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

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

№1
2026

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

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

1 (357)

JANUARY – MARCH 2026

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

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/

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/

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 КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*
Периодичность: *4 раза в год.*
http://www.physico-mathematical.kz/index.php/en/

CONTENTS

COMPUTER SCIENCE

Akhmetova S.T., Yunussova A.A., Alisheva S.S., Olzhataeva B.T., Mussirepova E.B.
Social network data mining for automated offensive language detection.....13

Amanov A.N., Kazbekova G.N., Zhunissov N.M., Abibullayeva A.A., Aben A.B.
Artificial intelligence-based intrusion detection for DDOS attacks in Software Defined Networking.....30

Amanzholova S.T., Ussatova O.A., Mutanov G.M., Mukhanov S.B., Aitmukash D.
Backend architecture of a hybrid blockchain-based academic credential verification system.....52

Amirkhanova G.A., Nurgazy T.N., Amirkhanov B.S., Tokhtassyn M.M., Nurgazy N.N.
Developing a predictive digital twin for a food product based on Edge ML and IoT sensors.....73

Bekarystankyzy A., Ussen D., Kassenkhan A., Chinibayev Y.
Cold-start in educational recommender systems: classical and LLM-Era strategies.....91

Bimoldina Zh., Mussiraliyeva Sh., Bagitova K., Tereikovska L.
Detection of cyber-propaganda content using machine learning and semantic models....106

Chezhimbayeva K.S.
Forecasting key 5G network KPIs using MLP and LSTM neural network models.....129

Dauitbayeva A.O., Konyrbaev N.B., Abildayeva Zh.T., Yessirkepova A.U., Karim N.A.
Development of an application to optimize the process of employment of graduates.....148

Dzhsupbekova G., Othman M., Ordabayeva G.
Comparative analysis of artificial intelligence algorithms to detect network attacks.....167

Issakhov A., Orazmoldayev N., Zharkynbek Y., Abylkassymova A.
Numerical modeling of the spread of viral infection by airborne droplets in confined spaces.....182

Kantureeva M., Omarova G.S., Duisen Z.D., Shekerbek A.A., Tulebayev Y.B.
Application of machine learning methods in forecasting and optimizing the processes of evacuation of people in high-rise buildings.....202

Khusain B., Telmanov M., Khusain A.B., Brodskiy A.R., Sass A.S.
Digital twin of an integrated emission purification and decarbonization system for thermal units.....218

Kulakayeva A., Ashurov A., Zhumazhanov B., Daineko Ye., Zylgara A.
Algorithm for determining the initial orbital parameters of KazeEOSat-1 for deorbiting.....236

Mimenbayeva A.B., Turebayeva R.D., Ospanova T.T., Aruova A.B., Naizagarayeva A.A.
Development and comparative analysis of machine learning models for urban traffic prediction.....253

Naumenko V.V., Mukanova Zh.A., Kiseleva O.V., Maintser D.A., Nerezov A.K.
The use of real-time polling to improve student academic performance.....271

Nazyrova A.E., Kaderkeyeva Z.K., Bekmanova G.T., Milosz M., Lamasheva Zh.
Transformation of education through digital technologies: advancing student academic performance across learning stages.....287

Oralbekova D., Mamyrbayev O., Akhmediyarova A., Kassymova D., Alibiyeva Z.
Development of a multi-level model for text summarization based on pretrained models.....316

Orazbayev B.B., Zhumadillayeva A.K., Kurbangalieva N.B., Yessirkessinov R.Zh., Orazbayeva K.N.
Synthesis of linguistic models for assessing sulfur quality and fuzzy modeling of the sulfur production process.....337

Sarsenbayeva A.K., Rakhimova D.R., Shormakova A.N., Mansurova M.E., Adali E.
Application of semantic methods in the field of legislation: an intellectual system for analysis of agglutinative texts.....354

Serek A., Shoiynbek A., Sharipov K., Kuanyshbay D., Mukhametzhanov A.
Analysis and classification of telephone fraud based on lexical features of speech transcriptions.....373

Shynzhigit B.B., Balabekova M.O., Amangeldy T.T.
Analysis and forecasting of brick product sales using machine learning models.....393

Tokhayeva A.O., Alzhanov A.K., Nezih Önal, Ziyatbekova G.Z., Begalieva K.B.
Formation of students virtualization competencies in higher education based on Proxmox VE.....412

Tukenova L.M., Auyelbekov O.A., Sapakova S.Z., Sametova A.A., Bostanov E.L.
Modelling and optimisation of hybrid power plant operating modes for unmanned aerial vehicles.....430

Yerimbetova A., Berzhanova U., Daiyrbayeva E., Sakenov B., Sambetbayeva M.
Sign language recognition using temporal convolutional network and MediaPipe.....443

Zhukabayeva T.K., Benkhelifa E., Mardenov Y.M., Baumuratova D., Karabayev N.
Decision support for responding to attacks in cyber-physical industrial internet-of-things systems.....461

МАЗМҰНЫ

ИНФОРМАТИКА

Ахметова С.Т., Юнусова А.А., Алишева С.С., Олжатаева Б.Т., Мүсірепова Э.Б.
Әлеуметтік желідегі бейәдеп пікірлерді автоматты анықтауда деректерді интеллектуалды талдау.....13

Аманов А.Н., Казбекова Г.Н., Жунисов Н.М., Абибуллаева А.А., Абен А.Б.
Бағдарламалық жасақтамамен анықталған желідегі DDOS шабуылдары үшін жасанды интеллектке негізделген шабуылдарды анықтау.....30

Аманжолова С.Т., Усатова О.А., Мутанов Г.М., Муханов С.Б., Айтмукаш Д.
Гибридтік блокчейнге негізделген академиялық сенімдік деректерді тексеру жүйесінің бекендік архитектурасы.....52

Амирханова Г.А., Нұрғазы Т.Н., Амирханов Б.С., Нұрғазы Н. Н.
EDGE ML және IOT сенсорлары негізінде азық-түлік өнімінің предиктивті цифрлық егізін әзірлеу.....73

Бекарыстанқызы А., Үсен Д., Қасенхан А., Чинибаев Е.
Білім беру саласындағы ұсынымдық жүйелеріндегі «Cold-start» мәселесі: классикалық әдістер және LLM дәуірінің стратегиялары.....91

Бимолдина Ж.А., Мусиралиева Ш.Ж., Багитова К.Б., Терейковская Л.З
Кибернасихаттық контентті анықтау үшін машиналық оқыту және семантикалық модельдер қолдану.....106

Чежимбаева К.С.
MLP және LSTM нейрондық желі модельдерін қолдана отырып, 5G желісінің негізгі KPI-лерін болжау.....129

Дәуітбаева А.О., Қоңырбаев Н.Б., Әбілдаева Ж.Т., Есіркепова А.У., Кәрім Н.Ә.
Бітіруші түлектердің жұмысқа орналастыру процесін оңтайландыру үшін қосымша әзірлеу.....148

Джусупбекова Г., Othman M., Ордабаева Г.
Жасанды интеллект алгоритмдерін желілік шабуылдарды анықтау үшін салыстырмалы талдау.....167

Исахов А.А., Оразмолдаев Н., Жаркынбек Е., Абылкасымова А.
Ауа тамшылары арқылы вирустық инфекцияның шектеулі кеңістікте таралуын сандық модельдеу.....182

Кантуреева М.А., Омарова Г.С., Дүйсен Ж.Д., Шекербек А.Ә., Тулебаев Е.Б.
Биік ғимараттардағы адамдарды эвакуациялау процестерін болжау және оңтайландыруда машиналық оқыту әдістерін қолдану.....202

Хусаин Б., Тельманов М.М., Хусаин А.Б., Бродский А.Р., Сасс А.С.
Жылу қондырғыларының шығарындыларын кешенді тазалау және
декарбонизациялау жүйесінің цифрлық егізі.....218

Кулакаева А.Е., Ашуров А.Е., Жумажанов Б.Р., Дайнеко Е.А., Зылғара А.Е.
KAZEOSAT-1 ғарыш аппаратының деорбитациясын жүзеге асыру үшін
бастапқы орбиталық параметрлерін анықтау алгоритмі.....236

Мименбаева А.Б., Туребаева А.Д., Оспанова Т.Т., Аруова А.Б.,
Найзағараева А.А.
Қалалық көлік ағынын болжауға арналған машиналық оқыту модельдерін
әзірлеу және салыстырмалы талдау.....253

Науменко В.В., Муканова Ж.А., Киселева О.В., Майнцер Д.А., Нерезов А.К.
Білім алушылардың үлгерімін арттыру үшін real-time
сауалнамаларын қолдану.....271

Назырова А.Е., Кадеркеева З.К., Бекманова Г.Т., Милош М., Ламашева Ж.Б.
Цифрлық білім және студенттердің академиялық жетістіктері:
деңгейлер бойынша білім беруді дамыту.....287

Оралбекова Д., Мамырбаев О., Ахмедиярова А., Қасымова Д.З, Алибиева Ж.,
Алдын ала оқытылған модельдер негізінде мәтінді резюмелеуге
арналған көпдеңгейлі модельді әзірлеу.....316

