural networks and word embedding. Computación y Sistemas, 2022. — Vol. 26. — No. 2. — P. 1007-1013 (in English).

Guo X., Adnan H. M., Abidin M. Z. Z. A Scientometric Analysis of Potentially Harmful Speech Detection in Muslim Communities: Trends, Challenges, and Future Directions. Society, 2025. — P. 1-21(in English).

Zhang Z., Robinson D., Tepper J. Detecting hate speech on twitter using a convolution-gru based deep neural network. European semantic web conference. — Cham: Springer International Publishing, 2018. — P. 745-760 (in English).

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.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

www.nauka-nanrk.kz

<http://physics-mathematics.kz/index.php/en/archive>

ISSN2518-1726 (Online),

ISSN 1991-346X (Print)

Ответственный редактор *А. Ботанқызы*

Редакторы: *Д.С. Аленов, Т. Апендиев*

Верстка на компьютере: *Г.Д. Жадырановой*

Подписано в печать 22.12.2025.

Формат 60x881/8. Бумага офсетная.

Печать –ризограф. 20,0 п.л. Заказ 4.