

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

3 (359)

JULY – SEPTEMBER 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 Жұмаділ 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>

СМОЛАРЖ 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>

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 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 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

Akhmetova A.Zh., Kantureyeva M.A., Serimbetov B.A., Zhasuzakova M.Zh., Mustafayeva N.T. Identification of communities and active users in social networks.....	11
Akzhigitov D., Mukhtarkhanov M., Zhetpisbayeva A., Kulakayeva A., Zhetpisbayev K. Analysis of the mechanical structure of a 12U CubeSat nanosatellite and assessment of its mechanical stability.....	29
Alibekov A., Nurimanov A., Mukhametkali A., Kassym D., Varol H.A. Qolda-Math: a small vision-language model for mathematical reasoning in Kazakh.....	50
Alisheva S.S., Ospanova A.O., Shaimerdenova G.S., Shaimerdenova L., Zhidebayeva A.N. Detecting Kazakh language offensive comments in online content using deep learning algorithms.....	68
Askar G., Mutaikhan A., Aitkozhasyeva M., Nauryzbaikyzy G., Makhataeva Z. Retrieval beats fine-tuning: adapting a Kazakh small language model for historical question answering.....	86
Beissenova G.I., Shynzhigit B.B., Beisebayeva A.Zh., Kaldarova B.S., Lesbayev A.U. Deep neural network-based speaker authentication for secure access control systems.....	106
Bekenuly B., Baimakhanova A., Baiturganova V., Bolatbekova D. Comparison of CNN architectures for automated ACL injury classification from knee MRI.....	123
Bekbolat A., Razak A., Alimbayeva Z., Ozhikenova A. Multi-task deep learning for knee MRI pathology classification.....	139
Issakhov A., Akylbekkyzy A., Abylkassymova A. Predicting the spread of airborne viral infections in enclosed spaces.....	158
Kisseleva O., Abdramanov K., Naumenko V., Mukanova Z., Maintser D. Evaluation of serverless cold start optimization techniques.....	177
Koshkarov Ye.Zh., Amirgaliyev B.Ye., Mulesa O. Physics-aware masked pretraining for cross-basin well log foundation models.....	197
Kurmetkan T., Turysbek Zh., Mukhitova A., Nauryzbaeva A. Research on methods for reducing knowledge conflicts in large language models in Kazakh question-and-answer systems.....	214
Makatov Ye.K., Abdul Razaque, Ismukanova A.N., Salykova A.K., Glock E.S. Machine learning-based model for phishing URL detection in online social networks.....	233
Mimenbayeva A.B., Naizagarayeva A.A., Ospanova T.T., Koikelova D.K., Kudabayeva R.N. Spatial and climate data analysis for machine learning-based crop yield prediction in Northern Kazakhstan.....	255
Mukhamediev R.I., Abdoldina F.N., Gertsen Y.A. Economic opportunities and constraints of artificial intelligence adoption in Kazakhstan's resource-oriented economy.....	274

Mutanov G.M., Karyukin V.I., Ussatova O.A., Begimbayeva Y.Y., Baispay G.B. The assessment of blockchain readiness for academic document verification.....	289
Mussiraliyeva Sh., Bolatbek M., Baisylbayeva K. Classification of cyber-extremist ideologies in the Kazakh language considering multi-annotator disagreement.....	306
Omarov B.S., Sultan D.R., Rakhymberdiev D.S., Abibullayeva A.A., Toktarova A.B. Deformable attention-guided Transformer for brain tumor segmentation.....	326
Oralbekova D., Bekarystankyzy A., Yerimbetova A., Berzhanova U. Development of a streaming Kazakh speech recognition system using end-to-end models.....	346
Serek A., Shoiynbek A., Zheludkov M., Kuanyshbay D., Meraliyev B. Comparative analysis of machine learning methods for detecting fraud in text messages.....	365
Smakhanova A.K., Uskenbayeva R.K., Kassymova A.B., Kalpeyeva Zh.B., Konyrbaev N.B. Experimental research using models and methods of the university digital ecosystem.....	384
Sultanali A.E., Kazbekova G.N. Machine learning-based detection of AI-generated content in students' written works.....	401
Tashenova Zh., Yergazin A., Nurlybaeva E., Abdugulova Zh., Amanzholova Sh. Experimental evaluation of a hybrid IDPS with adaptive learning for low-latency 5G networks using edge computing.....	418
Telman D., Yerimbetova A., Daiyrbayeva E. Leakage-resistant fake news detection for Kazakh: temporal evaluation of lexical, Transformer and hybrid models.....	441
Zhanegizov A., Mamyrbayev O., Abdigappar M. Application of YOLO12 for road defect detection in GPR images.....	457

МАЗМҰНЫ

Ахметова А.Ж., Кантуреева М.А., Серимбетов Б.А., Жасұзақова М.Ж., Мұстафаева Н.Т. Әлеуметтік желілердегі қауымдастықтар мен белсенді қолданушыларды анықтау.....	11
Ақжігітов Д.Ж., Мұхтарханов М.О., Жетписбаева А.Т., Кулакаева А.Е., Жетписбаев К.У. 12U CubeSat үлгісіндегі наноспутниктің механикалық құрылымын талдау және оның механикалық орнықтылығын бағалау.....	29
Алибеков А., Нуриманов А., Мұхаметқали А., Қасым Д., Варол Х.А. Qolda-Math: қазақ тіліндегі математикалық пайымдауға арналған кіші көру-тілдік модель.....	50
Алишева С.С., Оспанова А.О., Шаймерденова Г.С., Шаймерденова Л., Жидебаева А.Н. Онлайн-контенттегі қазақ тіліндегі бейәдеп пікірлерді терең оқыту алгоритмдері арқылы анықтау.....	68
Асқар Ғ., Мұтайхан А., Айтқожаева М., Наурызбайқызы Г., Махатаева Ж. Retrieval fine-tuning-тен басым: тарихи сұрақтарға жауап беру үшін қазақша шағын тілдік модельді бейімдеу.....	86
Бейсенова Г.И., Шынжігіт Б.Б., Бейсебаева А.Ж., Қалдарова Б.С., Лесбаев А.У. Терең нейрондық желіге негізделген қауіпсіз басқару жүйелерінде сөйлеушіні аутентификациялау.....	106
Бекенұлы Б., Баймаханова А., Байтурганова В., Болатбекова Д. Алдыңғы крест тәрізді байлам жарақатын анықтауға арналған CNN архитектураларын салыстыру.....	123
Бекболат А., Разак А., Алимбаева Ж., Ожикенова А. Tize MPT патологиясын жіктеуге арналған көп тапсырмалы терең оқыту.....	139
Исahов А.А., Ақылбекқызы А., Абылкасымова А. Шектеулі кеңістіктерде әуе тамшыларымен вирустық инфекцияның таралуын сандық модельдеу.....	158
Киселева О., Абдраманов Қ., Науменко В., Мұқанова Ж., Майнцер Д. Serverless жүйелеріндегі суық іске қосуды оңтайландыру әдістерін эксперименттік бағалау.....	177
Кошкарлов Е.Ж., Амиргалиев Б.Е., Мүлеса О. Бассейнаралық ұңғымалық диаграммалардың іргелі модельдерін физикалық тұрғыдан негізделген масқаланған алдын ала оқыту.....	197
Құрметқан Т., Тұрысбек Ж., Мұхитова А., Наурызбаева А. Қазақ тіліндегі сұрақ-жауап жүйелерінде үлкен тілдік модельдердегі білім қайшылықтарын азайту әдістерін зерттеу.....	214
Мақатов Е.К., Abdul Razaque, Исмуканова А.Н., Салыкова А.К., Глюк Е.С. Онлайн әлеуметтік желілердегі фишингтік URL-дарды анықтауға арналған машиналық оқыту негізіндегі модель.....	233

