

MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN
ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН
KAZAKHSTAN



**INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION
TECHNOLOGIES**

Published since 2020.
Volume 7. 1 (25). 2026
January–March

**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ
ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

2020 жылдан бері шығарылады
Том 7. 1 (25). 2026
Қаңтар-Наурыз

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

Издается с 2020 г.
Том 7. 1 (25). 2026
Январь-Март

Свидетельство о постановке на учет периодического печатного издания в Министерство информации и общественного развития Республики Казахстан № KZ82VPY00020475, выданное от 20.02.2020 г.

Зарегистрировано в Международном центре регистрации серийных изданий ISSN (ЮНЕСКО, Париж, Франция). ISSN 2708–2032 (print), ISSN 2708–2040 (online)

Журнал входит в Перечень научных изданий, рекомендуемых КОКНВО МНВО РК для публикации основных результатов научной деятельности.

EDITOR-IN-CHIEF:

Kateryna Kolesnikova — Doctor of Technical Sciences, professor, Vice-Rector for Research, International Information Technology University (Kazakhstan)

DEPUTY EDITOR-IN-CHIEF:

Madina Ipalakova — Candidate of Technical Sciences, associate professor, Director of the Research Department, International Information Technology University (Kazakhstan)

EDITORIAL BOARD:

Abdul Razak — PhD, professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)

Lucio Tommaso De Paolis — Director of the R&D Department of the AVR Laboratory, Department of Engineering for Innovation, University of Salento (Italy)

Liz Bacon — Professor, Deputy Vice-Chancellor, Abertay University (United Kingdom)

Michele Pagano — PhD, Professor, University of Pisa (Italy)

Mukhtarbay Otelbayev — Doctor of Physical and Mathematical Sciences, professor, academician of the National Academy of Sciences of the Republic of Kazakhstan, professor of the Department of Mathematical and Computer Modeling, International Information Technology University (Kazakhstan)

Bolatbek Rysbauly — Doctor of Physical and Mathematical Sciences, professor, professor of the Department of Computing and Data Science, Astana IT University (Kazakhstan)

Yevgeniya Daineko — PhD, research professor, Department of Information Systems, International Information Technology University (Kazakhstan)

Nurzhan Duzbayev — PhD, associate professor, Vice-Rector for Digitalization and Innovation, International Information Technology University (Kazakhstan)

Bakhtgerai Sinchev — Doctor of Technical Sciences, professor, Department of Information Systems, International Information Technology University (Kazakhstan)

Nurgul Seilova — Candidate of Technical Sciences, Dean of the Faculty of Computer Technologies and Cybersecurity, International Information Technology University (Kazakhstan)

Ardak Mukhamediyeva — Candidate of Economic Sciences, Dean of the Faculty of Business, Media and Management, International Information Technology University (Kazakhstan)

Zamira Abdikalikova — PhD, associate professor, Head of the Department of Mathematical and Computer Modeling, International Information Technology University (Kazakhstan)

Yerlan Shildibekov — PhD, associate professor, Head of the Department of Economics and Business, International Information Technology University (Kazakhstan)

Damilya Yeskendirowa — Candidate of Technical Sciences, associate professor, Head of the Department of Cybersecurity, International Information Technology University (Kazakhstan)

Aigul Niyazgulova — Candidate of Philological Sciences, Professor, Head of the Department of Media Communications and History of Kazakhstan, International Information Technology University (Kazakhstan)

Altai Aitmagambetov — Candidate of Technical Sciences, Professor, Department of Radio Engineering, Electronics and Telecommunications, International Information Technology University (Kazakhstan)

Yelena Bakhtiyarova — Candidate of Technical Sciences, associate professor, Head of the Department of Radio Engineering, Electronics and Telecommunications, International Information Technology University (Kazakhstan)

Kanibek Sansyzbay — PhD, research professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)

Sakhybay Tynymbayev — Candidate of Technical Sciences, Professor, Research Professor, Department of Computer Engineering, International Information Technology University (Kazakhstan)