Оразбаев Б.Б., Жумадиллаева А.К., Курбангалиева Н.Б., Оразбаева К.Н.
Күкірт сапасын бағалаудың лингвистикалық модельдерін синтездеу
және күкіртті өндіру процесін бұлыңғыр модельдеу.....337

Сарсенбаева А.К., Рахимова Д.Р., Шормакова А.Н., Мансурова М.Е., Адали Э.
Семантикалық әдістерді заңнама саласында қолдану: агглютинативті
мәтіндерді талдауға арналған интеллектуалды жүйе.....354

Серек А., Шойынбек А., Шарипов К., Қуанышбай Д., Мухаметжанов А.
Сөйлеу транскрипцияларының лексикалық белгілеріне негізделген телефон
алаяқтықтарын талдау және жіктеу.....373

Шынжігіт Б.Б., Балабекова М.О., Амангелді Т.Т.
Кірпіш өнімдерін сату көлемдерін машиналық оқытуда талдау
және болжамдау.....393

Тохаева А.О., Альжанов А.К., Nezih Ö., Зиятбекова Г.З., Бегалиева К.Б.
PROXMOX VE негізінде жоғары оқу орындарында білім алушыларды
виртуалдандыру құзыреттерін қалыптастыру.....412

Төкенова Л.М., Әуелбеков О.А., Сапақова С., Саметова А.А., Бостанов Е.Л.
Пилотсыз ұшу аппараттарына арналған гибридті электр станцияларының жұмыс режимдерін модельдеу және оңтайландыру.....430

Еримбетова А.С., Бержанова У.Г., Дайырбаева Э.Н., Сәкенов Б.Е., Самбетбаева М.А.
Уақытша конволюциялық желі мен mediaріре көмегімен ым тілін тану.....443

Жукабаева Т.К., Бенхелифа Э., Марденов Е.М., Баумуратова Д., Карабаев Н.
Киберфизикалық өнеркәсіптік интернет заттары жүйелеріндегі шабуылдарға әрекет ету кезінде шешім қабылдауды қолдау.....461

СОДЕРЖАНИЕ

ИНФОРМАТИКА

Ахметова С.Т., Юнусова А.А., Алишева С.С., Олжатаева Б.Т., Мүсірепова Э.Б.
Интеллектуальный анализ данных для автоматического выявления языка ненависти в социальных сетях.....13

Аманов А.Н., Казбекова Г.Н., Жунисов Н.М., Абибуллаева А.А., Абен А.Б.
Обнаружение вторжений на основе искусственного интеллекта для DDoS-атак в программно-определяемых сетях.....30

Аманжолова С.Т., Усатова О.А., Мутанов Г.М., Муханов С.Б., Айтмукаш Д.
Бэкенд-архитектура гибридной системы проверки академических достижений на основе блокчейна.....52

Амирханова Г.А., Нургазы Т.Н., Амирханов Б.С., Нургазы Н.Н.
Разработка предиктивного цифрового двойника пищевого продукта на основе Edge ML и IoT-сенсоров.....73

Бекарыстанқызы А., Үсен Д., Қасенхан А., Чинибаев Е.
Холодный старт в системах рекомендаций в области образования: классические подходы и стратегии эпохи LLM.....91

Бимолдина Ж.А., Мусиралиева Ш.Ж., Багитова К.Б., Терейковская Л.
Использование машинного обучения и семантических моделей для обнаружения киберпропагандистского контента.....106

Чежимбаева К.С.
Прогнозирование ключевых KPI сетей 5G на основе нейросетевых моделей MLP и LSTM.....129

Даутбаева А.О., Конырбаев Н.Б., Абильдаева Ж.Т., Есиркепова А.У., Карим Н.А.
Разработка приложения для оптимизации процесса трудоустройства выпускников.....148

Джусупбекова Г., Othman M., Ордабаева Г.
Сравнительный анализ алгоритмов искусственного интеллекта для обнаружения сетевых атак.....167

Исахов А.А., Оразмолдаев Н., Жаркынбек Е., Абылкасымова А.
Численное моделирование распространения вирусной инфекции воздушно-капельным путём в замкнутых помещениях.....182

Кантуреева М.А., Омарова Г.С., Дүйсен Ж.Д., Шекербек А.Ә., Тулебаев Е.Б.
Использование методов машинного обучения для прогнозирования
и оптимизации процессов эвакуации людей в высотных зданиях.....202

Хусаин Б., Тельманов М.М., Хусаин А.Б., Бродский А.Р., Сасс А.С.
Цифровой двойник комплексной системы очистки и декарбонизации
выбросов тепловых установок.....218

Кулакаева А.Е., Ашуров А.Е., Жумажанов Б.Р., Дайнеко Е.А., Зылғара А.Е.
Алгоритм определения начальных орбитальных параметров KazEOSat-1
для деорбитации.....236

**Мименбаева А.Б., Туребаева А.Д., Оспанова Т.Т., Аруова А.Б.,
Найзағараева А.А.**
Разработка и сравнительный анализ моделей машинного обучения
для прогнозирования городского трафика.....253

Науменко В.В., Муканова Ж.А., Киселёва О.В., Майнцер Д.А., Нерезов А.К.
Применение опросов в режиме реального времени для повышения
успеваемости обучающихся.....271

Назырова А.Е., Кадеркеева З.К., Бекманова Г.Т., Милош М., Ламашева Ж.Б.
Цифровое образование и академическая успеваемость учащихся:
межуровневый анализ.....287

Оралбекова Д., Мамырбаев О., Ахмедиярова А., Касымова Д., Алибиева Ж.
Разработка многоуровневой модели для абстрактивного резюмирования текста
на основе предварительно обученных моделей.....316

**Оразбаев Б.Б., Жумадиллаева А.К., Курбангалиева Н.Б., Есиркесинов Р.Ж.,
Оразбаева К.Н.**
Синтез лингвистических моделей оценки качества серы и нечёткое
моделирование процесса её производства.....337

Сарсенбаева А.К., Рахимова Д.Р., Шормакова А.Н., Мансурова М.Е., Адали Э.
Применение семантических методов в юридическом анализе: интеллектуальная
система для обработки агглютинативных текстов.....354

Серек А., Шойынбек А., Шарипов К., Куанышбай Д., Мухаметжанов А.
Анализ и классификация телефонного мошенничества на основе
лексических признаков речевых транскрипций.....373

Шынжігіт Б.Б., Балабекова М.О., Амангелді Т.Т.
Анализ и прогнозирование объёмов продаж кирпичной продукции
с использованием машинного обучения.....393

Тохаева А.О., Альжанов А.К., Nezih Ö., Зиятбекова Г.З., Бегалиева К.Б.
Формирование компетенций в области виртуализации у обучающихся
в высшем образовании на основе платформы Proxmox VE.....412

Тукенова Л.М., Ауелбеков О.А., Сапакова С.З., Саметова А.А., Бостанов Е.Л.
Моделирование и оптимизация режимов работы гибридных силовых установок
для беспилотных летательных аппаратов.....430

Еримбетова А.С., Бержанова У.Г., Дайырбаева Э.Н., Сакенов Б.Е., Самбетбаева М.А.
Распознавание языка жестов с использованием временных свёрточных
сетей и MediaPipe4.....43

Жукабаева Т.К., Бенхелифа Э., Марденов Е.М., Баумуратова Д., Карабаев Н.
Поддержка принятия решений при реагировании на атаки в киберфизических
промышленных системах интернета вещей.....461

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 1.
Namber 357 (2026). 373–392

https://doi.org/10.32014/2026.2518-1726.418

IRSTI 28.23.25
UDC 004.855.5

© Serek A.[1], Shoiynbek A.[2]*, Sharipov K.[2], Kuanyshbay D.[2], Mukhametzhanov A.[3], 2026.
[1]Astana IT University, Astana, Kazakhstan;
[2]Narxoz University, Almaty, Kazakhstan;
[3]SDU University, Kaskelen, Kazakhstan.
E-mail: aisultan.shoiynbek@gmail.com

ANALYSIS AND CLASSIFICATION OF TELEPHONE FRAUD BASED ON LEXICAL FEATURES OF SPEECH TRANSCRIPTIONS

Serek Azamat — PhD, Associate Professor, Astana IT University, Astana, Kazakhstan,
E-mail: azamat.serek@astanait.edu.kz, https://orcid.org/0000-0001-7096-6765;
Shoiynbek Aisultan — PhD, Professor, Narxoz University, Almaty, Kazakhstan,
E-mail: aisultan.shoiynbek@gmail.com, https://orcid.org/0000-0002-9328-8300;
Sharipov Karim — Master's student, Narxoz University, Almaty, Kazakhstan,
E-mail: karim.sharipov@narxoz.kz, https://orcid.org/0009-0003-2452-8803;
Kuanyshbay Darkhan — PhD, Assistant-Professor, Narxoz University, Almaty, Kazakhstan,
E-mail: darkhan.kuanyshbay @narxoz.kz, https://orcid.org/0000-0001-5952-8609;
Mukhametzhanov Assylbek — Master's student, SDU University, Kaskelen, Kazakhstan,
E-mail: 221107046@stu.sdu.edu.kz, https://orcid.org/0009-0009-8528-9985.

