

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 Жұмаіл 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 № **KZ77VPU00121154**. 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. Tіze MPT патологиясын жіктеуге арналған көп тапсырмалы терең оқыту.....	139
Исахов 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
Кошкарров 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. Көпанныаторлы 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). 233–254

<https://doi.org/10.32014/2026.2518-1726.461>

IRSTI 20.23.27

UDC 004.056.53:004.8

© Makatov Ye.K.^{*1}, Abdul Razaque², Ismukanova A.N.¹, Салыкова А.К.¹,
Glock E.S.¹, 2026.

¹Sh. Ualikhanov Kokshetau University Kokshetau, Kazakhstan;

²Seton Hall University, New Jersey, USA.

E-mail: yerkhanmakatov@gmail.com

MACHINE LEARNING-BASED MODEL FOR PHISHING URL DETECTION IN ONLINE SOCIAL NETWORKS

Makatov Yerkhan — Master of Pedagogical Sciences, Senior Lecturer, Sh. Ualikhanov Kokshetau University, Kokshetau, Kazakhstan,

E-mail: yerkhanmakatov@gmail.com, ORCID ID: <https://orcid.org/0000-0002-9223-3594>;

Abdul Razaque — Doctor of Computer Science, Professor, Seton Hall University, New Jersey, USA,
E-mail: arazaque@my.bridgeport.edu, ORCID ID: <https://orcid.org/0000-0003-0409-3526>;

Ismukanova Aigerim — Master of Technical Sciences, Senior Lecturer, Sh. Ualikhanov Kokshetau University, Kokshetau, Kazakhstan,

E-mail: aigera_ismukan@mail.ru, ORCID ID: <https://orcid.org/0009-0002-7156-8317>;

Salykova Aigerim — Master of Technical Sciences, Senior Lecturer, Sh. Ualikhanov Kokshetau University, Kokshetau, Kazakhstan,

E-mail: Aigera0709listru@gmail.com, ORCID ID: <https://orcid.org/0000-0002-2273-7666>;

Glock Elena — Master of Technical Sciences, Senior Lecturer, Sh. Ualikhanov Kokshetau University, Kokshetau, Kazakhstan,

E-mail: Glock_E@mail.ru, ORCID ID: <https://orcid.org/0009-0003-1960-5341>.

Abstract. Phishing URLs remain one of the most prevalent cybersecurity threats in modern digital environments. The rapid growth of online social networks (OSNs) and web-based platforms has facilitated the widespread deployment of phishing attacks aimed at stealing users' sensitive information, distributing malicious content, and redirecting users to fraudulent websites. Conventional phishing detection techniques, including WHOIS records, DNS information, and blacklist-based approaches, are effective in identifying known phishing domains; however, they often exhibit limited performance when detecting newly emerging or rapidly evolving phishing URLs. Furthermore, the reliance of such approaches on external services increases detection latency and limits system scalability. This study proposes a real-time phishing URL detection system based on the lexical features of URLs. To evaluate the effectiveness of different machine learning approaches, a comparative analysis was conducted using character-level TF-IDF vectorization

in conjunction with Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Random Forest, and Extreme Gradient Boosting (XGBoost) classifiers. The experiments were performed using the “Phishing and Legitimate” URLs dataset obtained from the Kaggle platform. Following data preprocessing, a total of 808,042 URLs were retained for model training and evaluation. The URLs were transformed into a TF-IDF feature space based on character n-grams with lengths ranging from 3 to 5. The experimental results demonstrated that the SVM classifier achieved the best classification performance, with an Accuracy of 97.69%, Precision of 97.41%, Recall of 98.25%, F1-score of 97.83%, and ROC-AUC of 99.66%. The average processing time per URL was approximately 1.9 ms. The proposed approach operates independently of external network services while maintaining high accuracy in phishing URL detection. The findings suggest that the proposed method can be effectively deployed for the automated detection of malicious links in online social networks OSNs, messaging platforms, and corporate cybersecurity systems.

Keywords: phishing URLs, phishing detection, machine learning, TF-IDF vectorization, cybersecurity

***Acknowledgments.** The authors express their sincere gratitude to the Department of Information and Communication Technologies of Sh. Ualikhanov Kokshetau University for the scientific and methodological support provided during the course of this research.*

For citations: Makatov Ye.K., Abdul Razaque, Ismukanova A.N., Салыкова А.К., Glock E.S. Machine learning-based model for phishing url detection in online social networks. Academic Scientific Journal of Computer Science, 2026. — No.3. — P. 233–254. DOI: <https://doi.org/10.32014/2026.2518-1726.461>

© Мақатов Е.К.^{1*}, Abdul Razaque², Исмуканова А.Н.¹, Салыкова А.К.¹, Глок Е.С.¹, 2026.

¹ Ш. Уәлиханов атындағы Көкшетау университеті, Көкшетау, Қазақстан;

² Сетон Холл университеті, Нью-Джерси, АҚШ.

E-mail: yerkhanmakatov@gmail.com

ОНЛАЙН ӘЛЕУМЕТТІК ЖЕЛІЛЕРДЕГІ ФИШИНГТІК URL-ДАРДЫ АНЫҚТАУҒА АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ НЕГІЗІНДЕГІ МОДЕЛЬ

Мақатов Ерхан — педагогика ғылымдарының магистрі, аға оқытушы, Ш. Уәлиханов атындағы Көкшетау университеті, Көкшетау, Қазақстан,

E-mail: yerkhanmakatov@gmail.com, <https://orcid.org/0000-0002-9223-3594>;

Abdul Razaque — компьютерлік ғылымдар докторы, профессор, Сетон Холл университеті, Нью-Джерси, АҚШ,

E-mail: arazaque@my.bridgeport.edu, <https://orcid.org/0000-0003-0409-3526>;

Исмуканова Айгерім — техника ғылымдарының магистрі, аға оқытушы, Ш. Уәлиханов атындағы Көкшетау университеті, Көкшетау, Қазақстан,

E-mail: aigera_ismukan@mail.ru, <https://orcid.org/0009-0002-7156-8317>;

Салыкова Айгерім — техника ғылымдарының магистрі, аға оқытушы, Ш. Уәлиханов атындағы Көкшетау университеті, Көкшетау, Қазақстан,

E-mail: Aigera0709listru@gmail.com, <https://orcid.org/0000-0002-2273-7666>;

Глок Елена — техника ғылымдарының магистрі, аға оқытушы, Ш. Уәлиханов атындағы Көкшетау университеті, Көкшетау, Қазақстан,

E-mail: Glock_E@mail.ru, <https://orcid.org/0009-0003-1960-5341>.

Аннотация. Фишингтік веб-мекенжайлар (URL) қазіргі цифрлық ортадағы ең кең таралған киберқауіптердің бірі болып табылады. Онлайн әлеуметтік желілердің (ОӘЖ) және веб-платформалардың қарқынды дамуы пайдаланушылардың құпия деректерін ұрлауға, зиянды контентті таратуға және пайдаланушыларды жалған веб-сайттарға бағыттауға арналған фишингтік шабуылдардың кеңінен таралуына ықпал етті. WHOIS тіркеу деректеріне, DNS ақпаратына және қара тізімдерге негізделген дәстүрлі анықтау әдістері белгілі фишингтік домендерді анықтауда тиімді болғанымен, жаңадан пайда болған немесе жылдам өзгеретін фишингтік URL-дарды анықтауда олардың тиімділігі шектеулі. Сонымен қатар сыртқы сервистерге тәуелділік анықтау кідірісін арттырып, жүйенің масштабталуын шектейді. Зерттеуде URL-дардың лексикалық белгілеріне негізделген және нақты уақыт режимінде жұмыс істейтін фишингтік URL-дарды анықтау жүйесі ұсынылады. Машиналық оқытудың әртүрлі әдістерінің тиімділігін бағалау үшін символдық деңгейдегі TF-IDF векторизациясы Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Random Forest және Extreme Gradient Boosting (XGBoost) классификаторларымен біріктірілді. Эксперименттер Kaggle платформасындағы ашық «Phishing and Legitimate URLs» деректер жиынтығында жүргізілді. Алдын ала өңдеуден кейін модельдерді оқыту және бағалау үшін 808 042 URL сақталды. URL-дар ұзындығы 3–5 символдан тұратын n-граммалар негізінде TF-IDF белгілер кеңістігіне түрлендірілді. Эксперимент нәтижелері SVM классификаторының ең жоғары нәтиже көрсеткенін көрсетті: Accuracy - 97,69%, Precision - 97,41%, Recall - 98,25%, F1-score - 97,83%, ROC-AUC - 99,66%. Бір URL-ды өңдеудің орташа уақыты шамамен 1,9 мс болды. Ұсынылған тәсіл сыртқы желілік сервистерге тәуелсіз жұмыс істей отырып, фишингтік URL-дарды жоғары дәлдікпен анықтайды. Алынған нәтижелер әдісті онлайн әлеуметтік желілерде, хабар алмасу платформаларында және корпоративтік киберқауіпсіздік жүйелерінде зиянды сілтемелерді автоматты түрде анықтау үшін қолдануға болатынын көрсетеді.