Ali Abd Almisreb — PhD, associate professor, Department of Cybersecurity, International Information Technology University (Kazakhstan)

Mohamed Ahmed Hamada — PhD, associate professor, Department of Information Systems, International Information Technology University (Kazakhstan)

Yang Im Chu — PhD, Professor, Gachon University (South Korea)

Tadeusz Wallas — PhD, Vice-Rector, Adam Mickiewicz University (Poland)

Orken Mamyrbayev — PhD, Deputy Director for Science, RSE Institute of Information and Computational Technologies, Committee for Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Kazakhstan)

Sergey Bushuyev — Doctor of Technical Sciences, professor, Director of the Ukrainian Project Management Association "UKRNET," Head of the Department of Project Management, Kyiv National University of Construction and Architecture (Ukraine)

Svetlana Beloshitskaya — Doctor of Technical Sciences, professor, Department of Computing and Data Science, Astana IT University (Kazakhstan)

MANAGING EDITOR

Raushan Mrzabayeva — Master of Science, editor, International Information Technology University (Kazakhstan)

International Journal of Information and Communication Technologies

Periodicity: 4 times a year.

Languages: Kazakh, Russian, English

DOI prefix: 10.54309

ISSN 2708-2032 (print)

ISSN 2708-2040 (online)

Thematic focus: "Information technology"; "Digital technologies in the development of socio-economic systems"; "Information security and communication technologies".

Distribution: Materials are distributed under the Creative Commons Attribution 4.0

Journal website: <https://journal.iitu.edu.kz>

Owner: International Information Technology University JSC (Almaty).

Copyright: © International Journal of Information and Communication Technologies, 2026

РЕДАКЦИЯ

БАС РЕДАКТОР:

Колесникова Катерина Викторовна — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университетінің ғылыми-зерттеу қызметі жөніндегі проректор (Қазақстан)

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

Ипалакова Мадина Түлегеновна — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университетінің ғылыми-зерттеу қызметі жөніндегі департамент директоры (Қазақстан)

РЕДАКЦИЯЛЫҚ АЛҚА:

Разак Абдул — PhD, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының профессоры (Қазақстан)

Лучино Томмазо де Паолис — Саленто Университеті (Италия) инновация және технологиялық инжиниринг департаменті AVR зертханасының зерттеу және әзірлеу бөлімінің директоры

Лиз Бэкон — профессор, Абертей Университеті (Ұлыбритания) вице-канцлерінің орынбасары

Микеле Пагано — PhD, Пиза Университетінің (Италия) профессоры

Өтелбаев Мұхтарбай Өтелбайұлы — физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының профессоры (Қазақстан)

Рысбайұлы Болатбек — физика-математика ғылымдарының докторы, профессор, Есептеу және деректер ғылымдары департаментінің профессоры, Astana IT University (Қазақстан)

Дайнеко Евгения Александровна — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессор-зерттеушісі (Қазақстан)

Дузаев Нуржан Тоқсуажасвич — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті цифрландыру және инновациялар жөніндегі проректор (Қазақстан)

Синчев Бахтгерей Кусанович — техника ғылымдарының докторы, профессор, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының профессоры (Қазақстан)

Сейлова Нургуль Абадуллаевна — техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университеті компьютерлік технологиялар және киберқауіпсіздік факультетінің деканы (Қазақстан)

Мухамедиева Ардак Габитовна — экономика ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті бизнес медиа және басқару факультетінің деканы (Қазақстан)

Абикаликова Замира Тұрсынбаевна — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті математика және компьютерлік модельдеу кафедрасының меңгерушісі (Қазақстан)

Шильдибеков Ерлан Жаржанович — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті экономика және бизнес кафедрасының меңгерушісі (Қазақстан)

Дамеля Максұтовна Ескендірова — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының меңгерушісі (Қазақстан)

Ниязгулова Айгуль Аскарбековна — филология ғылымдарының кандидаты, доцент, профессор, Халықаралық ақпараттық технологиялар университеті медиакоммуникация және Қазақстан тарихы кафедрасының меңгерушісі (Қазақстан)