Мименбаева А.Б., Найзағараева А.А., Оспанова Т.Т., Қойкелова Д.К., Қудабаева Р.Н. Солтүстік Қазақстандағы дақылдар өнімділігін машиналық оқыту арқылы болжауда кеңістіктік және климаттық деректерді талдау.....	255
Мухамедиев Р.И., Абдолдина Ф.Н., Герцен Е.А. Қазақстанның ресурстық бағытталған экономикасына жасанды интеллектті енгізудің экономикалық мүмкіндіктері мен шектеулері.....	274
Мутанов Г.М., Карюкин В.И., Усатова О.А., Бегимбаева Е.Е., Байспай Г.Б. Білім беру құжаттарының түпнұсқалығын тексеруге блокчейн технологиясының дайындығын бағалау.....	289
Мусиралиева Ш., Болатбек М., Байсылбаева К. Көпаннынотаторлы келіспеушілікті ескере отырып қазақ тіліндегі киберэкстремистік идеологияларды жіктеу.....	306
Омаров Б.С., Сұлтан Д.Р., Рахымбердиев Д.С., Абибуллаева А.А., Тоқтарова А.Б. Ми ісіктерін сегменттеуде деформацияланатын зейін механизміне негізделген мультимодальды трансформер.....	326
Оралбекова Д.О., Бекарыстанқызы А., Еримбетова А.С., Бержанова У.Г. Интегралды модельдерді қолдану арқылы қазақ тіліндегі ағындық сөйлеуді тану жүйесін әзірлеу.....	346
Серек А., Шойынбек А., Желудков М., Қуанышбай Д., Мералиев Б. Мәтіндік хабарламалардағы алаяқтықты анықтауға арналған машиналық оқыту әдістерінің салыстырмалы талдауы.....	365
Смаханова А.К., Ускенбаева Р.К., Касымова А.Б., Кальпеева Ж.Б., Қоңырбаев Н.Б. Университеттің цифрлық экожүйесінің модельдері мен әдістерін қолдана отырып жүргізілген эксперименттік зерттеу.....	384
Сұлтанәлі А.Е., Казбекова Г.Н. Білім алушылардың жазбаша жұмыстарындағы AI-контентті машиналық оқыту арқылы анықтау.....	401
Ташенова Ж.М., Ергазин Ә., Нұрлыбаева Ә.Н., Абдуғулова Ж.К., Аманжолова Ш.А. Edge Computing көмегімен төмен кідірісті 5G желілеріне арналған бейімделгіш оқытуы бар гибриді IDPS-ті эксперименттік бағалау.....	418
Тельман Д., Еримбетова А., Дайырбаева Ә. Қазақ тіліндегі жалған жаңалықтарды анықтау: лексикалық, трансформерлік және гибридті модельдерді уақыт бойынша бағалау.....	441
Жанегизов А.С., Мамырбаев Ә.Ж., Абдіғаппар М.Д. YOLO12 моделін георадар кескіндеріне қолдану.....	457

СОДЕРЖАНИЕ

Ахметова А.Ж., Кантуреева М.А., Серимбетов Б.А., Жасузакова М.Ж., Мустафаева Н.Т. Выявление сообществ и активных пользователей в социальных сетях.....	11
Акжигитов Д.Ж., Мухтарханов М.О., Жетписбаева А.Т., Кулакаева А.Е., Жетписбаев К.У. Анализ механической конструкции наноспутника модели 12U CubeSat и оценка его механической устойчивости.....	29
Алибеков А., Нуриманов А., Мухаметкали А., Касым Д., Варол Х.А. Qolda-Math: маленькая мультимодальная модель зрения и языка для математических рассуждений на казахском языке.....	50
Алишева С.С., Оспанова А.О., Шаймерденова Г.С., Шаймерденова Л., Жидебаева А.Н. Выявление оскорбительных комментариев на казахском языке в онлайн-контенте с помощью алгоритмов глубокого обучения.....	68
Аскар Г., Мутайхан А., Айткожаева М., Наурызбайкызы Г., Махатаева Ж. Retrieval эффективнее fine-tuning: адаптация малой языковой модели казахского языка для ответов на исторические вопросы.....	86
Бейсенова Г.И., Шынжигит Б.Б., Бейсебаева А.Ж., Калдарова Б.С., Лесбаев А.У. Аутентификация личности по голосу на основе глубокой нейронной сети для систем безопасного управления доступом.....	106
Бекенулы Б., Баймаханова А., Байтурганова В., Болатбекова Д. Сравнение архитектур сверточных нейронных сетей для автоматической классификации разрыва ПКС по МРТ коленного сустава.....	123
Бекболат А., Разак А., Алимбаева Ж., Ожикенова А. Многозадачное глубокое обучение для классификации патологий колена по МРТ.....	139
Исахов А.А., Акылбеккызы А., Абылкасымова А. Прогнозирование распространения вирусной инфекции воздушно-капельным путем в замкнутых помещениях.....	158
Киселева О., Абдраманов К., Науменко В., Муканова Ж., Майнцер Д. Экспериментальная оценка методов оптимизации холодного запуска в serverless-системах.....	177
Кошкарар Е.Ж., Амиргалиев Б.Е., Мулеса О. Физически-осознанное маскированное предобучение для межбассейновых фундаментальных моделей каротажных диаграмм.....	197
Курметкан Т., Турысбек Ж., Мухитова А., Наурызбаева А. Исследование методов снижения конфликтов знаний в больших языковых моделях казахских вопросно-ответных систем.....	214
Макатов Е.К., Abdul Razaque, Исмуканова А.Н., Салыкова А.К., Глок Е.С. Модель обнаружения фишинговых URL в онлайн-социальных сетях на основе машинного обучения.....	233