Түйін сөздер: фишингтік URL; фишингті анықтау; машиналық оқыту; TF-IDF векторизациясы; киберқауіпсіздік

© **Макатов Е.К.^{1*}, Abdul Razaque², Исмуканова А.Н.¹, Салыкова А.К.¹, Глок Е.С.¹, 2026.**

¹Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан;

² Университет Сетон Холл, Нью-Джерси, США.

E-mail: yerkhanmakatov@gmail.com

МОДЕЛЬ ОБНАРУЖЕНИЯ ФИШИНГОВЫХ URL В ОНЛАЙН-СОЦИАЛЬНЫХ СЕТЯХ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Макатов Ерхан — магистр педагогических наук, старший преподаватель, Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан,

E-mail: yerkhanmakatov@gmail.com, <https://orcid.org/0000-0002-9223-3594>;

Abdul Razaque — доктор компьютерных наук, профессор, Университет Сетон Холл, Нью-Джерси, США,

E-mail: arazaque@my.bridgeport.edu, <https://orcid.org/0000-0003-0409-3526>;

Исмуканова Айгерим — магистр технических наук, старший преподаватель, Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан,

E-mail: aigera_ismukan@mail.ru, <https://orcid.org/0009-0002-7156-8317>;

Салыкова Айгерим — магистр технических наук, старший преподаватель, Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан,

E-mail: Aigera0709listru@gmail.com, <https://orcid.org/0000-0002-2273-7666>;

Глок Елена — магистр технических наук, старший преподаватель, Кокшетауский университет имени Ш. Уалиханова, Кокшетау, Казахстан,

E-mail: Glok_E@mail.ru, <https://orcid.org/0009-0003-1960-5341>.

Аннотация. Фишинговые веб-адреса (URL) остаются одной из наиболее распространенных угроз кибербезопасности в современной цифровой среде. Стремительное развитие онлайн-социальных сетей (ОСС) и веб-платформ способствует широкому распространению фишинговых атак, направленных на хищение конфиденциальных данных пользователей, распространение вредоносного контента и перенаправление пользователей на мошеннические веб-сайты. Традиционные методы обнаружения, основанные на регистрационных данных WHOIS, DNS-информации и черных списках, эффективны для выявления известных фишинговых доменов, однако зачастую демонстрируют ограниченную эффективность при обнаружении новых или быстро изменяющихся фишинговых URL. Кроме того, зависимость от внешних сервисов увеличивает задержку обнаружения и ограничивает масштабируемость системы. В исследовании предложена система обнаружения фишинговых URL в режиме реального времени, основанная на их лексических характеристиках. Для сравнительной оценки различных методов машинного обучения символьная TF-IDF-векторизация применялась совместно с моделями Logistic Regression, Naive Bayes, Support Vector Machine (SVM), Random Forest и Extreme Gradient Boosting (XGBoost). Эксперименты проводились на открытом наборе данных «Phishing and Legitimate URLs» с платформы Kaggle. После предварительной обработки для обучения и

оценки моделей было сохранено 808 042 URL. URL преобразовывались в пространство признаков TF-IDF на основе символьных n-грамм длиной от 3 до 5 символов. Экспериментальные результаты показали, что наилучшее качество классификации обеспечивает SVM: Accuracy - 97,69%, Precision - 97,41%, Recall - 98,25%, F1-score - 97,83%, ROC-AUC - 99,66%. Среднее время обработки одного URL составило приблизительно 1,9 мс. Предложенный подход функционирует независимо от внешних сетевых сервисов и при этом обеспечивает высокую точность обнаружения фишинговых URL. Полученные результаты подтверждают возможность применения разработанного метода для автоматического выявления вредоносных ссылок в онлайн-социальных сетях, системах обмена сообщениями и корпоративных системах кибербезопасности.

Ключевые слова: фишинговые URL; обнаружение фишинга; машинное обучение; TF-IDF-векторизация; кибербезопасность

Кіріспе. Цифрлық коммуникациялардың қарқынды дамуы киберқауіптердің артуына әкелді. ОӘЖ-дегі негізгі киберқауіптерге фишингтік шабуылдар, әлеуметтік инженерия және зиянды URL-дер жатады (Aktayeva et al., 2023; Makatov et al., 2022). Фишингтік шабуылдар пайдаланушылардың жеке деректерін ұрлауға, зиянды веб-ресурстарды таратуға және жалған веб-сайттарға бағыттауға кеңінен қолданылады. Фишингтік инфрақұрылымдардың жылдам өзгеруі, яғни зиянды домендердің аз уақыт аралығында көп көлемде құрылып, кейін қысқа уақытта өшірілуі олардың анықталуын қиындатады.

Соңғы жылдары фишингтік URL-ды анықтауға арналған зерттеулер саны артқанымен, көпшілігінде жіктеу дәлдігін бағалауға бағытталған. Ал модельдердің нақты уақыт режиміндегі өнімділігі, өңдеу кідірістері, қайта жаңғыртылуы және практикада қолдану ерекшеліктері жеткілікті деңгейде зерттелмеген. Кейбір шешімдердің сыртқы сервистерге тәуелділігі олардың ауқымдылығын төмендетіп, үлкен көлемдегі URL ағынын өңдеу кезінде қосымша қиындықтар туғызады.

Зерттеудің мақсаты — символдық TF-IDF-векторизациясы мен жеңілдетілген машиналық оқыту модельдері негізінде нақты уақыт режимінде жұмыс істейтін фишингтік URL-ды анықтау жүйесін әзірлеу. Зерттеу барысында алдын ала өңделген ірі деректер жиынында әртүрлі машиналық оқыту алгоритмдеріне салыстырмалы талдау жүргіздік. Зерттеудің ғылыми жаңалығы символдық TF-IDF-векторизациясы мен жеңілдетілген машиналық оқыту модельдері негізінде құрылған өңдеу конвейерінің жіктеу сапасы мен орындау жылдамдығын кешенді бағалау. Зерттеуде әртүрлі машиналық оқыту модельдерінің нәтижелері салыстырылып, URL-ды нақты уақыт режимінде сүзуге арналған практикалық шешім ұсындық.

Әдебиеттерге шолу. Қара тізімдерге негізделген әдістер осы уақытқа дейін белгілі фишингтік URL-ды тиімді анықтай алады. Бірақ мұндай тәсілдер

жаңадан тіркелген немесе белгісіз болған фишингтік URL-ды анықтай алмайды. Бұндай әдістердің тиімді жұмыс істеу үшін деректер қорын үнемі жаңарту қажет (Odeh et al., 2025).

WHOIS-талдауға негізделген тәсілдер URL домендердің тіркелу сипаттамаларын талдайды. Оларға доменнің тіркелген күні, қолданылу мерзімі және иесі туралы мәліметтер жатады. Егерде домен деректері құпиялылық механизмдерін қолданса немесе қол жетімсіз болса, бұндай тәсілдер тиімсіз болады (Kaur et al., 2025).

DNS-талдауға негізделген әдістер домендік атаулардың IP-мекенжайларға түрлену ерекшеліктерін, домендік инфрақұрылым сипаттамаларын және күмәнді ресурстармен байланысты желілік белгілерің зерттейді. Алайда жылдам өзгерту технологиялар мен қысқа уақыт өмір сүретін зиянды URL-дер үшін тиімсіз (Deniz et al., 2025).

Эвристикалық ережелерге негізделген тәсілдерде алдын ала анықталған белгілерді қолданады. Оларға URL ұзындығын, ішкі домендер санын, күмәнді кілт сөздерді, IP-мекенжайларды, арнайы символдарды талдау жатады. Мұндай әдістерді іске асыруы қарапайым ал нәтижелері түсінікті. Егер фишингтік шабуылдардың стратегиялары жиі өзгеріп отырса, тәсілдің тиімділігі жоғары болу үшін эвристикалық ережелерді жиі жаңарту қажет (Korkmaz et al., 2020).