Айтмағамбетов Алтай Зуфарович — техника ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының профессоры (Қазақстан)

Бахтиярова Елена Ажибековна — техника ғылымдарының кандидаты, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті радиотехника, электроника және телекоммуникация кафедрасының меңгерушісі (Қазақстан)

Канибек Сансызбай — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының профессор-зерттеушісі (Қазақстан)

Тынымбаев Сахибай — техника ғылымдарының кандидаты, профессор, Халықаралық ақпараттық технологиялар университеті компьютерлік инженерия кафедрасының профессор-зерттеушісі (Қазақстан)

Алмисреб Али Абд — PhD, Халықаралық ақпараттық технологиялар университеті киберқауіпсіздік кафедрасының қауымдастырылған профессоры (Қазақстан)

Мохамед Ахмед Хамада — PhD, Халықаралық ақпараттық технологиялар университеті ақпараттық жүйелер кафедрасының қауымдастырылған профессоры (Қазақстан)

Янг Им Чу — PhD, Гачон университетінің профессоры (Оңтүстік Корея)

Тадеш Валлас — PhD, Адам Мицкевич атындағы (Польша) университеттің проректоры

Мамырбаев Оркен Жұмажанович — PhD, ҚР ҒЖБМ Ғылым комитеті ақпараттық және есептеу технологиялары институты ӨМК директорының ғылым жөніндегі орынбасары (Қазақстан)

Бушув Сергей Дмитриевич — техника ғылымдарының докторы, профессор, Украинаның "УКРНЕТ" жобаларды басқару қауымдастығының директоры, Киев ұлттық құрылыс және сулест университеті жобаларды басқару кафедрасының меңгерушісі (Украина)

Белюшицкая Светлана Васильевна — техника ғылымдарының докторы, доцент, Astana IT University есептеу және деректер ғылымы кафедрасының профессоры (Қазақстан)

ЖАУАПТЫ РЕДАКТОР:

Мрзабаева Раушан Жалиевна — магистр, Халықаралық ақпараттық технологиялар университетінің редакторы (Қазақстан)

Халықаралық ақпараттық және коммуникациялық технологиялар журналы

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Префикс DOI: 10.54309

Мерзімділігі: жылына 4 рет.

Басылым тілі: қазақ, орыс, ағылшын.

Тақырып бағыты: "Ақпараттық технологиялар"; "Ақпараттық қауіпсіздік және коммуникациялық технологиялар"; "Әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технология".

Журнал сайты: <https://journal.iitu.edu.kz>

Тарату: материалдар Creative Commons Attribution 4.0 лицензиясы бойынша таратылады

Меншік иесі: АҚ «Халықаралық ақпараттық технологиялар университеті» (Алматы қ.).

Авторлық құқық: © Халықаралық ақпараттық және коммуникациялық технологиялар журналы, 2026

РЕДАКЦИЯ

ГЛАВНЫЙ РЕДАКТОР:

Колесникова Катерина Викторовна — доктор технических наук, профессор, проректор по научно-исследовательской деятельности Международного университета информационных технологий (Казахстан)

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

Ипалакова Мадина Тулегеновна — кандидат технических наук, ассоциированный профессор, директор департамента по научно-исследовательской деятельности Международного университета информационных технологий (Казахстан)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Разак Абдул — PhD, профессор кафедры кибербезопасности Международного университета информационных технологий (Казахстан)

Лучио Томмазо де Паолис — директор отдела исследований и разработок лаборатории AVR департамента инноваций и технологического инжиниринга Университета Саленто (Италия)

Лиз Бэкон — профессор, заместитель вице-канцлера Университета Абертей (Великобритания)

Микеле Пагано — PhD, профессор Университета Пизы (Италия)

Отелбаев Мухтарбай Отелбайулы — доктор физико-математических наук, профессор, академик НАН РК, профессор кафедры математического и компьютерного моделирования Международного университета информационных технологий (Казахстан)