Мименбаева А.Б., Найзагараева А.А., Оспанова Т.Т., Койкелова Д.К., Кудабаяева Р.Н. Анализ климатических данных для прогнозирования урожайности пшеницы методами машинного обучения в Северном Казахстане.....	255
Мухамедиев Р.И., Абдолдина Ф.Н., Герцен Е.А. Экономические возможности и ограничения внедрения искусственного интеллекта в ресурсно-ориентированной экономике Казахстана.....	274
Мутанов Г.М., Карюкин В.И., Усатова О.А., Бегимбаева Е.Е., Байспай Г.Б. Оценка готовности технологии блокчейн к проверке подлинности документов об образовании.....	289
Мусиралиева Ш., Болатбек М., Байсылбаева К. Классификация киберэкстремистских идеологий на казахском языке с учетом разногласий нескольких аннотаторов.....	306
Омаров Б.С., Султан Д.Р., Рахымбердиев Д.С., Абибуллаева А.А., Токтарова А.Б. Мультимодальный трансформер на основе механизма деформируемого внимания для сегментации опухолей головного мозга.....	326
Оралбекова Д.О., Бекарыстанкызы А., Еримбетова А.С., Бержанова У.Г. Распознавание казахской речи с применением интегральных моделей.....	346
Серек А., Шойынбек А., Желудков М., Куанышбай Д., Мералиев Б. Сравнительный анализ методов машинного обучения для выявления мошенничества в текстовых сообщениях.....	365
Смаханова А.К., Ускенбаева Р.К., Касымова А.Б., Кальпеева Ж.Б., Конырбаев Н.Б. Экспериментальное исследование с использованием моделей и методов цифровой экосистемы университета.....	384
Султанали А.Е., Казбекова Г.Н. Обнаружение AI-контента в письменных работах обучающихся методами машинного обучения.....	401
Ташенова Ж.М., Ергазин А., Нурлыбаева Э.Н., Абдугулова Ж.К., Аманжолова Ш.А. Экспериментальная оценка гибридной системы обнаружения и предотвращения вторжений (IDPS) с адаптивным обучением для сетей 5G с низкой задержкой с использованием граничных вычислений.....	418
Тельман Д., Еримбетова А., Дайырбаева Э. Обнаружение фейковых новостей на казахском языке: временная оценка лексических, трансформерных и гибридных моделей.....	441
Жанегизов А.С., Мамырбаев О.Ж., Абдигаппар М.Д. Применение модели YOLOv12 для анализа данных георадара.....	457

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 3.
Number 359 (2026). 441–456

<https://doi.org/10.32014/2026.2518-1726.472>

IRSTI 20.23.27

UDC 004.912

© **Telman D.^{1,2}, Yerimbetova A^{1,3}, Daiyrbayeva E.^{1,2*}**, 2026.

¹Institute of Information and Computational Technologies, Almaty, Kazakhstan;

²Satbayev University, Almaty, Kazakhstan;

³META University, Astana, Kazakhstan.

E-mail: nurbekkyzyelmira@gmail.com

LEAKAGE-RESISTANT FAKE NEWS DETECTION FOR KAZAKH: TEMPORAL EVALUATION OF LEXICAL, TRANSFORMER AND HYBRID MODELS

Telman Duman — PhD student, Satbayev University, Astana, Kazakhstan; junior research assistant, Institute of Information and computing technologies, Almaty, Kazakhstan,
E-mail: telmand28@gmail.com, <https://orcid.org/0000-0003-1387-5351>;

Yerimbetova Aigerim — Leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD, Cand. Tech. Sci., Associate Professor, Research Professor, META University, Almaty, Kazakhstan,

E-mail: aigerian8888@gmail.com, <https://orcid.org/0000-0002-2013-1513>;

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. This study compares lexical, monolingual transformer and score-fusion classifiers on cleaned Kazakh-language titles and leads from Factcheck.kz. Explicit verdict expressions and source boilerplate were removed, and 1,265 records were split chronologically into 885 training, 190 validation and 190 test records. The operational positive class combines false, fake, manipulation, half-true and unproven verdicts; it is not a pure fabrication label. The TF-IDF + Linear SVM achieved macro-F1 = 0.6563, with a wide 95% bootstrap confidence interval of [0.5834, 0.7253]. The transformer achieved mean macro-F1 = 0.6585 (SD = 0.0442) across five seeds. The experimental score-fusion baseline achieved 0.6422 (SD = 0.0180) and failed the pre-specified requirement to exceed both component models in mean test macro-F1. Paired bootstrap intervals for the differences between these three approaches include zero; the results do not establish a superior model. In a post hoc sensitivity analysis restricting class 1 to fake and manipulation, a retrained SVM achieved macro-F1 = 0.7153 on 123 test records. The original broad-label SVM scored 0.7488 on the same subset, so the narrower

experiment does not demonstrate a benefit from retraining. Evaluation remains confined to one editorial source: chronological separation and text cleaning do not establish generalization to other publishers or original claims. The findings support cautious interpretation of small-data classification results; use as a triage aid requires independent external validation and human review.

Keywords: fake news detection, misinformation, Kazakh language, low-resource NLP, support vector machine, model fusion, data leakage

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. AP23484329).*

For citations: Telman D., Yerimbetova A., Daiyrbayeva E. Leakage-resistant fake news detection for kazakh: temporal evaluation of lexical, transformer and hybrid models. Academic Scientific Journal of Computer Science, 2026. — No.3. — P. 441–456. DOI: <https://doi.org/10.32014/2026.2518-1726.472>

© Тельман Д.^{1,2}, Еримбетова А.^{1,3}, Дайырбаева Э.^{1,2*}, 2026.

¹ ҚР Ғылым және жоғары білім министрлігі Ғылым комитетінің Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан;

² Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан;

³ META University, Алматы, Қазақстан.

E-mail: nurbekkyzyelmira@gmail.com

ҚАЗАҚ ТІЛІНДЕГІ ЖАЛҒАН ЖАҢАЛЫҚТАРДЫ АНЫҚТАУ: ЛЕКСИКАЛЫҚ, ТРАНСФОРМЕРЛІК ЖӘНЕ ГИБРИДТІ МОДЕЛЬДЕРДІ УАҚЫТ БОЙЫНША БАҒАЛАУ

Тельман Думан — PhD докторант, Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті; кіші ғылыми қызметкер ҚР ҒЖБМ Ғылым комитетінің Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан,

E-mail: telmand28@gmail.com, <https://orcid.org/0000-0003-1387-5351>;

Еримбетова Айгерім — жетекші ғылыми қызметкер, ҚР ҒЖБМ Ғылым комитетінің Ақпараттық және есептеу технологиялары институты; PhD, техника ғылымдарының кандидаты, қауымдастырылған профессор, профессор-зерттеуші, META University, Алматы, Қазақстан,

E-mail: aigerian8888@gmail.com, <https://orcid.org/0000-0002-2013-1513>;

Дайырбаева Эльмира — ғылыми қызметкер, ҚР ҒЖБМ Ғылым комитетінің Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан; магистр, аға оқытушы, Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан, E-mail: nurbekkyzyelmira@gmail.com, <https://orcid.org/0000-0002-4255-5456>.