Дәстүрлі фишингтік URL-ды анықтау тәсілдерінің шектеулігіне қарамастан бүгінгі күнде қолданыстан шыққан жоқ. Жоғарыдағы айтылған шектеулерді жою мақсатында URL құрылымын автоматты түрде талдауға арналған машиналық оқыту әдістері қарқынды дамуда. Дәстүрлі тәсілдерден ерекшелігі URL белгілері негізінде оқытылып, заңдылықтарды анықтайды және осы уақытқа дейін кездеспеген фишингтік URL-дерді анықтауда ықтималдылығы жоғары (Aljofey et al., 2022; Safi et al., 2023).

Зерттеулердің көбінде URL-дың лексикалық белгілері мен символдық TF-IDF-векторизациясы қолданылады. Лексикалық талдау URL-ды мәтіндік тізбек ретінде қарастырып, домен құрылымын, арнайы символдарды, ішкі домен үлгілерін және символдық n-граммаларды сыртқы желілік дереккөздерге жүгінбей зерттеуге мүмкіндік береді. Ал TF-IDF-векторизациясы URL-ды сандық белгілер кеңістігіне түрлендіріп, оларды машиналық оқыту модельдерінде тиімді пайдалануға жағдай жасайды. Мұндай тәсілдер жоғары дәлдік пен төмен есептеу шығындарын қатар қамтамасыз ете отырып, нақты уақыт режимінде жұмыс істейтін жүйелер үшін тиімді шешім болып табылады (Tashtoush et al., 2024; Benita et al., 2024).

Фишингтік URL-ды жіктеу үшін Logistic Regression, Naive Bayes классификаторы және SVM кеңінен қолданылады. Сызықтық модельдер сирек TF-IDF-векторларымен жұмыс істеуде жоғары тиімділік көрсетіп, есептеу шығындарының төмен болуымен ерекшеленеді. Фишингті анықтауға арналған зерттеулерде Random Forest мен XGBoost сияқты ансамбльдік

әдістер де жиі пайдаланылады. Бұл әдістер белгілер арасындағы күрделі байланыстарды модельдеуге мүмкіндік береді (Tian et al., 2025; Kytidou et al., 2025; Benita et al., 2024).

Жіктеу дәлдігі жоғары болғанына қарамастан, көптеген зерттеулер негізінен модельдердің дербес бағалануымен шектеледі. Ал практикалық енгізу, өңдеу кідірістері, модельдердің тұрақтылығы және өңдеу конвейерінің қайта жаңғыртылуы жеткілікті деңгейде зерттелмеген (Tian et al., 2025).

Соңғы жылдары фишингтік URL-ды анықтау үшін терең оқыту әдістері кеңінен қолданыла бастады. Олардың қатарына конволюциялық нейрондық желілер (CNN), рекурренттік нейрондық желілер (RNN), ұзақ және қысқа мерзімді жад желілері (LSTM), трансформер архитектуралары (Transformers) жатады. Мұндай модельдер URL құрылымының күрделі белгілерін автоматты түрде анықтап, оның ішіндегі тізбектік байланыстарды ескереді (Kulkarni et al., 2024). Терең оқыту модельдерінің басты артықшылығы URL құрылымындағы күрделі белгілерді автоматты түрде үйрену мүмкіндігі. Мұндай тәсілдер фишингтік URL-ды анықтауда жоғары нәтижелер көрсетіп келеді (Opara et al., 2024; Kibriya et al., 2025; Divakaran et al., 2022). Алайда терең оқыту әдістерін қолдану үшін жоғары есептеу ресурстар мен үлкен көлемдегі оқыту деректері қажет (Kibriya et al., 2025; Divakaran et al., 2022). Сондықтан өңдеу кідірісінің төмен болуы, енгізудің қарапайымдылығы және жүйенің ауқымдылығы маңызды болатын жағдайларда жеңілдетілген машиналық оқыту модельдері өзектілігін сақтап келеді. Осы себепті нақты уақыт режимінде жұмыс істейтін жүйелерде TF-IDF-векторизациясы мен жеңілдетілген машиналық оқыту модельдері терең оқыту тәсілдеріне тиімді балама ретінде қарастырылады (Opara et al., 2024; Kibriya et al., 2025; Divakaran et al., 2022).

Фишингтік URL-ды анықтау саласында айтарлықтай жетістіктерге қол жеткізілгеніне қарамастан, қолданыстағы зерттеулерде әлі де бірқатар шектеулер бар.

Біріншіден, басым бөлік зерттеулерде тек Accuracy, Precision, Recall және F1-score сияқты жіктеу метрикаларын бағалаумен шектеледі.

Екіншіден, көптеген зерттеулерде ұсынылған тәсілдердің тәжірибелік енгізілуін қайта жаңғырту мәселелері толық қарастырылмайды.

Үшіншіден, кейбір фишингті анықтау жүйелері WHOIS, DNS, қара тізімдер және веб-беттерді талдау сияқты сыртқы дереккөздерге тәуелді болады.

Төртіншіден, зерттеулерде модельдердің тұрақтылығына жан-жақты талдау жүргізілмейді.

Осыған байланысты жоғары жіктеу сапасын, есептеу тиімділігін, өңдеу конвейерінің қайта жаңғыртылуын және үлкен көлемдегі URL-ды өңдеу кезіндегі тұрақтылықты қатар қамтамасыз ететін тәсілдерді әзірлеу өзекті ғылыми міндет болып қала береді. Сонымен қатар ОӘЖ-дегі пайдаланушыларды қорғауға арналған заманауи қауіпсіздік архитектуралары фишингтік қауіптерді ерте кезеңде анықтайтын механизмдерді қамтуы тиіс.

Әлеуметтік платформалардың қауіпсіздігін қамтамасыз етуге арналған зерттеу көпдеңгейлі қорғаныс механизмдерін қолданудың маңыздылығын көрсетеді (Makatov et al., 2026). Алдыңғы зерттеуіміздегі ОӘЖ қауіпсіздік архитектурасына фишингтік URL-дарды нақты уақыт режимінде анықтауға арналған модельді енгізу, жұмысымыздың қажеттілігі мен практикалық маңыздылығын артады.

Материалдар мен негізгі әдістер. Зерттеу барысында URL құрылымының лексикалық сипаттамалары мен машиналық оқыту әдістеріне негізделген фишингтік URL-дарды анықтау жүйесі әзірленді. Жүйе пайдаланушы хабарламасынан URL-ды бөліп алуды, алдын ала өңдеуді, белгілер кеңістігін қалыптастыруды, машиналық оқыту моделі арқылы бинарлық жіктеуді және күмәнді URL анықталған жағдайда ескерту хабарламасын қалыптастыруды қамтитын өңдеу конвейері ретінде құрылды.



Сурет 1 – Фишингтік URL-ды анықтау жүйесінің өңдеу конвейері

Алғашқы кезеңде пайдаланушы хабарламасынан URL-дар шығарылып, алдын ала өңдеуден өткізілді. Одан кейін URL-дар character-level TF-IDF-векторизациясы арқылы сандық белгілер кеңістігіне түрлендірілді. Алынған белгілер машиналық оқыту классификаторына беріліп, URL фишингтік немесе сенімді классқа жіктелді. Соңғы кезеңде күмәнді деп анықталған URL үшін ескерту хабарламасы қалыптастырылды. Негізгі жіктеу архитектурасы URL құрылымынан алынатын белгілерге негізделді және inference кезеңінде сыртқы WHOIS, DNS немесе blacklist сервистеріне сұрау жіберуді талап етпеді.