Рысбайулы Болатбек — доктор физико-математических наук, профессор, профессор Astana IT University (Казахстан)

Дайнеко Евгения Александровна — PhD, профессор-исследователь кафедры информационных систем Международного университета информационных технологий (Казахстан)

Дузбаев Нуржан Токкужаевич — PhD, ассоциированный профессор, проректор по цифровизации и инновациям Международного университета информационных технологий (Казахстан)

Синчев Бахтгерей Куспанович — доктор технических наук, профессор, профессор кафедры информационных систем Международного университета информационных технологий (Казахстан)

Сейлова Нургуль Абадуллаевна — кандидат технических наук, декан факультета компьютерных технологий и кибербезопасности Международного университета информационных технологий (Казахстан)

Мухамедиева Ардак Габитовна — кандидат экономических наук, декан факультета бизнеса медиа и управления Международного университета информационных технологий (Казахстан)

Абдикаликова Замира Турсынбаевна — PhD, ассоциированный профессор, заведующая кафедрой математического и компьютерного моделирования Международного университета информационных технологий (Казахстан)

Шильдибеков Ерлан Жаржанович — PhD, ассоциированный профессор, заведующий кафедрой экономики и бизнеса Международного университета информационных технологий (Казахстан)

Дамеуш Максустовна Ескендирова — кандидат технических наук, ассоциированный профессор, заведующая кафедрой кибербезопасности Международного университета информационных технологий (Казахстан)

Ниязгулова Айгуль Аскарбековна — кандидат филологических наук, доцент, профессор, заведующая кафедрой медиакоммуникации и истории Казахстана Международного университета информационных технологий (Казахстан)

Айтмаганбетов Алтай Зуфарович — кандидат технических наук, профессор кафедры радиотехники, электроники и телекоммуникаций Международного университета информационных технологий (Казахстан)

Бахтиярова Елена Ажибековна — кандидат технических наук, ассоциированный профессор, заведующая кафедрой радиотехники, электроники и телекоммуникаций Международного университета информационных технологий (Казахстан)

Канибек Сансызбай — PhD, ассоциированный профессор, профессор-исследователь кафедры кибербезопасности, Международного университета информационных технологий (Казахстан)

Тынымбаев Сахиябай — кандидат технических наук, профессор, профессор-исследователь кафедры компьютерной инженерии, Международного университета информационных технологий (Казахстан)

Алимсиреб Али Абд — PhD, ассоциированный профессор кафедры кибербезопасности Международного университета информационных технологий (Казахстан)

Мохамед Ахмед Хамада — PhD, ассоциированный профессор кафедры информационных систем Международного университета информационных технологий (Казахстан)

Янг Им Чу — PhD, профессор университета Гачон (Южная Корея)

Тадеуш Валлас — PhD, проректор университета имен Адама Мицкевича (Польша)

Мамырбаев Оркен Жумажанович — PhD, заместитель директора по науке РГП Института информационных и вычислительных технологий Комитета науки МНВО РК (Казахстан)

Бушуев Сергей Дмитриевич — доктор технических наук, профессор, директор Украинской ассоциации управления проектами «УКРНЕТ», заведующий кафедрой управления проектами Киевского национального университета строительства и архитектуры (Украина)

Белошицкая Светлана Васильевна — доктор технических наук, доцент, профессор кафедры вычислений и науки о данных Astana IT University (Казахстан)

ОТВЕТСТВЕННЫЙ РЕДАКТОР:

Мрзабаева Раушан Жалиевна — магистр, редактор Международного университета информационных технологий (Казахстан)

Международный журнал информационных и коммуникационных технологий

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Префикс DOI: 10.54309

Периодичность: 4 выпусков в год.

Язык издания: казахский, русский, английский.

Тематическая направленность: "Информационные технологии"; "Информационная безопасность и коммуникационные технологии"; "Цифровые технологии в развитии социо-экономических систем".