Abstract. Telephone fraud (vishing) is one of the most common forms of social engineering, causing significant financial and psychological damage. In conditions of constant number changes and fraud scenarios, traditional protection methods are not effective enough, which makes it necessary to automatically analyze the contents of telephone conversations. This paper presents an experimental study of the use of machine learning methods for the automatic detection of fraudulent phone calls based on text transcriptions of speech. To conduct the experiment, a balanced Russian-language corpus of 1,400 telephone conversations was formed, including fraudulent and legitimate calls received from open sources. The audio recordings were automatically transcribed using the Whisper neural network speech recognition model, after which the texts were normalized and lemmatized. TF–IDF unigrams and bigrams were used as the feature representation. Based on the data obtained, several classical machine learning models were trained and compared, including Logistic Regression, Linear SVM, Multinomial Naive Bayes, Random Forest, and XGBoost. Experimental results showed that all the considered

models achieve high classification accuracy, while the best performance was demonstrated by linear models and the Multinomial Naive Bayes classifier with minimal smoothing (accuracy up to 94%, ROC–AUC up to 0.99). The analysis of lexical features made it possible to identify stable verbal markers of fraudulent speech, characteristic of typical scenarios of social engineering. The stability and generalizing ability of the models were confirmed using k-fold cross-validation and ROC–AUC analysis. The results obtained indicate the practical applicability of the proposed approach for the automatic detection of telephone fraud in real conditions.

Keywords: telephone fraud; vishing; machine learning; natural language processing; speech recognition; TF–IDF; text classification

Financing. *This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP27510301 "Development of technology for recognizing fraudulent actions during a telephone conversation and/or text message exchange in messengers based on artificial intelligence algorithms").*

For citations: Serek A., Shoiynbek A., Sharipov K., Kuanyshbay D., Mukhametzhanov A. Analysis and classification of telephone fraud based on lexical features of speech transcriptions. Academic Scientific Journal of Computer Science, 2026. — No.1. – P. 373–392. DOI: https://doi.org/10.32014/2026.2518-1726.418

© Серек А.[1], Шойынбек А.[2]*, Шарипов К.[2], Қуанышбай Д.[2], Мухаметжанов А.[3], 2026.

[1]Астана IT Университеті, Астана, Қазақстан;
[2] Нархоз университеті, Алматы, Қазақстан;
[3] SDU Университеті, Қаскелең, Қазақстан.
E-mail: aisultan.shoiynbek@gmail.com

СӨЙЛЕУ ТРАНСКРИПЦИЯЛАРЫНЫҢ ЛЕКСИКАЛЫҚ БЕЛГІЛЕРІНЕ НЕГІЗДЕЛГЕН ТЕЛЕФОН АЛАЯҚТЫҚТАРЫН ТАЛДАУ ЖӘНЕ ЖІКТЕУ

Серек Азамат — PhD, Қауымдастырылған профессор, Астана IT Университеті, Астана, Қазақстан,
E-mail: azamat.serek@astanait.edu.kz, https://orcid.org/0000-0001-7096-6765;
Шойынбек Айсұлтан — PhD, профессор, Нархоз университеті, Алматы, Қазақстан,
E-mail: aisultan.shoiynbek@gmail.com, https://orcid.org/0000-0002-9328-8300;
Шарипов Карим — магистрант, Нархоз университеті, Алматы, Қазақстан,
E-mail: karim.sharipov@narxoz.kz, https://orcid.org/0009-0003-2452-8803;
Қуанышбай Дархан — PhD, ассистент-профессор, Нархоз университеті, Алматы, Қазақстан,
E-mail: darkhan.kuanyshbay@sdu.edu.kz, https://orcid.org/0000-0001-5952-8609;
Мухаметжанов Асылбек — магистрант, SDU Университеті, Қаскелең, Қазақстан,
E-mail: 221107046@stu.sdu.edu.kz, https://orcid.org/0009-0009-8528-9985.

Аннотация. Телефондық алаяқтық (вишинг) - бұл қаржылық және психологиялық тұрғыдан айтарлықтай зиян келтіретін әлеуметтік инженерияның кең таралған түрлерінің бірі. Алаяқтардың нөмірлері мен сценарийлерінің үнемі өзгеруі жағдайында дәстүрлі қорғаныс әдістері тиімсіз болып шығады, бұл телефон қоңырауларының мазмұнын автоматты түрде талдауды қажет етеді. Бұл жұмыста мәтіндік сөйлеу транскрипцияларына негізделген алаяқтық телефон қоңырауларын автоматты түрде анықтау үшін машиналық оқыту әдістерін қолдану бойынша эксперименттік зерттеу ұсынылған. Эксперимент жүргізу үшін ашық көздерден алынған алаяқтық және заңды қоңырауларды қамтитын 1400 телефон қоңырауларынан тұратын теңдестірілген орыс тілді корпус құрылды. Дыбыстық жазбалар Whisper-дің сөйлеуді танудың нейрондық моделі арқылы автоматты түрде транскрипцияланды, содан кейін мәтіндер қалыпқа келтіріліп, лемматизацияланды. TF–IDF униграммалары мен биграммалары қолтаңба көрінісі ретінде пайдаланылды. Алынған мәліметтер Logistic Regression, Linear SVM, Multinomial Naive Bayes, Random Forest және XGBoost сияқты бірнеше классикалық Машиналық оқыту модельдерін оқыды және салыстырды. Эксперименттік нәтижелер барлық қарастырылған модельдер жоғары жіктеу дәлдігіне қол жеткізетінін көрсетті, ең жақсы көрсеткіштер сызықтық модельдер мен Multinomial Naive Bayes классификаторын минималды Тегістеу арқылы көрсетті (accuracy 94% дейін, ROC–AUC 0.99 дейін). Лексикалық белгілерді талдау әлеуметтік инженерияның типтік сценарийлеріне тән алаяқтық сөйлеудің тұрақты ауызша белгілерін анықтауға мүмкіндік берді. Модельдердің тұрақтылығы мен жалпылау қабілеті ROC-AUC кросс-валидациясы мен талдауының k–еселігі арқылы расталды. Нәтижелер нақты жағдайларда телефон алаяқтықтарын автоматты түрде анықтау үшін ұсынылған тәсілдің практикалық қолданылуын көрсетеді.

Түйін сөздер: телефон алаяқтық; вишинг; Машиналық оқыту; табиғи тілді өңдеу; сөйлеуді тану; TF-IDF; мәтіндерді жіктеу

© Серек А.[1], Шойынбек А.[2]*, Шарипов К.[2], Куанышбай Д.[2], Мухаметжанов А.[3], 2026.

[1] Астана IT Университет, Астана, Казахстан;

[2] Университет Нархоз, Алматы, Казахстан;

[3] SDU Университет, Каскелен, Казахстан.

E-mail: aisultan.shoiynbek@gmail.com

АНАЛИЗ И КЛАССИФИКАЦИЯ ТЕЛЕФОННОГО МОШЕННИЧЕСТВА НА ОСНОВЕ ЛЕКСИЧЕСКИХ ПРИЗНАКОВ РЕЧЕВЫХ ТРАНСКРИПЦИЙ

Серек Азамат — PhD, ассоциированный профессор, Астана IT Университет, Астана, Казахстан,

E-mail: azamat.serek@astanait.edu.kz, https://orcid.org/0000-0001-7096-6765;

Шойынбек Айсултан — PhD, профессор, Университет Нархоз, Алматы, Казахстан,
E-mail: aisultan.shoiynbek@gmail.com, https://orcid.org/0000-0002-9328-8300;
Шарипов Карим — магистрант, Университет Нархоз, Алматы, Казахстан,
E-mail: karim.sharipov@narxoz.kz, https://orcid.org/0009-0003-2452-8803;
Куанышбай Дархан — PhD, ассистент-профессор, Университет Нархоз, Алматы, Казахстан,
E-mail: darkhan.kuanyshbay@sdu.edu.kz, https://orcid.org/0000-0001-5952-8609;
Мухаметжанов Асылбек — магистрант, SDU Университет, Каскелен, Казахстан,
E-mail: 221107046@stu.sdu.edu.kz, https://orcid.org/0009-0009-8528-9985.