Эксперимент үшін Kaggle платформасындағы «Phishing and Legitimate URLs» ашық деректер жиыны пайдаланылды. Деректер жиынында URL және оған сәйкес бинарлық класс белгісі берілген: 0 – фишингтік URL, 1 – сенімді URL. Бастапқы деректер жиыны 822 010 жазбадан тұрды, оның 394 982-сі фишингтік және 427 028-і сенімді URL болды. Деректерді дайындау барысында

missing URL, empty URL, missing class label, exact duplicate URL және label conflict жағдайлары тексерілді. Қайталанатын URL-дар жойылғаннан кейін қалыптастырылған бірегей жазбалар негізгі эксперименттік деректер жиыны ретінде пайдаланылды. Деректер жиынының қолжетімді metadata-сы URL-дардың бастапқы collection period-ын сенімді анықтауға мүмкіндік бермегендіктен, temporal split қолданылған жоқ. Деректер random stratified holdout әдісімен training және test жиындарына 80:20 қатынасында бөлінді (random_state=42). Бөлудің дұрыстығын тексеру үшін row conservation, class stratification, row_id overlap және exact URL overlap аудиттері жүргізілді. Random URL-level split кезінде domain-level тәуелсіздік автоматты түрде қамтамасыз етілмейтіндіктен, hostname және registrable domain деңгейлерінде қосымша overlap audit орындалды. Анықталған domain overlap әсерін бақылау үшін training partition ішінде registrable domain бойынша топтастырылған cross-validation қолданылды.

Негізгі белгілерді ұсыну тәсілі ретінде character-level TF-IDF-векторизациясы қолданылды. Әр URL символдар тізбегі ретінде қарастырылып, ұзындығы 3-тен 5-ке дейінгі character n-gram-дар қалыптастырылды. Белгілер кеңістігінің жоғарғы шегі max_features=40000 параметрімен шектелді. Data leakage-гі болдырмау үшін TF-IDF vocabulary және оған қатысты статистикалар тек training деректерінде fit жасалып, test деректеріне дайын transformation қолданылды.

TF-IDF representation-нан бөлек, 24 қолмен анықталған лексикалық белгі есептелді: url_length, domain_dots, hyphen_count, underscore_count, slash_count, question_count, equal_count, ampersand_count, digit_count, letter_count, keyword_login, keyword_verify, keyword_secure, keyword_account, keyword_update, keyword_bank, keyword_paypal, keyword_signin, keyword_confirm, keyword_password, has_https, has_ip, is_shortened және shannon_entropy. Бұл белгілер URL құрылымын, символдық құрамын, фишингке тән кілт сөздердің кездесуін, HTTPS белгісін, IP-мекенжайдың қолданылуын, қысқартылған URL белгісін және символдар үлестірімінің энтропиясын сипаттау үшін пайдаланылды.

Фишингтік URL-дарды жіктеу үшін Logistic Regression, Multinomial Naive Bayes, Linear SVM, Random Forest және XGBoost алгоритмдері зерттелді. Модельдерге берілетін деректер бірдей preprocessing және feature-generation хаттамасы бойынша қалыптастырылды. Logistic Regression сызықтық базалық модель, Multinomial Naive Bayes ықтималдыққа негізделген классификатор, Linear SVM жоғары өлшемді сирек белгілер кеңістігіне арналған сызықтық классификатор ретінде қолданылды. Random Forest және XGBoost ансамбльдік тәсілдермен салыстыру үшін енгізілді.

Негізгі модель конфигурациялары алдын ала бекітілді: Logistic Regression үшін max_iter=1000 және n_jobs=-1; Multinomial Naive Bayes үшін MultinomialNB(); Linear SVM үшін LinearSVC(); Random Forest үшін n_estimators=50, random_state=42 және n_jobs=-1 қолданылды. Baseline XGBoost

конфигурациясында `n_estimators=30`, `max_depth=4`, `learning_rate=0.1`, `subsample=0.7`, `colsample_bytree=0.7`, `eval_metric='logloss'`, `random_state=42` және `tree_method='hist'` параметрлері пайдаланылды.

Naive Bayes классификаторының жұмысы Байес теоремасына негізделген:

$$P(c | x) = \frac{P(x|c)P(c)}{P(x)} \quad (1)$$

мұндағы:

- c — URL-дың класы;
- x — белгілер векторы;
- $P(c|x)$ — URL-дың c класына жату ықтималдығы;
- $P(x|c)$ — c класы берілген кездегі x белгілерінің байқалу ықтималдығы;
- $P(c)$ — кластың априорлық ықтималдығы;
- $P(x)$ — x белгілерінің байқалу ықтималдығы.

SVM-де жіктеу сызықтық бөлуші гипержазықтық негізінде жүзеге асырылады:

$$w^T x + b = 0 \quad (2)$$

мұндағы:

- w — салмақтар векторы;
- x — белгілер векторы;
- b — гипержазықтықтың ығысу коэффициенті.

Негізгі random holdout бағалауынан бөлек, модельдердің domain-level generalization қабілетін тексеру үшін training partition ішінде 5-fold domain-grouped cross-validation жүргізілді. Ол StratifiedGroupKFold (`n_splits=5`, `shuffle=True`, `random_state=42`) арқылы орындалды, ал grouping variable ретінде registrable domain пайдаланылды. Бір registrable domain-ға жататын URL-дардың бір fold ішінде training және validation бөліктеріне қатар түсуіне жол берілмеді. Әр fold-та TF-IDF vectorizer тек сол fold-тың training бөлігінде fit жасалып, validation бөлігіне дайын transformation қолданылды.

Cross-validation нәтижелерін сипаттау үшін fold-тар бойынша mean және standard deviation есептелді, ал 95% confidence interval Student's t-distribution негізінде анықталды. Linear SVM және Random Forest бірдей fold-тар бойынша paired design арқылы салыстырылды. Салыстыруда two-sided paired t-test және exact Wilcoxon signed-rank test қолданылып, multiple comparisons үшін Holm correction жүргізілді. Effect size ретінде Cohen's `d` және `z`, paired Hedges-type `g` және rank-biserial correlation есептелді. Fold санының шектеулі болуына байланысты статистикалық тесттер абсолютті басымдықты дәлелдеу құралы ретінде емес, модельдер арасындағы айырмашылықтың тұрақтылығын қосымша сипаттау үшін пайдаланылды.

XGBoost нәтижесінің hyperparameter configuration-ға тәуелділігін тексеру үшін training partition шегінде бөлек tuning процедурасы жүргізілді. `n_estimators`

үшін 50 және 100, max_depth үшін 3, 4 және 5 мәндерінің комбинацияларынан алты конфигурация бағаланды. Барлық конфигурациялар үшін learning_rate=0.10, subsample=0.8, colsample_bytree=0.8, eval_metric='logloss', random_state=42 және tree_method='hist' тұрақты сақталды. Конфигурация training partition ішіндегі 5-fold domain-grouped cross-validation бойынша phishing-class F1-score орташа мәнінің негізінде таңдалды. Test жиыны hyperparameter selection барысында пайдаланылған жоқ және тек таңдалған конфигурацияның қорытынды holdout бағалауы үшін сақталды.

Feature representation-ның үлесін бағалау үшін controlled ablation experiment жүргізілді. Бірдей LinearSVC классификаторымен үш конфигурация салыстырылды: TFIDF_ONLY, LEXICAL_ONLY және TFIDF_PLUS_LEXICAL. Классификатор параметрлері representation-дар арасында өзгертілмеді. Осылайша эксперименттің мақсаты classifier tuning әсерін емес, feature representation өзгерісінің үлесін бөлек бағалау болды.

URL құрылымын әдейі өзгертуге модельдің тұрақтылығын бағалау үшін post-training robustness experiment жүргізілді. Бұл кезеңде оқытылған LinearSVC моделі мен TF-IDF representation қайта оқытылмады, refit жасалмады және classification threshold өзгертілмеді. Бірінші сценарийде URL-дың path, query және fragment бөліктеріндегі символдарға детерминирленген substitution қолданылды: a→4, e→3, i→1, o→0, s→5. Екінші сценарийде URL синтетикалық shortener-style representation-ға түрлендірілді: https://bit.ly/ префиксі және детерминирленген 16-символдық token қолданылды. Бұл сценарий нақты third-party URL-shortening сервисінің redirect resolution механизмін модельдемейді және тек URL representation өзгерісіне модель реакциясын бағалау үшін пайдаланылды. Percent encoding, Unicode/homograph substitutions, multi-stage redirect chains және басқа adversarial transformations осы эксперименттің ауқымына енгізілген жоқ.

Барлық эксперименттер Python бағдарламалау ортасында орындалды. Қолданылған бағдарламалық құралдар мен кітапханалар 1-кестеде көрсетілген.

Кесте 1 – Бағдарламалық орта және қолданылған кітапханалар

Параметр	Мәні
Операциялық жүйе	Windows 10
Python	3.13.13
Scikit-learn	1.8.0
XGBoost	3.2.0
Flask	3.1.3
Pandas	3.0.3
NumPy	2.4.4
Ескерту: авторлар құрастырған	

Scikit-learn деректерді өңдеу, feature transformation, модельдерді оқыту, cross-validation және бағалау процедуралары үшін қолданылды.