Аннотация. Зерттеуде Factcheck.kz материалдарының тазартылған қазақ тіліндегі тақырыптары мен қысқаша сипаттамалары негізінде лексикалық,

монотілді Transformer және ықтималдық бағаларын біріктіруге негізделген модельдер салыстырылады. Вердиктті тікелей білдіретін сөздер мен дереккөзге тән шаблондық тіркестер алынып тасталып, 1 265 жазба уақыт тәртібімен 885 оқыту, 190 валидация және 190 тест жазбасына бөлінді. Операциялық 1-класс «жалған», «фейк», «манипуляция», «жартылай шындық» және «дәлелденбеген» санаттарын біріктіреді, сондықтан ол тек ойдан шығарылған ақпаратты білдіретін біртекті класс емес. TF-IDF + Linear SVM моделі $\text{macro-F1} = 0,6563$ нәтижесін көрсетті, ал 95% bootstrap сенімділік аралығы $[0,5834; 0,7253]$ болды. Transformer моделі бес random seed бойынша орташа $\text{macro-F1} = 0,6585$ ($\text{SD} = 0,0442$) нәтижесіне жетті. Эксперименттік score-fusion моделі $0,6422$ ($\text{SD} = 0,0180$) көрсеткішін көрсетті және тестілік жиындағы орташа macro-F1 бойынша екі құрамдас модельден де жоғары нәтиже көрсетуге арналған алдын ала белгіленген критерийді орындамады. Үш тәсіл арасындағы айырмашылықтарға арналған жұпталған bootstrap сенімділік аралықтары нөлді қамтыды, сондықтан қандай да бір модельдің басымдығы дәлелденген жоқ. Қосымша сезімталдық талдауында 1-класс тек «фейк» және «манипуляция» санаттарымен шектелді. Қайта оқытылған SVM 123 тест жазбасында $\text{macro-F1} = 0,7153$ нәтижесіне жетті, ал кеңейтілген бастапқы таңбалары бар SVM сол ішкі жиында $0,7488$ көрсетті. Демек, тар санаттар бойынша қайта оқытудың артықшылығы дәлелденбеді. Бағалау бір ғана редакциялық дереккөзбен шектелгендіктен, уақыт бойынша бөлу және мәтінді тазарту нәтижелердің басқа басылымдарға немесе бастапқы тұжырымдарға жалпылануын қамтамасыз етпейді. Алынған нәтижелер шағын деректер жиынтықтарындағы жіктеу нәтижелерін сақтықпен түсіндіру қажеттігін көрсетеді. Модельдерді алдын ала сүзгілеу құралы ретінде қолдану тәуелсіз сыртқы валидацияны және сарапшылық бақылауды талап етеді.

Түйін сөздер: жалған жаңалықтарды анықтау; дезинформация; қазақ тілі; ресурсы шектеулі NLP; тірек векторлар әдісі; модельдерді біріктіру; деректердің ағып кетуі

© Тельман Д.^{1,2}, Еримбетова А.^{1,3}, Дайырбаева Э.^{1,2*}, 2026.

¹ Институт информационных и вычислительных технологий Комитета науки Министерства науки и высшего образования Республики Казахстан, Алматы, Казахстан;

² Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан;

³ META University, Алматы, Казахстан.

E-mail: nurbekkyzyelmira@gmail.com

ОБНАРУЖЕНИЕ ФЕЙКОВЫХ НОВОСТЕЙ НА КАЗАХСКОМ ЯЗЫКЕ: ВРЕМЕННАЯ ОЦЕНКА ЛЕКСИЧЕСКИХ, ТРАНСФОРМЕРНЫХ И ГИБРИДНЫХ МОДЕЛЕЙ

Тельман Думан — младший научный сотрудник, Институт информационных и вычислительных технологий Комитета науки Министерства науки и высшего образования Республики Казахстан; PhD докторант, Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан,
E-mail: telmand28@gmail.com, <https://orcid.org/0000-0003-1387-5351>;

Еримбетова Айгерим — ведущий научный сотрудник, Институт информационных и вычислительных технологий Комитета науки Министерства науки и высшего образования Республики 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>;

Дайырбаева Эльмира — научный сотрудник, Институт информационных и вычислительных технологий Комитета науки Министерства науки и высшего образования Республики Казахстан, Алматы, Казахстан; магистр, старший преподаватель, Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан,
E-mail: nurbekkyzyelmira@gmail.com, <https://orcid.org/0000-0002-4255-5456>.

Аннотация. В работе сравниваются лексическая, монологическая Transformer-модель и экспериментальная модель объединения вероятностных оценок на очищенных казахскоязычных заголовках и кратких описаниях Factcheck.kz. После удаления явных формулировок вердиктов и шаблонных фраз источника 1 265 записей были хронологически разделены на 885 обучающих, 190 валидационных и 190 тестовых. Операционный положительный класс объединяет категории «ложь», «фейк», «манипуляция», «полуправда» и «непроверено» и поэтому не представляет собой однородный класс полностью вымышленных сообщений. Модель TF-IDF + Linear SVM достигла macro-F1 = 0,6563 при широком 95%-ном bootstrap-доверительном интервале [0,5834; 0,7253]. Transformer-модель показала среднее значение macro-F1 = 0,6585 (SD = 0,0442) по пяти random seeds. Экспериментальная модель score fusion получила 0,6422 (SD = 0,0180) и не выполнила заранее установленный критерий превосходства над обоими компонентными моделями по среднему значению test macro-F1. Парные bootstrap-интервалы различий между тремя подходами включали ноль, поэтому статистически подтвержденное превосходство какой-либо модели не установлено. В

дополнительном анализе положительный класс был ограничен категориями «фейк» и «манипуляция». Переобученная SVM достигла macro-F1 = 0,7153 на 123 тестовых записях, тогда как исходная модель с широким набором меток на той же подвыборке показала 0,7488. Следовательно, преимущество переобучения на более узком наборе категорий не подтверждено. Оценка ограничена одним редакционным источником: хронологическое разделение и очистка текста сами по себе не подтверждают переносимость результатов на другие публикационные источники или исходные утверждения. Полученные результаты указывают на необходимость осторожной интерпретации результатов классификации на малых выборках. Использование модели в качестве инструмента предварительного отбора материалов требует независимой внешней валидации и экспертного контроля.

Ключевые слова: фейковые новости; дезинформация; казахский язык; малоресурсная обработка естественного языка; метод опорных векторов; объединение моделей; утечка данных

Introduction. The digital information environment of Kazakhstan is multilingual, and public statements are heard in Kazakh, Russian and multi-linguistic forms. However, research on automatic misinformation has been confined to high resource languages. This imbalance is important since a translation-based or English-based system might not capture the morphology, named entities, pragmatic features or culturally sensitive context of Kazakh. The most recent research has started to create Kazakh-specific and bilingual resources, however, evaluation of these resources is still limited by the size of data, misheterogeneity of labels and learning of editorial and source artifacts instead of claim characteristics (Telman et al., 2026; Sambetbayeva et al., 2025).

This study is focused on the difference between classification and verification. A content classifier is only an estimation of patterns that are related to fact check verdicts from the past, it is not a statement of fact and it is not a substitute for evidence collection or for people to fact-check. Previous studies have demonstrated that content alone is not enough, as deceptive claims are able to seemingly be identical to genuine reports, and that signals from social context, publisher actions, and propagation are complementary signals (Shu et al., 2017; Shu et al., 2019). Our narrower aim is to assess the ability of a leakage-controlled text-only model to classify later Kazakh-language fact-checking instances from the same publisher.