Аннотация: Телефонное мошенничество (вишинг) представляет собой одну из наиболее распространённых форм социального инжиниринга, наносящую значительный финансовый и психологический ущерб. В условиях постоянной смены номеров и сценариев мошенников традиционные методы защиты оказываются недостаточно эффективными, что обусловливает необходимость автоматизированного анализа содержимого телефонных разговоров. В данной работе представлено экспериментальное исследование применения методов машинного обучения для автоматического выявления мошеннических телефонных звонков на основе текстовых транскрипций речи. Для проведения исследования сформирован сбалансированный русскоязычный корпус, включающий 1400 телефонных разговоров (мошеннические и легитимные звонки), собранных из открытых источников. Аудиозаписи были автоматически транскрибированы с использованием нейросетевой модели распознавания речи Whisper, после чего тексты подвергались нормализации и лемматизации. В качестве признакового представления использованы TF–IDF-признаки (униграммы и биграммы). На основе подготовленных данных были обучены и сравнены несколько моделей машинного обучения, включая Logistic Regression, Linear SVM, Multinomial Naive Bayes, Random Forest и XGBoost. Результаты экспериментов показали, что все модели демонстрируют высокую точность классификации, при этом наилучшие показатели достигнуты линейными моделями и классификатором Multinomial Naive Bayes с минимальным сглаживанием (accuracy до 94%, ROC–AUC до 0,99). Анализ лексических признаков позволил выявить устойчивые языковые маркеры мошеннической речи, характерные для типичных сценариев социального инжиниринга. Надёжность и обобщающая способность моделей подтверждены с использованием k-кратной кросс-валидации и анализа ROC–AUC. Полученные результаты свидетельствуют о высокой практической применимости предложенного подхода для автоматического выявления телефонного мошенничества в реальных условиях.

Ключевые слова: телефонное мошенничество; вишинг; машинное обучение; обработка естественного языка; распознавание речи; TF–IDF; классификация текстов

Введение. Телефонное мошенничество («вишинг», от voice phishing) стало серьезной угрозой, ежегодно приносящей жертвам значительные финансовые потери (Jones et al., 2021). Злоумышленники звонят, выдавая себя за банковских

сотрудников, коммунальные службы или другие организации, чтобы обманом получить конфиденциальные данные или деньги (Jones et al., 2021). Традиционные меры защиты – например, чёрные списки номеров – недостаточно эффективны, ведь мошенники часто меняют номера (Triantafyllopoulos et al., 2025). Поэтому все больше внимания уделяется автоматическому распознаванию мошеннических звонков по содержанию разговора (Triantafyllopoulos et al., 2025).

Современные технологии искусственного интеллекта (ИИ) и обработки речи позволяют анализировать сам текст разговора и манеру общения в реальном времени (Triantafyllopoulos et al., 2025; Kim et al., 2021; Lee et al., 2023). Уже появляются решения, способные во время звонка распознать характерные признаки «скрипта» мошенников и предупредить потенциальную жертву о угрозе. Такие системы внедряются сотовыми операторами – например, «СберМобайл», «Тинькофф Мобайл» и др. – и показывают высокую точность выявления злоумышленников, по данным открытых материалов операторов связи (Sberbank, 2026; T-Bank, 2026). Автоматизация распознавания мошенничества с помощью ИИ имеет значительный потенциал в повышении эффективности защиты, так как позволяет определять именно мошеннические паттерны речи, а не просто блокировать заранее известные номера. В данной работе представлено экспериментальное исследование применения методов машинного обучения для автоматического обнаружения мошеннических телефонных разговоров по их расшифровкам. Нами был собран корпус звонков, выполнена их транскрипция с помощью нейросетевой модели распознавания речи и обучение нескольких моделей бинарной классификации по тексту разговора.

Вклад настоящей работы заключается в следующем. Во-первых, был сформирован и подготовлен специализированный русскоязычный корпус телефонных разговоров, включающий как мошеннические, так и легитимные звонки, с последующей автоматической транскрипцией и текстовой обработкой. Во-вторых, проведено систематическое экспериментальное сравнение классических методов машинного обучения (Logistic Regression, Linear SVM, Multinomial Naive Bayes, Random Forest, XGBoost) для задачи обнаружения телефонного мошенничества на основе текстовых транскрипций. В-третьих, выполнен детальный анализ лексических паттернов мошеннической речи, что позволило выявить устойчивые словесные маркеры, характерные для сценариев социального инжиниринга. В-четвёртых, эмпирически показано, что относительно простые и интерпретируемые модели (в частности, Multinomial Naive Bayes и линейные классификаторы) способны достигать качества, сопоставимого или превосходящего более сложные ансамблевые методы. Наконец, устойчивость и обобщающая способность моделей были подтверждены с использованием k-кратной кросс-валидации и анализа ROC–AUC, что повышает надёжность полученных выводов и их практическую применимость.

Литературный обзор

Проблема телефонного мошенничества и социальной инженерии активно исследуется в последние годы в связи с глобальным ростом киберпреступности. Исследователи выделяют несколько ключевых направлений в разработке систем противодействия вишингу: анализ сетевых характеристик звонка, биометрический анализ голоса и семантический анализ содержания разговора.

Традиционные методы защиты, такие как создание «черных списков» номеров, демонстрируют снижающуюся эффективность из-за массового использования злоумышленниками технологий подмены номера (Caller ID spoofing) и IP-телефонии. Как отмечают Triantafyllopoulos и соавторы (2025), динамическая природа мошеннических сценариев требует перехода от статической фильтрации к интеллектуальному анализу контента в реальном времени.

Важным этапом в автоматизации процесса обнаружения мошенничества стало развитие технологий распознавания речи (ASR). Использование современных нейросетевых моделей, таких как Whisper от OpenAI, позволяет получать высокоточные текстовые транскрипции даже в условиях зашумленного аудиоканала телефонной связи. Исследования Radford et al. (2023) подтверждают, что масштабное обучение на слабо контролируемых данных делает такие модели устойчивыми к различным акцентам и диалектам, что критически важно для анализа русскоязычного сегмента звонков.

В области классификации текстов для выявления мошенничества (fraud detection) исследователи часто сравнивают эффективность классических алгоритмов машинного обучения и глубоких нейронных сетей. Несмотря на развитие трансформеров, работы Kowsari et al. (2019) и Ferhati et al. (2025) показывают, что методы на основе TF–IDF векторизации в сочетании с линейными моделями (SVM, Logistic Regression) или наивным байесовским классификатором (Multinomial Naive Bayes) остаются крайне эффективными для задач с высокой размерностью признаков и ограниченным объемом данных.

Особое внимание в литературе уделяется лексическим маркерам и сценариям общения. Исследования Jones et al. (2021) подчеркивают, что мошеннические звонки характеризуются использованием специфических «скриптов», направленных на создание чувства срочности или авторитетности. Ключевые слова, связанные с банковской безопасностью, подтверждением транзакций и персональными кодами, служат устойчивыми предикторами класса «мошенничество».

Таким образом, современные системы обнаружения вишинга эволюционируют в сторону комплексных решений, объединяющих автоматическое распознавание речи и продвинутые методы обработки естественного языка (NLP) для выявления скрытых паттернов в поведении злоумышленников.

Материалы и методы

Создание корпуса.

Для эксперимента сформирован сбалансированный датасет из 1400 записей телефонных разговоров на русском языке. Из них 700 – мошеннические звонки, например распространенные сценарии звонков псевдосотрудников банка, коммунальных служб и т.д., а 700 – обычные телефонные разговоры (Рис. 1).

```
Распределение классов:
label
fraud        700
not_fraud    700
Name: count, dtype: int64
```

Рисунок 1 – Распределение классов

Аудиозаписи были найдены на открытых источниках YouTube – например, на каналах, посвященных разоблачению телефонных мошенников, а также записи бытовых телефонных разговоров. Каждая аудиозапись была снабжена меткой класса: «fraud» (мошенничество) или «not_fraud» (нет мошенничества). Метки присваивались автоматически исходя из источника и контекста (например, файлы в папке с мошенническими звонками помечены как fraud).

Транскрипция аудио.

Все 1400 звонков были автоматически распознаны (транскрибированы в текст) с помощью модели Whisper (OpenAI) – нейросетевого преобразователя речи в текст (Radford, A. et al., 2023). Использовалась модель Whisper Small, обученная для русского языка, что обеспечило достаточно высокое качество распознавания при приемлемом времени обработки. Для каждого звонка получена текстовая расшифровка. Пример фрагмента транскрипта мошеннического звонка: «… Система контроля звонков в онлайне не будет работать в личных беседах клиента…» (звонящий пытается говорить официально, ссылаясь на некую «систему контроля»). Все полученные пары «текст – метка» сохранены в таблицу для дальнейшего использования в обучении моделей.

Предобработка текста.

Для последующего анализа тексты звонков были приведены к стандартизированному виду. Выполнены следующие шаги:

1. приведение к нижнему регистру;
2. удаление посторонних символов и пунктуации, оставлены только слова (включая последовательности букв и цифр);
3. лемматизация – приведение слов к начальной форме. Лемматизация выполнялась с помощью библиотеки Pymorphy3: например, фраза «картой в карте карт» преобразуется в «карта в карта». Это уменьшает вариативность словоформ в русском языке (Kowsari et al., 2019).
4. удаление стоп-слов – наиболее частотных и малоинформативных слов.

Мы использовали кастомный список из ~15 слов, включающий междометия, приветствия и общие слова (например, «здравствуйте», «алло», «да», «нет», «вот», «это», «как», «уже» и т.п.), которые часто встречаются в речи и не помогают отличить мошенников. После очистки и нормализации текстов каждая транскрипция была представлена как набор значимых токенов (слов) в базовой форме.

Признаки и представление данных.