Pandas деректерді жүктеу және құрылымдық өңдеу, NumPy сандық есептеулер, XGBoost градиенттік бустинг эксперименттері, ал Flask фишингтік URL-дарды анықтау жүйесінің прототипін әзірлеу үшін пайдаланылды.

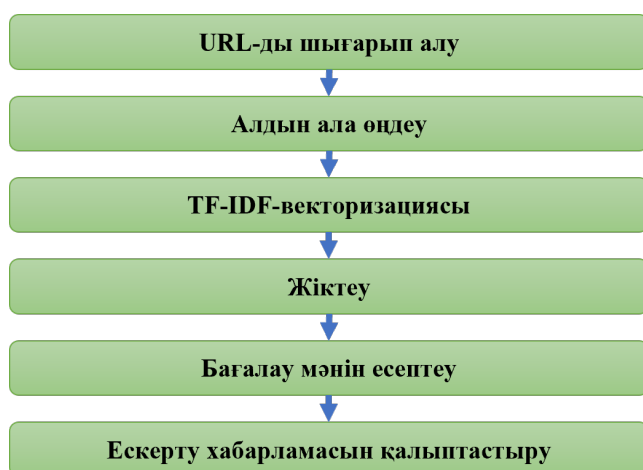
Эксперименттер HP Pavilion Power Laptop 15-cb0xx ноутбугында Windows 10 Pro 64-bit операциялық жүйесінде жүргізілді.

Кесте 2 – Есептеу жүйесінің аппараттық сипаттамалары

Аппараттық параметр	Мәні
Процессор	Intel Core i7-7700HQ @ 2.80 GHz
Жедел жад	16 GB
Архитектура	x64
Жинақтауыш	SSD + HDD
Ескерту: авторлар құрастырған	

Модельдердің жіктеу сапасын бағалау үшін Accuracy, Precision, Recall, F1-score және ROC-AUC метрикалары қолданылды. Precision, Recall және F1-score фишингтік URL класына қатысты есептелді. Деректер жиынында фишингтік URL label=0 ретінде белгіленгендіктен, бұл класс бағалау кезінде positive class ретінде қарастырылды. ROC-AUC фишингтік және сенімді URL-дарды әртүрлі decision threshold жағдайларында ажырату қабілетін бағалау үшін қолданылды.

Модельдердің есептеу тиімділігі training time және бір URL үшін inference time арқылы бағаланды. Толық өңдеу конвейерінің практикалық latency көрсеткішін анықтау үшін URL-ды бөліп алу, preprocessing, TF-IDF transformation, classification, score calculation және warning generation кезеңдері бірге өлшенді.



Сурет 2 — Толық өңдеу конвейерінің кезеңдері

Бір URL-ды жеке өңдеу және бірнеше URL-ды бір хабарлама құрамында өңдеу уақыты өлшенді. Жүйенің практикалық прототипінде preprocessing, vectorization және classification кезеңдері Scikit-learn pipeline арқылы біріктірілді. Linear SVM ықтималдық мәндерін тікелей есептемейтіндіктен, модельдің decision_function шығысы probability ретінде емес, салыстырмалы confidence score ретінде пайдаланылды.

Жасанды интеллект [<https://chatgpt.com/>] технологисы мақаланың мәтінін оқылымын жақсартуға, әдебиеттер тізімін бастапқы деңгейде құруға және эксперименттік зерттеу барысында бағдарламадығы кодта анықталған қателерді талқылауға көмегіші құрал ретінде қолдандық. Авторлар мақаланың ғылыми мазмұны мен нәтижелеріне толық жауапкершілік алады.

Нәтижелер. Бастапқы деректер жиынында 822 010 жазба болды: 394 982 фишингтік және 427 028 сенімді URL. Бастапқы аудит барысында URL немесе класс белгісі жоқ жазбалар анықталған жоқ. 13 968 қайталанатын URL жойылғаннан кейін 808 042 бірегей URL қалды, олардың 381 014-і (47,153%) фишингтік және 427 028-і (52,847%) сенімді URL болды. Бірдей URL-ға қарама-қайшы класс белгісі берілген жағдайлар анықталған жоқ. Домендік аудит барысында 807 901 URL үшін hostname сәтті анықталды, ал 141 URL үшін hostname анықталмады. Жалпы 439 166 бірегей hostname және 317 590 бірегей registrable domain анықталды. Train/test бөлінуінен кейін 29 887 hostname және 31 775 registrable domain екі жиында да кездесті. Тест жиынындағы hostname сәтті анықталған URL-дардың 82 827-сінің hostname-ы оқыту жиынында бұрын кездескен, бұл 51,258%-ды құрады. Осылайша, train және test жиындары арасында exact URL overlap болмағанымен, random holdout кезінде domain-level overlap сақталды.

Зерттеуді WHOIS ақпараты мен DNS-талдауға негізделген базалық детекторды бағалаудан бастадық.

Кесте 3 – WHOIS негізіндегі базалық детектордың нәтижелері

Метрика	Мәні
Accuracy	50.00%
Precision	0.00%
Recall	0.00%
F1-score	0.00%
Орташа талдау уақыты	0.196007 сек
Ескерту: авторлар құрастырған	

Базалық детектор фишингтік URL-дарды ажырата алмады. Нәтиженің құрылымын нақтылау үшін confusion matrix талданды.

Кесте 4 – WHOIS негізіндегі базалық детектордың қателер матрицасы

Нақты класс / Болжанған класс	Заңды	Фишингтік
Заңды	5000	0
Фишингтік	5000	0
Ескерту: авторлар құрастырған		

Қателер матрицасы базалық детектордың барлық URL-ды бір классқа жатқызғанын көрсетті. Сондықтан бұл конфигурацияда WHOIS және DNS ақпараты фишингтік және сенімді URL-дарды ажыратуға жеткіліксіз болды.

Базалық тәсілді бағалағаннан кейін символдық TF-IDF-векторизациясы негізінде оқытылған машиналық оқыту модельдеріне салыстырмалы талдау жүргізілді.

Кесте 5 – Машиналық оқыту модельдерінің салыстырмалы талдау нәтижелері

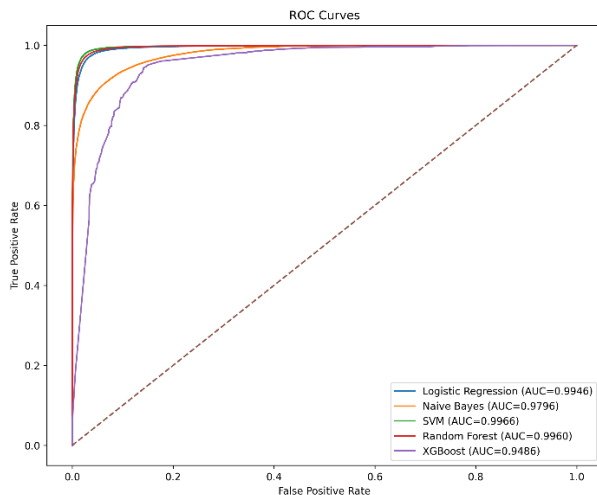
Модель	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	96.84%	96.35%	97.72%	97.03%	99.46%
Naïve Bayes классификаторы	91.51%	89.72%	94.80%	92.19%	97.96%
SVM	97.69%	97.41%	98.25%	97.83%	99.66%
Random Forest	97.26%	97.11%	97.72%	97.42%	99.60%
XGBoost	89.43%	89.44%	90.70%	90.07%	94.86%
Ескерту: авторлар құрастырған					

Random holdout бағалауында Linear SVM ең жоғары нүктелік көрсеткіштерді көрсетті, ал Random Forest нәтижелері оған жақын болды. Сондықтан бір реттік holdout нәтижесіне сүйеніп модельдердің статистикалық басымдығы туралы қорытынды жасамау үшін қосымша domain-grouped cross-validation жүргізілді.