The study documents a cleaning and binary-label construction procedure, evaluates lexical and monolingual transformer classifiers on a chronological holdout, and reports variation across five transformer seeds. It also evaluates a pre-specified score-fusion baseline that fails its primary success criterion. The contribution is a transparent evaluation protocol and an account of its limitations, rather than a claim of a superior architecture. A post hoc narrower-label sensitivity analysis examines how the operational definition of class 1 affects the lexical baseline.

Literary review. The literature reviewed for this study can be segmented into misinformation detection, low-resource language modelling and sparse lexical classification. Linguistic features, classic machine learning methods, neural text encoders, social graphs and multimodal information have been investigated to detect fake news. Early surveys focused on intentional creation of false content and its subsequent challenges of detection by text alone (Shu et al., 2017). TriFN integrated relationships between publishers, news and users (Shu et al., 2019). These findings impel the need to interpret any text-only finding in a cautious manner; text-only models can be used for triage, but should not be considered as truth engines. When using low-resource or Kazakh language, the primary methodological background is transfer learning and language-specific pretraining. For example, in the case of transformer pretraining, an unsupervised task that generated contextual representations, these representations were then fine-tuned to a downstream supervised classification task using a comparatively small amount of labeled data (Devlin et al., 2019). XLM-R furthered this by scaling up the multilingual pretraining with 100+ languages and found good cross-lingual transfer for lower-resource languages (Conneau et al., 2020). Recently, monolingual LTG-BERT encoders, such as the Kazakh model, were released by the HPLT project, which may be used as a language-specific starting point for fine-tuning (HPLT Project, 2024). The papers (Telman et al., 2026) made a comparison of translation, TF-IDF with logistic regression and XLM-R in the detection of fake news in the Kazakh language in a directly relevant 2026 study. The results reported in the paper support the importance of direct Kazakh modeling, as well as the inability of a simple TF-IDF-to-XLM-R comparison to serve as a new contribution. The present work focuses on the activities of leakage auditing, temporal generalization, a monolingual Kazakh encoder and a lexical-semantic fusion architecture. A multi-level Kazakh-Russian annotation framework further demonstrates that binary labels are an operational approximation and not an adequate description of misinformation (Sambetbayeva et al., 2025). Sparse lexical models continue to serve as base models in small data misinformation detection. Societally, misinformation researchers are driven by data to demonstrate that misinformation can have a negative impact on public decision-making and can spread quickly through online networks. Research on fake news during the 2016 U.S. election provides a further motivation for studying misinformation (Allcott and Gentzkow, 2017). Lazer et al. (2018) thus conceptualize fake news as a socio-technical challenge that calls for collaboration between computational technologies and platforms, journalists, and domain experts. For the present Kazakh-language study, a small sample and a train/test split that does not reflect deployment time are important evaluation concerns. This is the reason why the present work adopts chronological evaluation and explicit leakage removal. While neural architectures have enhanced the ability of text classification systems to capture contextual representations of textual data, they have also made them more sensitive to the instability of fine-tuning and hyperparameter selection. Pretraining

proved useful for downstream language understanding tasks in the cases of BERT and XLM-R (Devlin et al., 2019; Conneau et al., 2020) and optimization details can have a significant impact on performance in the case of RoBERTa and recent fine-tuning studies (Liu et al., 2019; Mosbach et al., 2021). This encourages the reporting of more than one random seed for small Kazakh corpora, instead of a single good run. Data augmentation has also been investigated for low-resource learning (Fadaee et al., 2017). The approach used for misinformation detection, however, is prone to adding false information to the factual cues or making label-preserving assumptions that do not hold true for claims. This negative result in this study is not just due to implementation failure, but empirical warning to augment fact-checking metadata conservatively.

Last but not least, the case of a classifier if used for triage is one where the probability calibration is important. The usefulness of a model with similar macro-F1 may vary based on the level of confidence it provides. Classical calibration tasks for SVM outputs and supervised probability estimation validate the use of validation-set calibration, and Brier score reporting (Platt, 1999; Niculescu-Mizil and Caruana, 2005). The motivation for the hybrid experiment is this consideration, and the hybrid experiment is conducted using a combination of lexical and semantic probabilities after they are validated and calibrated.

TF-IDF is competitive in small text-classification tasks as it characterizes discriminative lexical patterns without estimating millions of parameters for each task (Salton and Buckley, 1988). The linear SVMs are well suited to high dimensional sparse features (Cortes and Vapnik, 1995). In addition, character n-grams can be useful for Kazakh, since they may represent affixes and orthographic variations and subword regularities without having to employ a full morphological analyzer.

Materials and methods. *Source data.* There are 2,007 records of articles published in the source collection (in Kazakh) that are related to Factcheck.kz. The records contain title, short lead, date of publication, URL of the article, metadata for category and author and an editorial verdict, if available. In the present study, only the text from the titles and leads is taken as a model input. Full fact-check explanations are not included as they often contain statements or reasons that directly result in the judgement, which cause major target leakage. Model features are also not mapped to the author, category, URL and the publication date. Documentary evidence of the project owner's permission for training and generated derived data was to be included in the submission version of the project, where applicable. The dataset filtering and leakage audit results are summarized in Table 1.

Table 1 — Dataset filtering and audit summary

Stage or issue	Count
Scraped Kazakh records	2,007
Records with missing editorial verdict	518
WordPress migration placeholders	146
Records after verdict-category filtering	1,397

Stage or issue	Count
Final trainable records after invalid-lead removal	1,265
Class 0: true or mostly-true	279
Class 1: false/fake/manipulation/half-true/unproven	986

Label construction. Editorial verdicts were mapped to an operational binary target. Class 0 contains true and mostly-true; class 1 contains false, fake, manipulation, half-true and unproven. Records without usable verdicts and records labelled technical-error, satire or scam were excluded. Class 1 denotes a heterogeneous set of adverse or unresolved editorial verdicts, not a uniform class of fabricated claims. In particular, lack of verification does not establish falsity, and manipulation and partial truth differ from fabrication. Combining these categories may obscure distinct linguistic patterns and reduce classification quality, but the present design does not isolate that causal effect. As a post hoc sensitivity analysis, class 0 was retained and class 1 was restricted to the source categories fake and manipulation. The separate false category, half-true and unproven were excluded rather than reassigned; the original chronological partition membership was preserved.

Leakage control. Explicit verdict terms were discovered in 448 titles and fact-checking or source boilerplate in hundreds of descriptions in an audit. We eliminated Kazakh verdict expressions, reference to Factcheck.kz, the entire factcheck/фактчек word family and frequently used phrases introducing that a claim was checked. The placeholder 'Материал импортирован из WordPress. Нужен ручной лид.' appeared 146 times in descriptions - these records were not included in the training-ready data set. The last cleaned_text field includes the cleaned title and short description. No fact-checking terms were left, no verdict terms were found, no empty descriptions were found, and no duplicate cleaned texts were found.