Для преобразования текстов в формат, пригодный для обучения моделей, применен метод «Bag of Words» с взвешиванием TF–IDF (Kowsari et al., 2019). С помощью TfidfVectorizer были извлечены униграммы и биграммы (ngram_range = (1,2)) из текстов, учитывая слова, встречающиеся как минимум в 5 разговорах, и игнорируя слова, присутствующие в большинстве звонков (Kowsari et al., 2019). Таким образом, очень редкие слова и чрезмерно частые не учитывались как признаки.

Отбор терминов осуществлялся на этапе построения словаря признаков: в него включались только те униграммы и биграммы, которые встречались не менее чем в пяти различных разговорах (параметр *min_df = 5*), что позволяло исключить случайные, шумовые токены, не обладающие статистической значимостью для задачи классификации. Одновременно из словаря исключались термины, присутствующие более чем в 85% разговоров (параметр *max_df = 0.85*), поскольку такие слова не несут дискриминативной информации и одинаково характерны как для мошеннических, так и для легитимных звонков (Kowsari et al., 2019).

Данный механизм фильтрации не искажает семантическое представление данных, поскольку модель ориентируется не на отдельные редкие или общеупотребительные слова, а на устойчивые лексические паттерны, характерные для мошеннических сценариев общения. Напротив, исключение нерелевантных признаков снижает размерность пространства признаков, уменьшает уровень шума и способствует повышению обобщающей способности модели, снижая риск переобучения (Kowsari et al., 2019; Ferhati et al., 2025).

После векторизации каждый звонок представлен разреженным вектором большой размерности, где каждому допустимому слову или биграмме соответствует признак – TF–IDF значение (важность этого термина для данного разговора относительно всего корпуса).

Разбиение на выборки.

Исходный набор из 1400 примеров разделен на обучающую и тестовую выборки в отношении 70/30 с сохранением баланса классов. Затем тестовая часть (30% данных) была поровну разделена на Valid и Test (по 15% каждый). Валидационная выборка использовалась для контроля обучения ансамблевых моделей и подбора гиперпараметров, тогда как финальная оценка проводилась на независимой тестовой выборке. В итоге: 980 звонков в обучающей выборке и по 210 звонков в валидационной и тестовой.

Модели машинного обучения.

Для решения задачи бинарной классификации телефонных звонков на мошеннические и легитимные были протестированы несколько классических алгоритмов машинного обучения. Выбор данных методов обусловлен их широкой распространённостью в задачах анализа текстов, устойчивостью к высокой размерности признакового пространства и возможностью интерпретации полученных результатов (Kowsari et al., 2019; Ferhati et al., 2025).

1. Логистическая регрессия

Логистическая регрессия (Logistic Regression) является линейным классификатором, который моделирует вероятность принадлежности объекта к определённому классу как логистическую функцию от взвешенной линейной комбинации входных признаков. Несмотря на относительную простоту, данный метод демонстрирует высокую эффективность при работе с разреженными высокоразмерными признаками, характерными для текстовых данных (Kowsari et al., 2019; Ferhati et al., 2025).

В рамках данного исследования логистическая регрессия применялась с L2-регуляризацией, позволяющей ограничить величину коэффициентов модели и снизить риск переобучения. Интенсивность регуляризации задавалась параметром C = 0.01, что соответствует усиленному штрафу на сложность модели. Дополнительно использовалась балансировка классов, направленная на повышение чувствительности модели к мошенническому классу.

2. Random Forest

Random Forest представляет собой ансамблевый метод, основанный на построении множества решающих деревьев, обучаемых на различных бутстрап-выборках исходных данных (Ferhati et al., 2025). Итоговое решение формируется путём голосования отдельных деревьев.

Поскольку отдельные решающие деревья обладают высокой дисперсией, ансамблирование позволяет существенно повысить устойчивость модели. В данном исследовании глубина деревьев была ограничена значением 5, а количество деревьев установлено равным 1000, что позволило сдерживать переобучение при сохранении достаточной выразительной способности модели.

3. Линейный SVM

Метод опорных векторов (Support Vector Machine, SVM) является мощным алгоритмом обучения с учителем, основная цель которого заключается в нахождении оптимальной разделяющей гиперплоскости между классами. Для задач классификации текстов особенно эффективен линейный SVM, хорошо работающий с разреженными и высокоразмерными признаковыми представлениями.

В рамках эксперимента использовалась линейная версия SVM с коэффициентом регуляризации C = 0.025, обеспечивающим баланс между максимизацией зазора и минимизацией ошибки классификации. Максимальное число итераций оптимизации было увеличено до 3000 для гарантированной сходимости алгоритма.

4. XGBoost

XGBoost является одной из наиболее эффективных реализаций градиентного бустинга на решающих деревьях (Ferhati K. et al., 2025). Метод последовательно строит ансамбль моделей, каждая из которых корректирует ошибки предыдущих, минимизируя регуляризованную функцию потерь.

Для обеспечения устойчивого обучения были выбраны следующие гиперпараметры: 500 деревьев глубиной 3, малый шаг обучения (learning_rate = 0.02), а также бутстрапирование 80% объектов и признаков. Дополнительно применялась L1- и L2-регуляризация ($\alpha = 0.5$, $\lambda = 1.5$). Обучение проводилось с контролем качества на валидационной выборке; при этом признаков существенного переобучения выявлено не было.

5. Классификатор Multinomial Naive Bayes

Классификатор Multinomial Naive Bayes основан на теореме Байеса и предположении условной независимости признаков. Несмотря на данное упрощение, байесовские методы широко применяются в задачах классификации текстов и часто демонстрируют высокую эффективность.

В ходе эксперимента особое внимание уделялось параметру Лапласовского сглаживания (α), регулирующему влияние редко встречающихся терминов. Были протестированы значения $\alpha = 1.0$, $\alpha = 10.0$ и $\alpha = 0.1$. Большие значения α обеспечивают более «гладкие» вероятностные оценки и снижают влияние редких слов, тогда как малые значения позволяют модели учитывать редкие, но потенциально информативные лексические маркеры мошеннической речи.

Обучение и оценка всех моделей проводились в среде Jupyter Notebook с использованием библиотеки scikit-learn, а для градиентного бустинга применялся официальный Python-пакет xgboost. Для каждой модели были получены предсказания на тестовой выборке, рассчитаны метрики качества, построены матрицы ошибок и ROC-кривые.

Результаты

1. Общее качество классификации.

Полученные модели продемонстрировали высокую точность в распознавании мошеннических звонков по тексту. Практически все алгоритмы достигли accuracy выше 0.8, а лучшие – выше 0.9 (Lee et al., 2023; Chichwadia et al., 2024). Наилучший результат показал Naive Bayes с минимальным сглаживанием (α=0.1): 94% точности на тестовой выборке. При этом баланс метрик по классам также выдающийся: precision и recall около 0.90–0.98 для обоих классов, то есть модель почти одинаково хорошо выявляет мошеннические звонки и распознаёт легитимные разговоры. Это подтверждается матрицей ошибок для модели Naive Bayes ($\alpha = 0.1$), представленной на рисунке 2, где число ложных срабатываний и пропусков минимально.

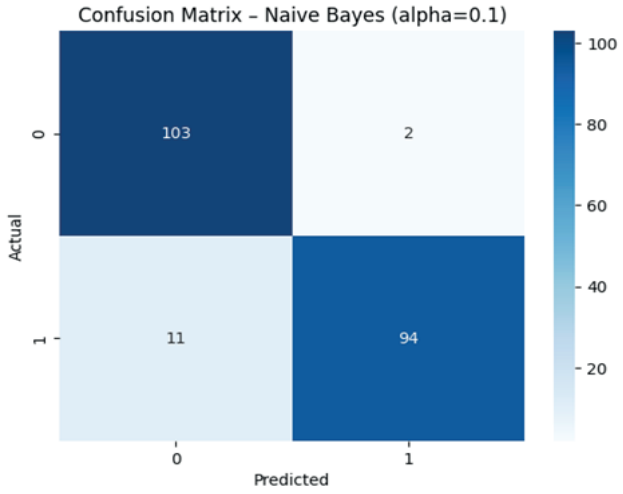

Рисунок 2 – Confusion Matrix - Naive Bayes (alpha = 0.1)

Linear SVM и XGBoost уступили совсем немного – они показали ~92% точности. Логистическая регрессия правильно классифицировала около 90% звонков. Несколько хуже выступил Random Forest (более сложная нелинейная модель) – порядка 84% правильных классификаций. Для наглядности сводка результатов представлена в таблице 1. Помимо accuracy в таблице представлены значения F1-score для класса *fraud*, что позволяет более объективно оценить баланс между точностью и полнотой моделей при обнаружении мошеннических звонков.