Сайт журнала: <https://journal.iitu.edu.kz>

Распространение: материалы распространяются по лицензии Creative Commons Attribution 4.0

Собственник: АО «Международный университет информационных технологий» (г. Алматы).

Авторские права: © Международный журнал информационных и коммуникационных технологий, 2026

CONTENTS

DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF SOCIOECONOMIC SYSTEMS

A.B. Zhalgas, Y.N. Kalpakov, B.Ye. Amirgaliyev MACHINE LEARNING-DRIVEN OPTIMIZATION OF LOGISTICS IN SMART CITIES: A CASE STUDY OF ASTANA	11
L. Kurmangazyieva, Sh. Kodanova, M. Urazgaliyeva, O. Findik, S. Iskakova INTEGRATING FUZZY LOGIC AND ARTIFICIAL INTELLIGENCE IN OPTIMIZING BUSINESS PROCESS AUTOMATION DECISIONS	26
Y. Mailybayev, U. Adilbayeva, R. Amanova ORGANIZATION OF AN ONLINE SURVEY OF PARTICIPANTS IN THE EDUCATIONAL PROCESS AND ANALYSIS OF THE RESULTS BASED ON THE MODIFIED DELPHI METHOD	48
V.A. Takizhanov, A.Z. Ibragimov, A. Shalakhmetov SIMULATION-BASED ROBUSTNESS ASSESSMENT OF ASTANA'S BUS NETWORK UNDER RANDOM AND TARGETED FAILURES	63
G. Hamada, M. Hamada LEVERAGING MACHINE LEARNING MODELS TO BOOST THE MANIPULATION OF UNSTRUCTURED DECISIONS AND UNSTRUCTURED DATA IN DIGITAL BUSINESS FIRMS	78

INFORMATION TECHNOLOGY

M. Zh. Aitimov, G. K. Muratova, Zh. K. Bissenbayeva, I.M. Bapiyev, M. Kassim SEMANTIC COMPLETENESS IN KAZAKH-LANGUAGE EXTRACTIVE QA THROUGH ONTOLOGY AND RETRIEVAL MECHANISMS	96
A.Aitim, A. Auyezova ENCODING NP-COMPLETE PROBLEM INSTANCES INTO NATURAL LANGUAGE FOR LLM PROCESSING	109
O.N. Akylbekov, Y.T. Dauletbek, A.N. Moldagulova, G.S. Zakariya, D.A. Gura MACHINE LEARNING METHODS FOR ANALYSING THREE-DIMENSIONAL SPATIAL DATA IN KAZAKHSTAN'S LAND USE PLANNING	126
S.Zh. Aliaskarov, R.K. Uskenbayeva, A. Razaque, A.B. Kassymova, A.M. Anartayeva TOWARDS EFFICIENT BIG DATA ANALYTICS IN REGIONAL SYSTEMS: PRACTICAL INSIGHTS FROM HYBRID ARCHITECTURE DEPLOYMENT	146
Zh. Bilyalova, N. Shazhdekeyeva, K. Uteuliyeva, T.Akhmurzina, L. Nurgaliyeva A MATHEMATICAL MODEL FOR DECISION-MAKING UNDER UNCERTAINTY HYBRID BASED ON FUZZY LOGIC AND GENERATIVE ARTIFICIAL INTELLIGENCE	165
A. Ismailova, G. Yessenbayeva, K. Kadyrkulov, R. Moldasheva, A. Amangeldi DEVELOPMENT OF A HYBRID DEEP LEARNING MODEL FOR MULTICLASS CLASSIFICATION OF MICROSCOPIC IMAGES OF BACTERIA	186
G. Kalman, J. Kultan, A.N. Ismukamova, N.M. Ausilova, Y.V. Makhatova A DOMAIN-KNOWLEDGE-BASED MODEL FOR REFERENCE RESOLUTION IN LOW-RESOURCE LANGUAGES.....	199
Y. Kamen, Zh. Yessendauletova, L. Fazylova, M. Rakhimzhanova, A.M. Nedzved USING NEURAL NETWORKS FOR OBJECTIVE ASSESSMENT OF ATTENTION IN CHILDREN BASED ON EEG DATA	216
A.Ye. Kulakayeva, Ye.A. Bakhtiyarova, G.T. Jakanova, Sh. Nursultan COMPARATIVE ANALYSIS OF VARIOUS RADIO WAVE PROPAGATION MODELS FOR MOBILE NETWORK COVERAGE PREDICTION	231
A. Kalykulova, K. Bakirov, A. Shoman, A. Nugumanova FORMATION OF A UNIFIED DATA SET FOR ANALYZING THE RELATIONSHIP BETWEEN DOG BREEDS AND DISEASES	247
D. Kantayeva, A. Seisenbekkyzy, S. Biloshchytska COMPARATIVE EVALUATION MODEL BASED ON EXPERT SYSTEMS FOR ABSTRACTIVE SUMMARIZATION OF SCIENTIFIC TEXTS	269
N. Mordvintsev, Zh. Oralbekova, A. Bakiyeva DEEP LEARNING-BASED VEHICLE MAKE AND MODEL RECOGNITION: DATASET CHALLENGES	