Temporal partition. The records were divided by chronology, such that the oldest 70% were used for training, the next 15% for validation, and the newest 15% for testing. This protocol mimics deployment, where models are trained on previous topics and writing patterns, and then are faced with new topics and patterns. The hyperparameters and thresholds were determined using the validation data, and the test labels were not used in choosing a model. The chronological split and class distribution of the dataset are presented in Table 2.

Table 2 — Chronological split and class distribution

Partition	Class 0	Class 1	Total
Training	197	688	885
Validation	32	158	190
Test	50	140	190
Total	279	986	1,265

Controlled augmentation. There was an ablation of the train-only robustness augmentation. For each of the 197 class-0 training, one lowercased variant was

generated. The 145 variants of the transformed text that were different from their parent text were the result of the use of telegram-style punctuation and spacing perturbations. This resulted in 342 artificial instances and 539 class-0 and 688 class-1 instances were added to the training partition. There were no changes to the validation and test sets. For every synthetic row, an identifier for the parent, an augmentation method and a synthetic-data flag are stored. No text was found spanning a partition boundary that was written synthetically.

Experimental design. TF-IDF + Linear SVM (Experiment 1). The lexical baseline is a baseline model based on word TF-IDF features (unigrams and bigrams) and character-within-word TF-IDF features (3- to 5-grams). A sublinear term frequency, and a minimum document frequency of two are shared by both branches. The word branch is limited to 50,000 features and has a max document frequency of 0.98; the character branch is limited to 75,000 features. A Linear SVM trained using equal class weights.

The value of the regularization parameter, C , was chosen among $\{0.1, 0.5, 1, 2, 5\}$. The validation macro-F1 was selected to determine decision threshold for each candidate. The thresholds that were preferred had values around 0 and subsequent smaller C . The key metric is macro-F1, with other metrics being balanced accuracy, MCC, ROC-AUC, class-1 PR-AUC, per-class precision/recall/F1, and the confusion matrix. The 95% macro-F1 confidence interval was estimated using 10000 bootstrap samples. Original and augmented variants were compared on the same test examples using paired bootstrap resampling. The same random seed (42) and scikit-learn (Pedregosa et al., 2011) were used for all runs.

For the revised comparison, paired bootstrap intervals were also computed for differences between SVM macro-F1 and mean transformer or fusion macro-F1. Each of 10,000 item resamples used the same test indices for all models, and seed-specific macro-F1 values were averaged within each resample. These percentile 95% intervals condition on the five trained seeds and do not represent the full uncertainty of retraining on new datasets. A single-model confidence interval, or overlap between such intervals, is not by itself a test of a paired model difference.

Experiment 2. Monolingual Kazakh transformer. We fine-tuned the Kazakh LTG-BERT encoder HPLT/hplt_bert_base_kk (HPLT Project, 2024) using only clean_text. Maximum sequence length was 256, with AdamW optimization, weight decay 0.01, warmup ratio 0.1, mixed precision and an effective batch size of 8 (4 examples per device and two gradient accumulation steps). The learning rate was selected from $\{1e-5, 2e-5, 3e-5\}$ using seed-42 validation macro-F1; 2e-5 was selected. Training used class-weighted loss and early stopping with patience of two epochs. The selected setting was evaluated on original and augmented training data with seeds $\{13, 21, 42, 87, 101\}$. The stopping epoch and decision threshold for each seed were chosen using validation macro-F1 only. Reported means and standard deviations describe variation across these five runs, not uncertainty from sampling new test records.

Experiment 3. Experimental score-fusion baseline. The fusion combines Platt-

calibrated SVM probabilities with transformer class-1 probabilities: $p_{\text{hybrid}} = \alpha p_{\text{transformer}} + (1 - \alpha) p_{\text{svm}}$. It is a baseline combination of existing models, not an established superior detector. The SVM calibrator was fitted to validation decision scores and validation labels; test labels were not used. The same validation set was reused for calibration and fusion selection, which can increase selection optimism. The original-data transformer branch was retained because its mean validation macro-F1 (0.6507) exceeded that of the augmented branch (0.6422). For each seed, α was chosen from $\{0.0, 0.1, \dots, 1.0\}$, and the decision threshold was selected using validation macro-F1. Both were fixed before test evaluation. Comparisons included SVM only, transformer only, equal fusion and validation-selected fusion. Success was defined in advance as mean test macro-F1 exceeding both the SVM and transformer components; secondary metrics do not replace this criterion.

Results. Hyperparameter selection. For the original training data, $C = 0.1$ and $\text{threshold} = 0.2571041$ achieved the highest validation macro-F1 (0.6470). Larger C values reduced validation macro-F1. For the augmented training data, $C = 0.1$ and $\text{threshold} = 0.3920509$ were selected, with validation macro-F1 = 0.6193. Thus, augmentation had already failed to improve the selection metric before test evaluation. The performance of all three experiments on the chronological test set is reported in Table 3.

Table 3 — Model performance on the chronological test set

Model and training	Accuracy	Balanced acc.	Macro-F1	MCC	ROC-AUC
SVM (original)	0.6895	0.6993	0.6563	0.3549	0.7580
SVM (augmented)	0.6474	0.6836	0.6244	0.3234	0.7590
Kazakh transformer (original; mean)	0.6958	0.6959	0.6585 ± 0.0442	0.3570	0.7691
Score fusion (mean)	0.6632	0.7071	0.6422 ± 0.0180	0.3650	0.7753
Kazakh transformer (augmented; mean)	0.6737	0.6796	0.6392 ± 0.0453	0.3234	0.7558

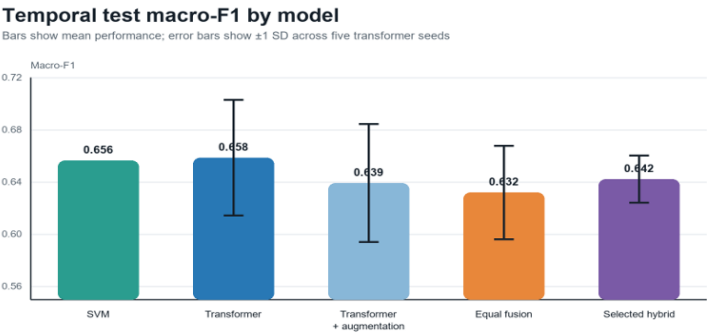


Figure 1 – Macro-F1 on the chronological test set. Error bars show one standard deviation across five trained seeds, not test-sampling confidence intervals. The deterministic SVM also has sampling uncertainty: its 95% bootstrap interval is [0.5834, 0.7253].