Таблица 1 – Метрики качества моделей машинного обучения при классификации телефонных звонков

Модель	Accuracy (точность)	Precision (fraud)	Recall (fraud)	F1-score (fraud)	Precision (not_fraud)	Recall (not_fraud)
Naive Bayes (α=0.1)	0.94	0.90	0.98	0.94	0.98	0.90
Linear SVM	0.92	0.94	0.90	0.92	0.90	0.94
XGBoost	0.92	0.92	0.92	0.92	0.92	0.92
Logistic Regression	0.90	0.92	0.88	0.90	0.88	0.92
Random Forest (деревья=1000)	0.84	0.85	0.83	0.84	0.83	0.86
Naive Bayes (α=1.0)	0.87	0.80	0.98	0.88	0.98	0.76
Naive Bayes (α=10.0)	0.76	0.67	1.00	0.80	1.00	0.51

Как видно, все модели, кроме сильно сглаженного Naive Bayes (α=10), обеспечили высокий recall для мошеннических звонков – 0.88 и выше, что важно с практической точки зрения (не пропускать мошенников)

(Triantafyllopoulos et al., 2025; Chichwadia et al., 2024). Однако при сильном сглаживании (α=10) Naive Bayes фактически склонен всегда относить звонок к классу «мошенничество», отсюда recall(fraud)=1.0 ценой очень низкого recall для нормальных звонков (лишь ~51%). Напротив, без сглаживания (α=0.1) Naive Bayes добивается баланса ~0.9–0.98 на обоих классах, показывая, что учет даже редких слов сыграл роль в идентификации.

Другие алгоритмы также достаточно сбалансированы: например, SVM имеет precision и recall ~0.90–0.94 для обеих категорий, что значит минимум ошибок как первого, так и второго рода. Logistic Regression дает f1-score = 0.90 для каждого класса– близко к SVM, подтверждая, что линейные модели хорошо справляются с разделением данных. Хуже баланс у Random Forest: его precision/recall ~0.84–0.86, то есть он чаще ошибается (в сравнении с линейными методами). Вероятно, ограничение глубины деревьев (до 5) не позволило Random Forest уловить некоторых важных комбинаций признаков, а большее углубление могло вести к переобучению на небольшой выборке.

2. Анализ ключевых признаков.

Модели машинного обучения позволили также выяснить, какие слова или фразы наиболее характерны для мошеннических звонков по сравнению с обычными. Логистическая регрессия и SVM (линейные модели) присвоили наименьшие отрицательные веса следующим словам: «**код**», «**номер**», «**посылка**», «**ключ**», «**письмо**» и т.п. – то есть лексике, связанной с банковскими операциями, безопасностью счетов, заявками на услуги и др. Эти термины часто фигурируют в речевых скриптах мошенников (Jones et al., 2021; Lee et al., 2023). Например, требование сообщить код из СМС, разговоры о блокировке номера телефона, заявке на отключение услуг и т.д. Данный вывод наглядно иллюстрируется графиками лексических признаков класса fraud, представленными на рисунках 3 и 4, где показаны слова с наибольшими по модулю весами в линейных моделях.

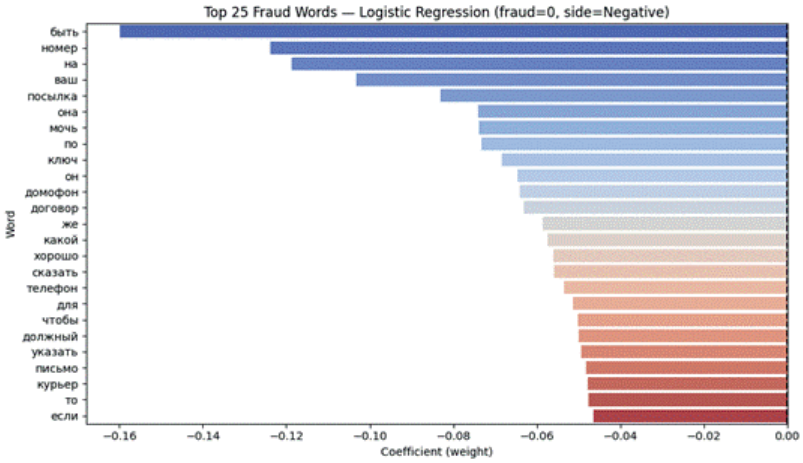

Рисунок 3 – Лексические признаки класса fraud (Logistic Regression)

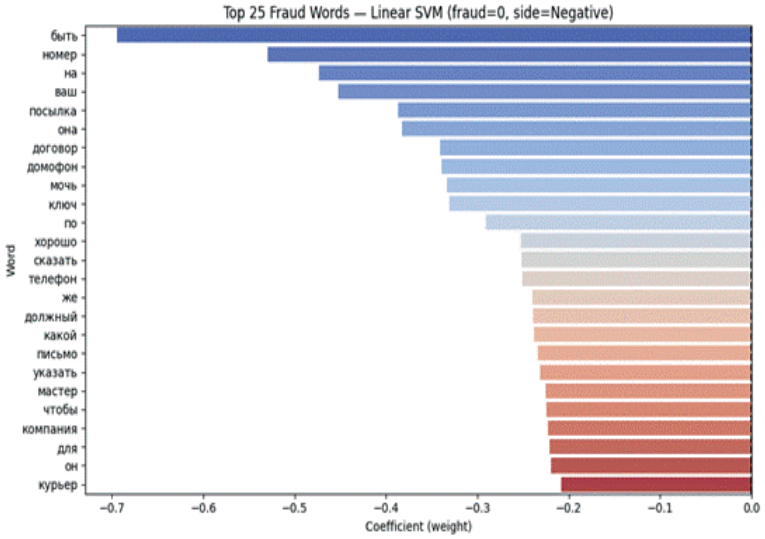

Рисунок 4 – Лексические признаки класса fraud (Linear SVM)

Напротив, слова, получившие положительные веса (ассоциированные с классом обычных звонков) – это, как правило, обращения по имени, разговорная лексика, неформальные выражения. Это наглядно показано на рисунках 5 и 6, где приведены топ-25 лексических признаков класса not_fraud, выявленных линейным SVM и логистической регрессией соответственно. Интересно, что встречаемость имени-отчества собеседника оказалась существенным индикатором: мошенники часто обращаются формально («Галина Петровна», «Роман Шарафудинович»), тогда как в бытовых разговорах такой стиль редок.

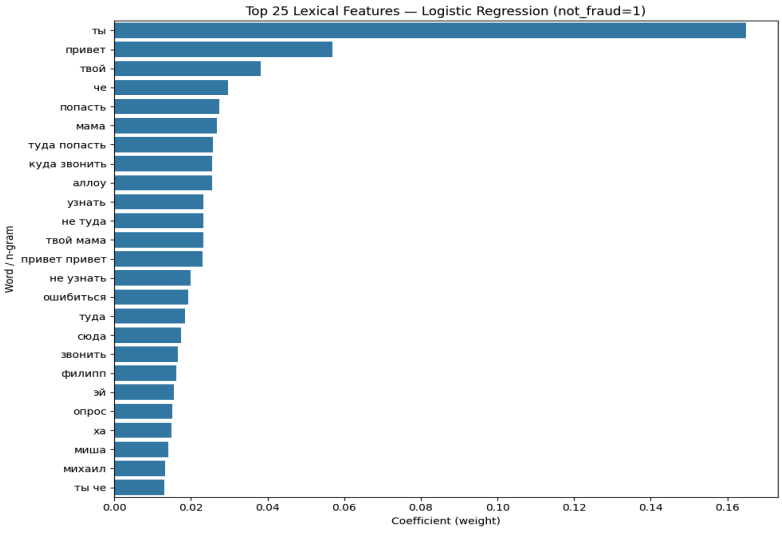

Рисунок 5 - Лексические признаки класса not_fraud (Logistic Regression)

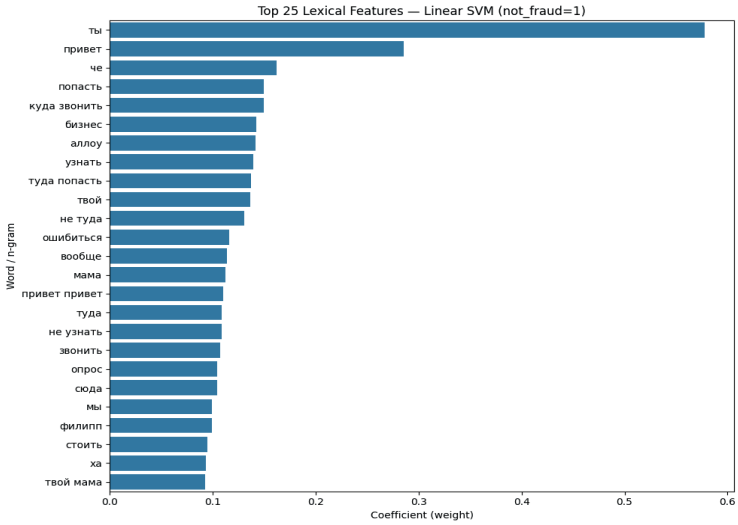

Рисунок 6 – Лексические признаки класса not_fraud (Linear SVM)

Классификатор Naive Bayes подтвердил эти наблюдения: разница лог-вероятностей показала, что слова вроде «ключ», «посылка», «курьер», «письмо» и т.п. намного более вероятны в мошеннических звонках, тогда как слова «привет», «дома», «алло» и т.п чаще встречаются в нормальных разговорах. Это наглядно иллюстрируется на рисунке 7, где представлены наиболее информативные лексические признаки, отсортированные по разнице лог-вероятностей между классами для модели Naive Bayes. Таким образом, модели научились узнавать мошенничество «по словам» – специфическому набору терминов и формулировок.