Кесте 6 – Domain-grouped cross-validation нәтижелері

Модель	Accuracy, mean±SD	F1-score, mean±SD	ROC-AUC, mean±SD
Linear SVM	95,15±1,07%	94,92±1,06%	98,98±0,32%
Random Forest	94,87±0,82%	94,54±0,85%	98,71±0,25%
Ескерту: авторлар құрастырған			

Domain-grouped cross-validation екі модельдің нәтижелері бір-біріне жақын екенін көрсетті. Accuracy және F1-score бойынша confidence interval нөлді қамтыды және paired statistical tests статистикалық мәнді айырмашылықты растаған жоқ. Recall және ROC-AUC бойынша қорытынды қолданылған статистикалық тестке тәуелді болды. Сондықтан бес fold нәтижесі Linear SVM-нің Random Forest-тен абсолютті статистикалық басымдығы туралы қорытынды жасауға негіз бермейді. Модельдердің ажырату қабілетін қосымша бағалау үшін ROC қисықтары талданды.



Сурет 3 - Машиналық оқыту модельдерінің ROC қисықтары.

ROC қисықтары Linear SVM, Random Forest және Logistic Regression модельдерінің фишингтік және сенімді URL-дарды ажырату қабілетінің жоғары екенін көрсетті. Naive Bayes нәтижесі салыстырмалы түрде төменірек болды, ал baseline XGBoost қарастырылған конфигурациялар арасында ең төмен ROC-AUC көрсетті. Бұл айырмашылықтың XGBoost конфигурациясына тәуелділігін тексеру үшін қосымша hyperparameter tuning жүргізілді.

Кесте 7 – Baseline және tuned XGBoost нәтижелерін салыстыру

Конфигурация	Accuracy	Precision	Recall	F1-score	ROC-AUC
Baseline XGBoost	89,43%	89,44%	90,70%	90,07%	94,86%
Tuned XGBoost	92,91%	94,38%	90,34%	92,32%	97,68%

Ескерту: авторлар құрастырған.

Hyperparameter tuning XGBoost көрсеткіштерінің басым бөлігін жақсартты. Бұл бастапқы төмен нәтиженің алгоритмнің жалпы қасиеті ретінде емес, қолданылған baseline конфигурациясына сезімтал нәтиже ретінде қарастырылуы керектігін көрсетеді.

Linear SVM жіктеу қателерінің құрылымын бағалау үшін confusion matrix талданды.

Кесте 8 – Linear SVM моделінің қателер матрицасы

Нақты класс / Болжанған класс	Сенімді	Фишингтік
Фишингтік	2234	73 969
Сенімді	83 913	1493

Ескерту: авторлар құрастырған.

Қателер матрицасы екі класта да дұрыс жіктеулердің басым екенін көрсетті. Қате жағдайлардың сипатын анықтау үшін жалған оң және жалған теріс URL-дарға жеке талдау жүргізілді.

Кесте 9 – Жалған оң және жалған теріс классификациялардың мысалдары

Жалған оң URL-дар	Жалған теріс URL-дар
https://www.ycrwjs.com/	au.rozee.pk/
209.193.86.177/static12.pdf	scripts.webmastersite.net/wsnforum/
www.newyahoobuy.org/	https://tytnetwork.com/backbone/
<i>Ескерту: авторлар құрастырған.</i>	

Жалған оң қателер әдеттен тыс домен құрылымы, IP-мекенжайлардың қолданылуы және ерекше символдық тізбектер кездесетін URL-дарда байқалды. Жалған теріс жағдайлардың бір бөлігі құрылымы бойынша сенімді URL-дарға ұқсас фишингтік URL-дармен байланысты болды. Бұл лексикалық representation үшін екі класс арасындағы құрылымдық ұқсастықтың негізгі қате көздерінің бірі екенін көрсетеді.

Модельдердің есептеу тиімділігін бағалау үшін оқыту уақыты салыстырылды.

Кесте 10 – Машиналық оқыту модельдерінің оқыту уақыты

Модель	Оқыту уақыты
Logistic Regression	115,98 сек
Naive Bayes	108,15 сек
Linear SVM	127,20 сек
Random Forest	2734,65 сек
XGBoost	302,54 сек
<i>Ескерту: авторлар құрастырған.</i>	

Сызықтық модельдер мен Naive Bayes ансамбльдік модельдермен салыстырғанда аз оқыту уақытын қажет етті. Random Forest үшін есептеу шығыны ең жоғары болды, бұл жоғары өлшемді сирек белгілер кеңістігінде көптеген шешім ағаштарын құру қажеттілігімен байланысты.

Модельдердің inference жылдамдығын бағалау үшін бір URL-ды жіктеуге кететін уақыт өлшенді.

Кесте 11 – Бір URL-ды жіктеу уақыты

Модель	Өңдеу уақыты
Logistic Regression	0,000108 сек
Naive Bayes	0,000099 сек
Linear SVM	0,000095 сек
Random Forest	0,000154 сек
XGBoost	0,000157 сек
<i>Ескерту: авторлар құрастырған.</i>	

Барлық модельдер бір URL үшін төмен inference latency көрсетті. Сызықтық модельдер мен Naive Bayes ансамбльдік тәсілдерге қарағанда жылдамырақ жұмыс істеді, бұл сирек TF-IDF representation-мен есептеу күрделілігінің төмен болуымен байланысты.

Character-level TF-IDF және қолмен анықталған лексикалық белгілердің үлесін бағалау үшін controlled feature ablation жүргізілді.

Кесте 12 – Белгілер үлесін талдау нәтижелері

Белгілер конфигурациясы	Accuracy	F1-score	ROC-AUC
TFIDF_ONLY	97,6938%	97,5426%	99,6642%
LEXICAL_ONLY	77,1844%	73,0924%	84,4081%
TFIDF_PLUS_LEXICAL	97,7507%	97,6038%	99,6686%
<i>Ескерту: авторлар құрастырған.</i>			

Ablation нәтижелері негізгі ақпараттық үлесті character-level TF-IDF representation қамтамасыз ететінін көрсетті. Лексикалық белгілерді жеке пайдалану төмен нәтиже берді, ал оларды TF-IDF representation-ға қосу жіктеу сапасын аз ғана жақсартты. Сондықтан қолмен анықталған лексикалық белгілер негізгі representation емес, TF-IDF-ті толықтыратын қосымша белгілер ретінде қарастырылады.

URL құрылымын әдейі өзгерту жағдайындағы тұрақтылықты бағалау үшін frozen-model robustness experiment жүргізілді.

Кесте 13 – Character-substitution robustness нәтижелері

Көрсеткіш	Бастапқы	Character substitution
Бағаланған URL саны	8169	8169
Recall	99,58%	96,94%
Recall өзгерісі	—	–2,64 пайыздық тармақ
Prediction flip	—	218
Evasion rate	—	2,67%
<i>Ескерту: авторлар құрастырған.</i>		

Character substitution модель сапасының төмендеуіне және бастапқыда дұрыс анықталған URL-дардың бір бөлігінің кейін анықталмай қалуына әкелді. Демек, character-level representation қарапайым символдық өзгерістерге толық инвариантты емес.

Synthetic shortener-style сценарийінде барлық бағаланған фишингтік URL фишингтік классқа жатқызылды. Алайда бұл эксперимент нақты third-party shortening сервисінің redirect resolution механизмін модельдемейді. Сондықтан алынған нәтиже нақты URL-shortening сервисіне немесе redirect chains жағдайларына тікелей жалпыланбайды.

Ұсынылған тәсілдің практикалық қолданылуын бағалау үшін толық өңдеу конвейерінің latency көрсеткіштері өлшенді.

Кесте 14 – Толық өңдеу конвейерінің latency нәтижелері

Өңдеу сценарийі	Уақыт
Бір URL	шамамен 1,9 мс
10 URL бар хабарлама	шамамен 11 мс
<i>Ескерту: авторлар құрастырған.</i>	

Толық конвейердің өлшенген latency көрсеткіштері зерттелген аппараттық және бағдарламалық ортада URL-дарды жедел өңдеуге мүмкіндік берді. Практикалық прототипте алдын ала өңдеу, векторизация және жіктеу кезеңдері бір Scikit-learn конвейеріне біріктірілді.

Ұсынылған тәсілдің бірнеше шектеуі бар. Біріншіден, модель URL құрылымына негізделген және веб-бет мазмұнын, визуалдық белгілерді немесе пайдаланушы хабарламасының контекстін ескермейді. Екіншіден, зерттелген тәсіл CNN, RNN, LSTM және Transformer архитектураларымен бірдей эксперименттік хаттама бойынша салыстырылған жоқ. Үшіншіден, robustness evaluation character substitution және synthetic shortener-style transformation сценарийлерімен шектелді; percent encoding, Unicode/homograph substitutions, multi-stage redirect chains және басқа күрделі adversarial transformations бағаланған жоқ. Synthetic shortener-style эксперимент нақты third-party shortening сервисінің redirect resolution механизмін модельдемейді.