Per-class behavior and uncertainty. The original-data SVM correctly classified 36 of 50 class-0 records and 95 of 140 class-1 records. Class-0 precision, recall and F1 were 0.4444, 0.7200 and 0.5496; class-1 values were 0.8716, 0.6786 and 0.7631. The 95% macro-F1 bootstrap interval was [0.5834, 0.7253], a width of 0.1419. With only 190 test records, including 50 in class 0, point estimates have substantial uncertainty. The transformer-SVM mean macro-F1 difference was 0.0022, with paired 95% interval [-0.0482, 0.0536]. The corresponding intervals were [-0.0537, 0.0268] for fusion minus SVM and [-0.0513, 0.0191] for fusion minus transformer. All include zero, so the observed macro-F1 ranking does not establish superiority among these approaches. This is absence of clear evidence for a difference, not proof of equivalence. Figure 1 shows the point estimates and seed variability; Table 4 reports the SVM confusion matrix.

Table 4 — Confusion matrix for the selected original-data SVM

Actual / predicted	Class 0	Class 1
Class 0	36	14
Class 1	45	95

Augmentation ablation. The simple robustness augmentation did not lead to an improved SVM. The macro-F1 value was found to have dropped by about 0.0321. The paired bootstrap 95% interval for augmented minus original macro-F1 was (-0.0685, 0.0031) and the estimated probability that augmentation was better was 0.0366. The difference may not be considered to be definitely significant since the interval does include zero, but the evidence seen does favor the original data baseline. The majority of lowercase examples for a TF-IDF pipeline that is set to lowercase the examples is pretty much redundant, and they will have the general effect of oversampling the minority examples, but not representatively enriching them. The confusion matrix for the selected original-data SVM model is shown in Table 4.

Kazakh transformer results. The learning-rate search selected $2e-5$: validation macro-F1 was 0.6315 for $1e-5$, 0.6699 for $2e-5$, and 0.6506 for $3e-5$. With original training data, test macro-F1 values for seeds 13, 21, 42, 87, and 101 were 0.6245, 0.6927, 0.5989, 0.6977, and 0.6787, respectively. The mean was 0.6585 (SD = 0.0442), with mean balanced accuracy 0.6959, MCC 0.3570, ROC-AUC 0.7691, and class-1 PR-AUC 0.9022. Mean class-0 recall was 0.6960. The mean gain of 0.0022 over SVM is small relative to both seed variability and test-sampling uncertainty. The paired interval reported above does not establish transformer superiority.

The same five-seed protocol with augmented training produced mean macro-F1 = 0.6392 (SD = 0.0453), balanced accuracy = 0.6796, MCC = 0.3234, and ROC-AUC = 0.7558. The variation of transformer macro-F1 scores across different random seeds is presented in Figure 2. Neither model family showed a higher mean test macro-F1 with this augmentation; these point estimates do not establish that augmentation is generally harmful.

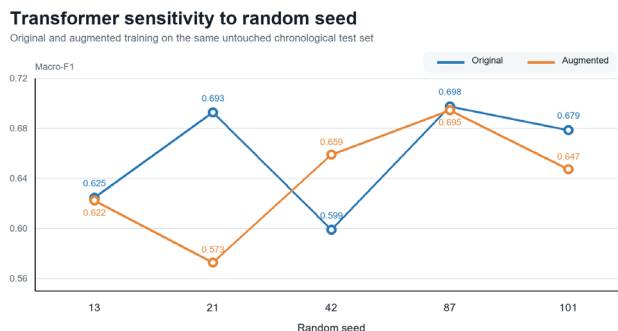


Figure 2 — Transformer macro-F1 by random seed for original and augmented training data

Score-fusion results. Validation-selected alpha values were 0.2, 0.5, 0.3, 0.3 and 0.3 for seeds 13, 21, 42, 87 and 101. The fusion baseline achieved mean macro-F1 = 0.6422 (SD = 0.0180), below SVM (0.6563) and the transformer mean (0.6585). It therefore failed the pre-specified success criterion. Mean balanced accuracy was 0.7071, MCC 0.3650, ROC-AUC 0.7753 and Brier score 0.1605. These secondary point estimates are descriptive and do not establish a general advantage or overturn the primary negative finding. Mean class-0 recall increased to 0.8000 while class-1 recall fell to 0.6143, indicating a change in the error trade-off. Equal fusion achieved mean macro-F1 = 0.6319. The paired macro-F1 intervals include zero for comparisons with both components.

Narrower-label sensitivity analysis. Restricting class 1 to fake and manipulation retained 693 records: 449 training (197 class 0, 252 class 1), 121 validation (32, 89) and 123 test (50, 73). A new TF-IDF vocabulary and SVM were fitted only on the retained training records, using the same feature settings and C grid as Experiment 1. The retained validation set selected C = 0.1 and threshold = 0.1981114 (validation macro-F1 = 0.7705). On the retained chronological test set, macro-F1 was 0.7153, with 95% bootstrap interval [0.6315, 0.7923], balanced accuracy 0.7167 and ROC-AUC 0.8126. The confusion matrix was [[34, 16], [18, 55]] in class order [0, 1]. For comparison, the original broad-label SVM, with its original predictions and threshold unchanged, achieved macro-F1 = 0.7488 (95% interval [0.6674, 0.8225]) on exactly these 123 records. Thus the higher score relative to 0.6563 on all 190 records cannot be attributed to narrower-label training; the evaluated population and class proportions also changed. This post hoc baseline analysis does not prove that label pooling caused lower performance. Transformer and fusion models were not retrained on the narrow definition, and the original primary results remain unchanged.



Figure 3 — Mean class recall for the component models and validation-selected hybrid

Discussion. The models show a limited association between cleaned text and editorial verdicts within Factcheck.kz. The broad SVM confidence interval and the paired intervals spanning zero prevent a firm ranking of the lexical, transformer and fusion approaches by macro-F1. Five seeds reveal training variability but do not enlarge the 190-record test set. The secondary ranking and calibration metrics are descriptive; their numerical differences should not be presented as established superiority. These results concern later titles and leads from the same editorial source, not independent truth verification or transfer to other publishers.

The operational label mapping is an additional source of uncertainty. Fabrication, manipulation, partial truth and unresolved verification are different tasks, and combining them may dilute learnable patterns. The narrower-label analysis makes this dependency visible but changes both training size and the evaluated population; it cannot isolate the effect of semantic heterogeneity. Moreover, a text-only model lacks the external evidence needed to verify new numerical or time-sensitive claims. Any future triage application should expose uncertainty and support evidence retrieval and human review.

Augmentation and score fusion are reported as negative primary results. Surface perturbations did not increase mean test macro-F1, and fusion did not meet its pre-specified criterion. Its higher class-0 recall came with lower class-1 recall. The study therefore treats fusion as an experimental baseline rather than a validated new detection method. Further model changes should be selected on development data and evaluated with uncertainty estimates on an independent holdout.

Limitations and ethical considerations. The most important external-validity limitation is the single source, Factcheck.kz. Titles and leads reflect its editorial selection, summarization, vocabulary and topic preferences rather than necessarily reproducing original claims verbatim. Removing explicit verdicts and boilerplate cannot remove all these cues. Chronological separation tests transfer across time within this source; it does not test transfer across publishers. No independent second-source evaluation of the present models is reported, so source-independent generalization remains unestablished. A suitable external test should contain both

target classes within an independent source, use compatible verdict definitions and input construction, and exclude duplicate claims or events shared with development data. Models, calibration and thresholds must remain fixed during that evaluation. Further limitations include the heterogeneous binary target, the small minority class, reuse of validation data for calibration and model selection, and the absence of images, propagation information or retrieved evidence. Editorial verdicts are not annotations of deceptive intent.