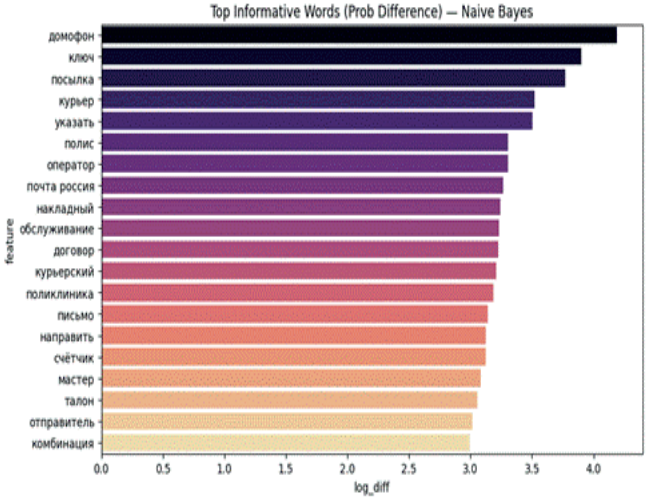

Рисунок 7 - Лексические признаки класса fraud (Naive Bayes)

3. ROC–AUC и устойчивость моделей.

Для оценки устойчивости и обобщающей способности моделей дополнительно была проведена *k-кратная кросс-валидация* на обучающей части датасета. Использовалась стратифицированная схема разбиения, обеспечивающая сохранение баланса классов «fraud / not_fraud» в каждом фолде. Значение параметра *k* было установлено равным 5, что является компромиссом между вычислительной сложностью и надёжностью оценки.

В рамках кросс-валидации для каждой модели вычислялись средние значения accuracy и стандартное отклонение по всем фолдам. В качестве признакового представления использовались те же TF–IDF униграммы и биграммы, а гиперпараметры моделей соответствовали настройкам, описанным в разделе «Методы». Это позволило оценить именно устойчивость моделей, а не эффект подбора параметров под конкретное разбиение данных.

Результаты кросс-валидации показали, что все рассмотренные модели демонстрируют стабильное качество классификации с относительно малым разбросом метрик между фолдами. Наиболее устойчивыми оказались линейные модели и классификатор Multinomial Naive Bayes с минимальным сглаживанием (α=0.1). Так, Linear SVM и Logistic Regression продемонстрировали среднюю accuracy на уровне ~0.90–0.92 при стандартном отклонении менее 0.02, что указывает на хорошую обобщающую способность и низкую чувствительность к выбору обучающей подвыборки.

Классификатор Multinomial Naive Bayes (α = 0.1), показавший наилучший результат на тестовой выборке, также подтвердил свою устойчивость в условиях кросс-валидации: средняя accuracy составила более 0.93 при минимальном разбросе значений. Это свидетельствует о том, что высокая точность данной модели не является следствием удачного разбиения данных, а отражает наличие устойчивых лексических паттернов мошеннической речи в корпусе.

Ансамблевые методы (Random Forest и XGBoost) продемонстрировали несколько больший разброс результатов между фолдами. Особенно это характерно для Random Forest, где вариативность accuracy была выше по сравнению с линейными моделями. Вероятной причиной является чувствительность деревьев решений к конкретному составу обучающих данных при высокой разреженности признакового пространства.

В целом результаты кросс-валидации подтверждают выводы, полученные на отложенной тестовой выборке: модели не переобучаются, демонстрируют устойчивое качество и способны обобщать выявленные закономерности на ранее невидимые данные. Это также подтверждается ROC-кривыми, представленными на рисунке 8, которые демонстрируют высокую площадь под кривой для всех рассмотренных моделей и близость кривых к левому верхнему углу, что указывает на хорошую разделимость классов. Это особенно важно с практической точки зрения, поскольку в реальных условиях

распределение телефонных разговоров может отличаться от обучающего корпуса.

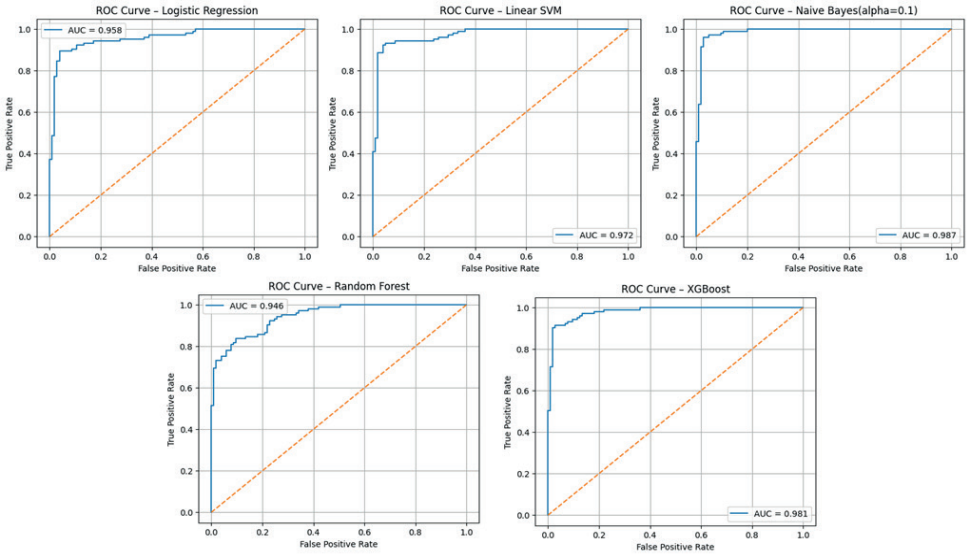

Рисунок 8 - ROC-кривые моделей машинного обучения для классификации телефонных звонков

Для количественной оценки качества помимо Accuracy рассмотрены площади под ROC-кривой. Все модели продемонстрировали очень высокие значения AUC: от ~0.95 у Random Forest до ~0.98–0.99 у лучших моделей. Так, ROC–AUC для SVM составил ~0.972, для XGBoost ~0.981, а наивысший результат дал Naive Bayes (α=0.1) – 0.987. Это указывает на отличную разделимость классов: даже при изменении порога классификации алгоритмы способны достичь высокой чувствительности при приемлемом количестве ложных тревог. В нашем сбалансированном наборе выбор порога 0.5 оказался близким к оптимальному (что отражено в precision/recall), но на практике при сильном дисбалансе (мошенников намного меньше, чем обычных звонков) порог можно было бы настроить для снижения false positives.

Высокий AUC Naive Bayes даже при среднем Accuracy=0.87 (при α=1) намекает, что эта модель ошибалась систематически из-за несбалансированности или выбора порога, хотя ранжировала разговоры довольно хорошо. Правильная калибровка вероятностей или подбор α позволили существенно повысить его качество до уровня лучших моделей.

Обсуждение

Эксперимент показал, что автоматическое распознавание телефонного мошенничества по речевым данным является практически осуществимой задачей и может достигать высоких значений F1-score (0,90–0,94). Использование данной метрики является принципиально важным, поскольку

в задачах обнаружения мошенничества классы, как правило, распределены неравномерно, и метрика accuracy может давать завышенную оценку качества модели за счёт доминирующего класса (Triantafyllopoulos et al., 2025).

Особенно показателен высокий F1-score, достигнутый наивным байесовским классификатором (до 0,94 при оптимальной настройке), несмотря на простоту модели. Это указывает на то, что наличие отдельных ключевых слов и устойчивых лексических маркеров играет решающую роль в определении мошеннического характера звонка (Jones et al., 2021; Lee et al., 2023; Chichwadia et al., 2024). Байесовский подход эффективно использует такие слова-«триггеры» (например, упоминания банковских терминов или призывов к срочным действиям), предполагая условную независимость признаков и практически не учитывая сложные взаимосвязи между ними.

Линейные модели, включая SVM и логистическую регрессию, также продемонстрировали высокие значения F1-score, что подтверждает их способность назначать информативные веса отдельным терминам и тем самым выявлять характерный «лексический почерк» мошеннических сценариев.

В то же время более сложные нелинейные методы, такие как XGBoost, не обеспечили значимого прироста F1-score (Moussavou Boussougou et al., 2023). Это позволяет предположить, что в рассматриваемой задаче дискриминативная информация сосредоточена преимущественно в отдельных словах и простых n-граммах, а не в сложных нелинейных комбинациях признаков. Аналогичный вывод подтверждается и относительно скромными результатами случайного леса: увеличение глубины деревьев не приводит к росту F1-score без существенного расширения обучающей выборки, тогда как неглубокие ансамбли уступают линейным моделям по способности точно балансировать между полнотой и точностью.

1. Практическая значимость.