Төртіншіден, random URL-level holdout кезінде hostname және registrable-domain overlap сақталды. Сондықтан негізгі holdout нәтижелері толық domain-disjoint generalization ретінде интерпретацияланбайды; бұл мәселе қосымша domain-grouped cross-validation арқылы бақыланды. Бесіншіден, бастапқы деректер жиынының қолжетімді метадеректері URL-дардың сенімді collection period-ын анықтауға мүмкіндік бермеді, сондықтан temporal split жүргізілген жоқ. Алтыншыдан, Linear SVM ықтималдық мәндерін тікелей есептемейді; decision_function шығыстары салыстырмалы confidence score ретінде пайдаланылды.

Болашақ зерттеулерде percent encoding, Unicode/homograph transformations және redirect chains сияқты күрделі adversarial өзгерістерге тұрақтылықты бағалау, domain-disjoint және temporal validation хаттамаларын кеңейту, URL белгілерін веб-бет мазмұны және визуалдық сипаттамалармен біріктіретін мультимодальды модельдерді зерттеу, сондай-ақ жүйені нақты веб-трафик жағдайында валидациялау қажет.

Талқылау. Зерттеу нәтижелері character-level TF-IDF-векторизациясының URL құрылымындағы ақпаратты машиналық оқыту модельдеріне жарамды белгілер кеңістігіне тиімді түрлендіретінін көрсетті. Feature ablation негізгі ақпараттық үлесті TF-IDF representation қамтамасыз ететінін, ал қолмен анықталған лексикалық белгілердің қосымша үлесі шектеулі екенін көрсетті. Бұл фишингтік және сенімді URL-дарды ажыратуда жекелеген қолмен анықталған сипаттамалардан гөрі URL құрамындағы символдық тізбектер мен олардың комбинацияларының маңызды екенін көрсетеді.

Бұл нәтиже фишингтік URL-ды анықтауда TF-IDF негізіндегі белгілерді

пайдаланатын алдыңғы зерттеулермен сәйкес келеді. TF-IDF пен Random Forest комбинациясын қолданса, біздің зерттеуде character-level TF-IDF әртүрлі машиналық оқыту модельдері үшін ортақ representation ретінде бағаланды (Benita et al. 2024). Random holdout-та Linear SVM жоғары нүктелік көрсеткіштер көрсеткенімен, domain-grouped cross-validation және paired statistical analysis оның Random Forest-тен абсолютті статистикалық басымдығын растаған жоқ. Сондықтан модельдерді бір ғана train/test бөлінуіндегі көрсеткіштер бойынша бағалау жеткіліксіз және олардың жаңа домендерге жалпылану қабілетін бөлек тексеру қажет.

Бұл мәселе train/test domain overlap аудитінің нәтижелерімен де байланысты. Exact URL overlap анықталмағанымен, бір hostname немесе registrable domain-ға жататын әртүрлі URL-дар training және test жиындарында қатар кездесті. Мұндай жағдай random URL-level holdout көрсеткіштеріне белгілі бір domain familiarity әсерін енгізуі мүмкін. Осы себепті негізгі holdout нәтижелері толық domain-disjoint generalization ретінде қарастырылмайды. Registrable domain бойынша топтастырылған cross-validation бұл тәуелділікті азайтып, модельдердің бұрын кездеспеген домендерге жалпылануын қатаңырақ жағдайда бағалауға мүмкіндік берді.

Кателерді талдау фишингтік және сенімді URL-дардың құрылымдық ұқсастығы лексикалық модельдер үшін негізгі қиындықтардың бірі екенін көрсетті. Фишингтік URL сенімді ресурстарға тән атауларды, кілт сөздерді немесе URL құрылымдарын имитациялаған жағдайда тек URL деңгейіндегі representation арқылы оларды ажырату күрделене түседі. Фишингтік ресурстардың заңды веб-ресурстарға ұқсас белгілерді пайдалануынан туындайтын осындай қиындықтар басқа зерттеулерде де қарастырылған (Opara et al., 2024; Kibriya et al., 2025).

Robustness evaluation бұл шектеуді қосымша нақтылады. Character substitution қолданылған кезде модель сапасының төмендеуі character-level representation-ның әдейі жасалған символдық өзгерістерге толық инвариантты емес екенін көрсетті. Сонымен бірге зерттелген сценарийлер adversarial URL manipulation тәсілдерінің барлық спектрін қамтымайды. Percent encoding, Unicode/homograph substitutions және multi-stage redirect chains сияқты күрделі transformation түрлері бағаланған жоқ. Synthetic shortener-style сценарийі де нақты third-party shortening сервисінің redirect resolution механизмін модельдемейді. Сондықтан бұл нәтижелер тек зерттелген transformation түрлерінің шегінде интерпретациялануы тиіс.

XGBoost бойынша жүргізілген қосымша эксперимент модель нәтижесінің конфигурацияға сезімтал екенін көрсетті. Training partition шегінде орындалған hyperparameter tuning бастапқы көрсеткіштерді жақсартқандықтан, baseline XGBoost нәтижесін алгоритмнің жалпы әлсіздігінің дәлелі ретінде қарастыруға болмайды. Бұл жоғары өлшемді және сирек TF-IDF representation-мен жұмыс істейтін алгоритмдерді салыстыру кезінде тек модель атауын емес, оның

конфигурациясын, hyperparameter selection процедурасын және validation protocol-ын да ескеру қажеттігін көрсетеді.

Ұсынылған тәсілдің практикалық ерекшелігі — URL-ды жіктеу кезінде сыртқы желілік сервистердің міндетті болмауы. Кейбір тәсілдер DNS және желілік инфрақұрылым белгілерін пайдаланады (Deniz et al., 2025), ал басқа шешімдер кара тізімдерге сүйенеді (Odeh et al., 2025). Мұндай тәсілдердің жұмысы сыртқы деректердің қолжетімділігі мен өзектілігіне тәуелді болуы мүмкін. Бұл зерттеуде жіктеу URL құрылымынан тікелей алынатын белгілерге негізделгендіктен, inference кезеңінде WHOIS, DNS немесе blacklist сервистеріне сұрау жіберу талап етілмейді.

Терең оқыту мен үлкен тілдік модельдер күрделі representation-дарды автоматты түрде үйренуге мүмкіндік береді, бірақ олардың есептеу ресурстарына қойылатын талаптары жоғары болуы мүмкін (Kibriya et al., 2025). Бұл зерттеуде веб-бет мазмұны, HTML сипаттамалары және визуалдық белгілер пайдаланылмай, URL деңгейіндегі representation зерттелді. Мұндай шектеу өңдеу конвейерін қарапайым етеді, бірақ URL құрылымынан тыс ақпараттың пайдаланылмауы анықтау мүмкіндігін де шектейді. Веб-бет мазмұны мен визуалдық сипаттамаларды интеграциялау қосымша ақпарат көздерін қамтамасыз етуі мүмкін, сондықтан мультимодальды талдау перспективалық бағыт болып табылады (Aljofey et al., 2022; Opara et al., 2024).

Нәтижелерді жалпылауда деректер жиынының уақыттық сипаттамасын да ескеру қажет. Қолжетімді metadata URL-дардың сенімді collection period-ын анықтауға мүмкіндік бермегендіктен, temporal split жүргізілген жоқ. Сондықтан жүргізілген эксперименттер модельдің уақыт бойынша кейін пайда болатын жаңа фишингтік кампанияларға жалпылануын тікелей бағаламайды. Мұндай қабілетті тексеру үшін timestamp-тары сақталған деректер жиындарында temporal validation жүргізу қажет.

Бұл зерттеу авторлардың ОӘЖ пайдаланушыларын қорғауға арналған қауіпсіздік архитектурасы бойынша бұрынғы жұмысының логикалық жалғасы болып табылады (Makatov et al., 2026). Алдыңғы зерттеуде ОӘЖ-ді қорғаудың тұжырымдамалық архитектурасы мен негізгі қағидаттары ұсынылған болса, осы жұмыста сол архитектура аясында фишингтік URL-дарды автоматты анықтауға арналған практикалық компонент зерттелді. Осы тұрғыдан ұсынылған модель ОӘЖ пайдаланушыларын қорғауға бағытталған кешенді қауіпсіздік жүйесінің URL деңгейіндегі анықтау компоненті ретінде қарастырылуы мүмкін.