The automated output can be problematic if it is used to make definitive statements that may result in damage to a person's reputation. Deployment should raise uncertainty, should contain evidence links, should be open to human inspection and should have a 'not enough information' outcome. Only the verified permission and copyright of the data and derived artefacts should be shared. For a reproducibility release, IDs, URLs, cleaning scripts, split definitions, aggregate results and prediction files should be preferred, as long as redistribution of text is explicitly allowed.

Additional analyses and future work. The narrower-label experiment is a post hoc sensitivity analysis prompted by review, not a pre-specified confirmatory result. The next priorities are an independently annotated, source-balanced external holdout and replication of the narrower definition for transformer and fusion models. Raw versus cleaned input, random versus chronological splitting, and structured manual review of errors would help distinguish leakage, temporal shift and label ambiguity. These analyses should be reported separately from the original primary comparison.

Future deployment research should evaluate evidence retrieval, calibrated abstention and multilingual routing on separately validated data. The present experiment does not establish performance for a bilingual Kazakh–Russian Telegram system. More elaborate augmentation or ensembling requires evidence of benefit on independent, untouched evaluation data, with model-selection decisions separated from final testing.

Conclusion. This study provides a chronological comparison of text classifiers on cleaned Kazakh-language Factcheck.kz metadata. The SVM macro-F1 was 0.6563, with a wide 95% interval [0.5834, 0.7253]; the transformer mean was 0.6585 across five seeds. Paired bootstrap comparisons do not establish macro-F1 superiority among SVM, transformer and score fusion. The experimental fusion baseline failed its pre-specified primary criterion and is retained as a negative result. Restricting class 1 to fake and manipulation produced retrained SVM macro-F1 = 0.7153 on 123 records, but the original SVM scored 0.7488 on that same subset, so a benefit from narrower-label retraining is not demonstrated. The main contribution is transparent reporting of evaluation uncertainty and label sensitivity. Independent second-source validation is still required before claiming transfer beyond Factcheck.kz or suitability for operational triage.

Data and Code Availability

The project retains the cleaning and split scripts, augmentation pipeline, model runners, validation and test predictions, hybrid grids and machine-readable metric reports. The revision additionally retains the narrower-label analysis script, test predictions, dataset checksum and paired-bootstrap results. No public repository or DOI is specified for this manuscript. Redistribution of source texts and derived artefacts is subject to the applicable permissions.

References

- Allcott H., & Gentzkow M. (2017) Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-236. <https://doi.org/10.1257/jep.31.2.211> (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. In *Proceedings of ACL*. — P. 8440-8451. <https://doi.org/10.18653/v1/2020.acl-main.747> (in Eng.)
- Cortes C., & Vapnik V. (1995) Support-vector networks. *Machine Learning*, 20. — P. 273-297. <https://doi.org/10.1007/BF00994018> (in Eng.)
- Devlin J., Chang M.-W., Lee K., & Toutanova K. (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*. — P. 4171-4186. <https://doi.org/10.18653/v1/N19-1423> (in Eng.)
- Fadaee M., Bisazza A., & Monz C. (2017) Data augmentation for low-resource neural machine translation. In *Proceedings of ACL*. — P. 567-573. <https://doi.org/10.18653/v1/P17-2090> (in Eng.)
- HPLT Project. (2024) First language models trained: HPLT Deliverable 4.1. Available at: https://hplt-project.org/HPLT_D4_1_First_language_models_trained.pdf; Kazakh model card: https://huggingface.co/HPLT/hplt_bert_base_kk (in Eng.)
- Lazer D.M.J., Baum M.A., Benkler Y., Berinsky A.J., Greenhill K.M., Menczer F., Metzger M.J., Nyhan B., Pennycook G., Rothschild D., Schudson M., Sloman S.A., Sunstein C.R., Thorson E.A., Watts D.J., & Zittrain J.L. (2018) The science of fake news. *Science*, 359(6380). — P. 1094-1096. <https://doi.org/10.1126/science.aao2998> (in Eng.)
- Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., & Stoyanov V. (2019) RoBERTa: A robustly optimized BERT pretraining approach. *arXiv*. <https://doi.org/10.48550/arXiv.1907.11692> (in Eng.)
- Mosbach M., Andriushchenko M., & Klakow D. (2021) On the stability of fine-tuning BERT: Misconceptions, explanations, and strong baselines. In *Proceedings of ICLR*. Available at: <https://openreview.net/forum?id=nzpLWnVAyah> (in Eng.)
- Niculescu-Mizil A., & Caruana R. (2005) Predicting good probabilities with supervised learning. In *Proceedings of ICML*. — P. 625-632. <https://doi.org/10.1145/1102351.1102430> (in Eng.)
- Oshikawa R., Qian J., & Wang W.Y. (2020) A survey on natural language processing for fake news detection. In *Proceedings of LREC* (pp. 6086-6093). Available at: <https://aclanthology.org/2020.lrec-1.747/> (in Eng.)
- Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., & Duchesnay E. (2011) Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12. — P. 2825-2830 (in Eng.)
- Pérez-Rosas V., Kleinberg B., Lefevre A., & Mihalcea R. (2018) Automatic detection of fake news. In *Proceedings of COLING* (pp. 3391-3401). Available at: <https://aclanthology.org/C18-1287/> (in Eng.)
- Platt J.C. (1999) Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*. — P. 61-74. (in Eng.)
- Potthast M., Kiesel J., Reinartz K., Bevendorff J., & Stein B. (2018) A stylistic inquiry into hyperpartisan and fake news. In *Proceedings of ACL*. — P. 231-240. <https://doi.org/10.18653/v1/P18-1022> (in Eng.)

Salton G., & Buckley C. (1988) Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5). — P. 513-523. [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0) (in Eng.)

Sambetbayeva M., et al. (2025) A multi-level annotation model for fake news detection: Implementing Kazakh-Russian corpus via Label Studio. *Big Data and Cognitive Computing*, 9(8). — 215 p. <https://doi.org/10.3390/bdcc9080215> (in Eng.)

Shu K., Sliva A., Wang S., Tang J., & Liu H. (2017) Fake news detection on social media: A data mining perspective. *SIGKDD Explorations*, 19(1). — P. 22-36. <https://doi.org/10.1145/3137597.3137600> (in Eng.)

Shu K., Wang S., & Liu H. (2019) Beyond news contents: The role of social context for fake news detection. In *Proceedings of WSDM*. — P. 312-320. <https://doi.org/10.1145/3289600.3290994> (in Eng.)

Telman D., Yerimbetova A., Sambetbayeva M., Daiyrbayeva E., Abdygalym B., & Turganbayev A. (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.)

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: September 30, 2026

Format: 70×90 1/16. 26.5 printed sheets. Order No. 3.