Полученные результаты вселяют уверенность, что интеграция подобных моделей в системы телефонии сможет повысить защиту абонентов (Kim et al., 2022). Например, модель может в режиме реального времени анализировать транскрибируемый разговор и выводить предупреждение, если обнаружен типичный паттерн мошенничества (как уже реализовано в пилотных проектах операторов 2) (Tlahun et al., 2025). При точности ~90% ложные срабатывания будут случаться, но их доля может быть приемлемой с учетом ценности предотвращенного мошенничества. Важным аспектом здесь является баланс: что хуже – пропустить мошенника или зря потревожить клиента предупреждением? Наши модели можно настроить под нужный компромисс. Например, при чуть пониженном пороге вероятности мошенничества recall возрастет почти до 100%, но и вероятность возникновения ложных тревог станет больше. Напротив, повышение порога обеспечит практически полное отсутствие ложных тревог ценой некоторого снижения чувствительности. В реальной эксплуатации оптимум зависит от терпимости пользователей к предупреждениям и от вреда от пропущенного звонка.

2. Ограничения исследования.

Следует отметить, что наше исследование проводилось на относительно небольшом датасете (Dai et al., 2025). Собранные нами 700 мошеннических разговоров могут не охватывать всего разнообразия сценариев. Возможно, многие из них – вариации популярных схем, таких как «звонок из банка», «служба безопасности карты», «лжесотрудник энергокомпании» и прочие, и модель могла обучиться распознавать именно эти сюжеты. Встреться ей иной тип мошенничества, качество может снизиться (Li et al., 2025; Figueiredo et al., 2024). Аналогично, нормальные разговоры в датасете в основном бытовые и дружеские беседы. Они явно контрастируют с официально вежливым стилем мошенников, что облегчает классификацию. Но в реальной жизни бывают легитимные звонки от банков, магазинов, сервисных служб – по лексике похожие на мошеннические. Наши модели пока не учились отличать настоящего банковского оператора от мошенника, если оба говорят о картах и счетах. Для решения этой более тонкой задачи понадобятся дополнительные признаки (например, акустические: интонация, уверенность голоса; или анализ диалога: мошенники чаще давят на срочность, не дают перезвонить и т.д.) (Kim et al., 2025; Kang et al., 2022; Elizalbe et al., 2021).

Кроме того, качество автоматической транскрипции влияет на результаты. Мы использовали модель Whisper Small, и, хотя она достаточно точна, ошибки распознавания речи всё же присутствуют. В наших экспериментах это не помешало алгоритмам – даже с опечатками и пропусками модель выявляла ключевые слова. Однако в условиях шума или плохой связи распознавание может сильно ухудшиться, что снизит надежность классификации. В будущем стоит протестировать более крупные модели ASR (Whisper Large) либо специальные модели, обученные на телефонных переговорах, чтобы повысить точность транскрипции (Sim et al., 2025).

Заключение

В работе показана возможность эффективного автоматического обнаружения телефонного мошенничества на основе текстовых транскрипций разговоров с использованием методов машинного обучения. На основе сбалансированного русскоязычного корпуса и TF–IDF представления текстов были обучены и сравнены несколько классических моделей, продемонстрировавших высокое качество классификации. Наилучшие результаты показал классификатор Multinomial Naive Bayes с минимальным сглаживанием (accuracy выше 0.9 и ROC–AUC до 0.99), при этом линейные модели также подтвердили свою устойчивость и эффективность.

Анализ лексических признаков выявил устойчивые словесные маркеры, характерные для мошеннических сценариев социального инжиниринга, что подтверждает интерпретируемость и практическую применимость предложенного подхода. Результаты k-кратной кросс-валидации и анализа ROC–AUC показали хорошую обобщающую способность моделей и отсутствие переобучения. В целом полученные результаты свидетельствуют

о возможности использования текстового анализа телефонных разговоров в практических системах выявления мошенничества и создают основу для дальнейшего развития мультимодальных решений.

References

Chichwadia A.E., & Mpekoa N. (2024) Detecting smishing and vishing attacks using machine learning. International Journal of Intelligent Computing Research, 15(1). — P. 1234–1241. https://doi.org/10.20533/ijicr.2042.4655.2024.0151 (in Eng.).

Dai H., Liu Z., Liao W., et al. (2025) AugGPT: Leveraging ChatGPT for text data augmentation. IEEE Transactions on Big Data, 11(5). — P. 907–918 (in Eng.).

Elizalde B., & Emmanouilidou D. (2021) Detection of robocall and spam calls using acoustic features of incoming voicemails. Proceedings of Meetings on Acoustics, 45, 060004 (in Eng.).

Ferhati K., Burlea-Schiopoiu A., & Nascu A.-G. (2025) A text-based project risk classification system using multi-model AI: Comparing SVM, logistic regression, random forests, naive Bayes, and XGBoost. Systems, 13(12). — 1078 p. https://doi.org/10.3390/systems13121078 (in Eng.).

Figueiredo J., Carvalho A., Castro D., Gonçalves D., & Santos N. (2024) On the feasibility of fully AI-automated vishing attacks. arXiv preprint. https://arxiv.org/abs/2409.13793 (in Eng.).

Jones K.S., Armstrong M.E., Tornblad M.K., & Siami Namin A. (2021) How social engineers use persuasion principles during vishing attacks. Information & Computer Security, 29(2). — P. 314–331 (in Eng.).

Kang Y., Kim W., Lim S., Kim H., & Seo H. (2022) DeepDetection: Privacy-enhanced deep voice detection and user authentication for preventing voice phishing. Applied Sciences, 12(22), 11109. https://doi.org/10.3390/app122211109 (in Eng.).

Kim J., Gu S., Kim Y., Lee S., & Kang C. (2025) A multimodal voice phishing detection system integrating text and audio analysis. Applied Sciences, 15(20). — 11170 p. https://doi.org/10.3390/app152011170 (in Eng.).

Kim J., Kim J., Wi S., Kim Y., & Son S. (2022) HearMeOut: Detecting voice phishing activities in Android. In Proceedings of the 20th International Conference on Mobile Systems, Applications and Services (MobiSys). — P. 422–435 (in Eng.).

Kim J.-W., Hong G.-W., & Chang H. (2021) Voice recognition and document classification-based data analysis for voice phishing detection. Human-Centric Computing and Information Sciences, 11, Article 45. https://doi.org/10.1186/s13673-021-00245-9 (in Eng.).

Kowsari K., Jafari Meimandi K., Heidarysafa M., Mendu S., Barnes L., & Brown D. (2019) Text classification algorithms: A survey. Information, 10(4). — 150 p. https://doi.org/10.3390/info10040150 (in Eng.).

Lee M., & Park E. (2023) Real-time Korean voice phishing detection based on machine learning approaches. Journal of Ambient Intelligence and Humanized Computing, 14. — P. 1–13. https://doi.org/10.1007/s12652-021-03587-x (in Eng.).

Li W., Manickam S., Chong Y.W., & Karuppayah S. (2025) Talking like a phisher: LLM-based attacks on voice phishing classifiers. arXiv preprint. https://arxiv.org/abs/2507.16291(in Eng.).

Moussavou Boussougou M.K., & Park D.J. (2023) Attention-based 1D CNN–BiLSTM hybrid model enhanced with FastText word embedding for Korean voice phishing detection. Mathematics, 11(14), 3217. https://doi.org/10.3390/math11143217 (in Eng.).

Radford A., Kim J.W., Xu T., Brockman G., McLeavey C., & Sutskever I. (2023) Robust speech recognition via large-scale weak supervision. In Proceedings of the International Conference on Machine Learning. — P. 28492–28518. PMLR (in Eng.).

Sberbank (2026) Telefonnoye moshennichestvo: masshtaby problemy i mery protivodeystviya [Telephone fraud: Scale of the problem and countermeasures]. Official website of PJSC Sberbank. https://www.sberbank.ru/ru/sberpress/all/article?blockID=1303&lang=ru&newsID=8066ed65-189c-4b5e-a2b6-15513e6b62e8 (in Russian)

Sim J.-Y., & Kim S.-H. (2025) Detecting voice phishing with precision: Fine-tuning small language models. arXiv preprint. https://arxiv.org/abs/2506.06180 (in Eng.).

T-Bank (2026) Kak zashchitit' sebya ot moshennikov i sokhranit' den'gi [How to protect yourself from fraudsters and save money]. Official blog of T-Bank. https://www.tbank.ru/finance/blog/save-money/#q3 (in Russian)

Tlahun A.Z., Sumbiri D., & Jonathan K.N. (2025) Identifying and evaluating the best ML predictive models for detecting voice (phone-call) vishing attacks on MoMo users in real time. Journal of Information and Technology, 5(6). — P. 34–43. https://doi.org/10.70619/vol5iss6pp34-43 (in Eng.).

Triantafyllopoulos A., Spiesberger A.A., Tsangko I., Jing X., Distler V., Dietz F., Alt F., & Schulle B.W. (2025) Vishing: Detecting social engineering in spoken communication—A first survey and urgent roadmap. Computer Speech & Language, 94. — 101802 p. (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.

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 31.03.2026.
Формат 60x881/8.
20,0 п.л. Заказ 1.

«Central Asian Academic Research Center» LLP
Алматы, Шевченко к-сі, 28