Қорытынды. Бұл жұмыста character-level TF-IDF-векторизациясы мен машиналық оқыту модельдеріне негізделген фишингтік URL-ды анықтау тәсілі зерттелді. Эксперименттік дизайн негізгі random stratified holdout бағалауын domain-level аудитпен, domain-grouped cross-validation-мен, статистикалық салыстырумен, hyperparameter sensitivity analysis-пен, feature ablation және robustness evaluation-мен толықтырды. Мұндай дизайн модельдердің жіктеу

сапасымен қатар нәтижелердің domain деңгейіндегі тұрақтылығы мен URL құрылымының өзгерістеріне сезімталдығын бағалауға мүмкіндік берді.

Зерттеу character-level TF-IDF representation-ның фишингтік URL-ды анықтауда негізгі ақпараттық үлесті қамтамасыз ететінін көрсетті. Сонымен бірге нәтижелер модельдерді тек бір реттік random holdout көрсеткіштері бойынша салыстыру жеткіліксіз екенін көрсетті. Domain-level тәуелсіздік, модель конфигурациясы және URL құрылымының әдейі өзгеруі сияқты факторлар эксперименттік бағалаудың нәтижесіне әсер етуі мүмкін және модельдің жалпылану қабілетін бағалау кезінде ескерілуі тиіс.

Ұсынылған тәсіл inference кезеңінде сыртқы WHOIS, DNS және blacklist сервистерін қажет етпейді және зерттелген есептеу ортасында URL-дарды жедел өңдеуге мүмкіндік берді. Сондықтан оны ОӘЖ, хабар алмасу платформалары және корпоративтік киберқауіпсіздік жүйелеріндегі URL ағынын бастапқы автоматты талдау компоненті ретінде қарастыруға болады. Бұл қорытынды зерттелген деректер жиыны, есептеу ортасы және эксперименттік сценарийлер шегінде жасалады.

Болашақ зерттеулер timestamp-тары сақталған деректерде temporal validation жүргізуге, толық domain-disjoint бағалауды кеңейтуге, percent encoding, Unicode/homograph substitutions және redirect chains сияқты күрделі adversarial transformations жағдайындағы тұрақтылықты тексеруге, сондай-ақ URL құрылымын веб-бет мазмұны мен визуалдық сипаттамалармен біріктіретін мультимодальды тәсілдерді зерттеуге бағытталады. Бұл модельдің жаңа домендерге, уақыт бойынша өзгертін фишингтік кампанияларға және күрделі айналып өту тәсілдеріне жалпылану қабілетін неғұрлым қатаң бағалауға мүмкіндік береді.

Мүдделер қақтығысы. Авторлар зерттеу нәтижелеріне әсер етуі мүмкін қаржылық, коммерциялық немесе өзге де мүдделер қақтығысының жоқ екенін мәлімдейді.

Зерттеудің қайта жаңғыртылуын қамтамасыз ету мақсатында бағдарламалық кодпен эксперименттік материалдар GitHub репозиторийінде жарияланды: https://github.com/ItBilim/nanrk_behavioral-risk-ranking.

References

Makatov Y.E., Aktayeva A.I., Niyazova R., Konyrkhanova A., Beissekov A., Aitkenova M., Pleskachev D., Zhamankarin M. (2022) Evaluation of the contemporary state of social network security issues: publication content analysis. Journal of Theoretical and Applied Information Technology. — Vol. 100. — No. 21. — P. 6598-6608. <https://jatih.org/volumes/Vol100No21/27Vol100No21.pdf> (in Eng.)

Aktayeva A., Makatov Y., Tulegenovna A.K., Dautov A., Niyazova, Vol. 8. — No. 10. 156. — P. 1-18. <https://doi.org/10.3390/data8100156> (in Eng.)

Aljofey A., Jiang Q., Rasool A., Chen H., Liu W., Qu Q., Wang Y. (2022) An effective detection approach for phishing websites using URL and HTML features. Scientific Reports, Vol. 12. — No. 1. — 8842. — P. 1–19. <https://doi.org/10.1038/s41598-022-10841-5> (in Eng.)

Benita J., Balaji C., Kamalesh D., Krishna D.S., Narasimharao C.M. (2024) Hybrid Approach

for Malicious URL Detection by Integrating TF-IDF and Random Forest Techniques. In: 2024 4th International Conference on Mobile Networks and Wireless Communications (ICMNWC). IEEE. — P. 1-7. <https://doi.org/10.1109/ICMNWC63764.2024.10872386> (in Eng.)

Deniz F., Nabeel M., Yu T., Khalil I. (2025) MANTIS: Detection of Zero-Day Malicious Domains Leveraging Low Reputed Hosting Infrastructure. In: 2025 IEEE Symposium on Security and Privacy (SP). IEEE. — P. 1789-1807. <https://doi.org/10.1109/SP61157.2025.00067> (in Eng.)

Divakaran D.M., Oest A. (2022) Phishing detection leveraging machine learning and deep learning: A review. IEEE Security & Privacy, Vol. 20. — No. 5. — P. 86-95. <https://doi.org/10.1109/MSEC.2022.3175225> (in Eng.)

Kaur K., Jain A.K. (2025) A Survey on Phishing Attack Taxonomy, Detection Techniques, Datasets, and Security Measures. Journal of Applied Security Research, Vol. 20. — No. 4. — P. 506-557. <https://doi.org/10.1080/19361610.2025.2508726> (in Eng.)

Kibriya H., Amin R., Alshamrani S.S., Rehman S., Hassan M., Alsubaei F.S. (2025) Lightweight malicious URL detection using deep learning and large language models. Scientific Reports, Vol. 15. — No. 1. — Article 43044. — P. 1-18. <https://doi.org/10.1038/s41598-025-26653-2> (in Eng.)

Korkmaz M., Sahingoz O.K., Diri B. (2020) Detection of phishing websites by using machine learning-based URL analysis. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE. — P. 1-7. <https://doi.org/10.1109/ICCCNT49239.2020.9225561> (in Eng.)

Kulkarni A., Balachandran V., Das T. (2024) Phishing webpage detection: Unveiling the threat landscape and investigating detection techniques. IEEE Communications Surveys & Tutorials, Vol. 27. — No. 2. — P. 974-1007. <https://doi.org/10.1109/COMST.2024.3441752> (in Eng.)

Kytidou E., Tsirikli T., Drosatos G., Rantos K. (2025) Machine learning techniques for phishing detection: A review of methods, challenges, and future directions. Intelligent Decision Technologies, Vol. 19. — No. 6. — P. 4356-4379. <https://doi.org/10.1177/18724981251366763> (in Eng.)

Makarov Y.K., Razaque A., Makatova A.Y. (2026) Cognitive model for protecting social media users: Architecture and principles. Bulletin of the Civil Aviation Academy, Vol. 1. — No. 40. — P. 189-197. https://doi.org/10.53364/24138614_2026_40_1_16 (in Eng.)

Odeh N., Hijazi S. (2025) Detection and prevention of phishing short URLs using machine learning and blacklist approaches. International Journal of Wireless and Microwave Technologies, Vol. 15. — No. 3. — P. 37-53. <https://doi.org/10.5815/ijwmt.2025.03.03> (in Eng.)

Opara C., Chen Y., Wei B. (2024) Look before you leap: Detecting phishing web pages by exploiting raw URL and HTML characteristics. Expert Systems with Applications, Vol. 236. — 121183. <https://doi.org/10.1016/j.eswa.2023.121183> (in Eng.)

Safi A., Singh S. (2023) A systematic literature review on phishing website detection techniques. Journal of King Saud University – Computer and Information Sciences, Vol. 35. — No. 2. — P. 590-611. <https://doi.org/10.1016/j.jksuci.2023.01.004> (in Eng.)

Tashtoush Y., Alajlouni M., Albalas F., Darwish O. (2024) Exploring low-level statistical features of n-grams in phishing URLs: a comparative analysis with high-level features. Cluster Computing, Vol. 27. — No. 10. — P. 13717-13736. <https://doi.org/10.1007/s10586-024-04655-5> (in Eng.)

Tian Y., Yu Y., Sun J., Wang Y. (2025) From past to present: A survey of malicious URL detection techniques, datasets and code repositories. Computer Science Review, Vol. 58. — 100810 p. <https://doi.org/10.1016/j.cosrev.2025.100810> (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.