AND INFLUENCING FACTORS IN URBAN TRAFFIC	269
M.B. Nurpeissova, Sh.K. Aitkazinova, A.M. Abenov, N.S. Donenbayeva METHODOLOGY FOR TRANSFORMING SATELLITE COORDINATES INTO A TOPOCENTRIC RECTANGULAR COORDINATE SYSTEM	315
L. Naizabayeva, G. Turken, R.A. Druba EFFECTIVE METHOD FOR OPTIMAL CLUSTERING IN UNSUPERVISED LEARNING	328
S. Nurgaliyeva, Zh. Seitkaliyev, A. Sakhipov, M. Kali, R. Hrmo YIELD FORECASTING FOR FRUITS AND VEGETABLES USING RANDOM FOREST ON LARGE- SCALE AGROCLIMATIC DATA	339
A. Nurgalykov, S. Sapakova, D. Zhalmukhanbetov, G. Nurgazy, A. Myrzabayev LANGUAGE ASYMMETRY IN LLMS: THE IMPACT OF NON-ENGLISH DATA ON ACCURACY AND RELIABILITY	355
A. Ospanov, P. Alonso-Jordá, A. Zhumadillayeva BLOCKCHAIN-ENABLED ERP WAREHOUSE INTEGRATION WITH IOT DIMENSIONERS AND MACHINE LEARNING–OPTIMIZED DIMENSIONAL WEIGHT RECONCILIATION	371
A.A. Sakhipov, R.B. Seıitbek EVENT-DRIVEN MICROSERVICES FOR INCIDENT DETECTION AND RESPONSE IN INTELLIGENT TRAFFIC SYSTEM	387
A. Tursynova, B. Omarov EVOLUTION OF THE EFFECTIVENESS OF CONDITIONAL GAN FOR THE SYNTHESIS OF CT IMAGES OF THE BRAIN IN THE DIAGNOSIS OF STROKE	400
G. Yusupova, K.S. Shadinova, D. Ussipbekova, Zh.Zh. Azhibekova, P. Schmidt DETERMINATION OF SOIL PROFILE STRATIFICATION AT 0–200 CM DEPTH USING A MULTILEVEL STACKING MODEL	414

INFORMATION SECURITY AND COMMUNICATION TECHNOLOGIES

S.A. Adilzhanova, M.Zh. Sakypbekova, L.Sh. Cherikbaeva, G.A. Tyulepberdinova, G.T. Zhubanysheva SYSTEMATIC ANALYSIS OF RISK ASSESSMENT METHODS AND MODELS IN INFORMATION SECURITY.....	427
T. K. Zhukabayeva, D.B. Baumuratova, E. Benkhelifa , N.A. Niyetbayeva EDGE COMPUTING-BASED TECHNIQUE FOR CONSTRUCTION OF ATTACK DETECTION MEANS IN CYBER-PHYSICAL SYSTEMS OF INDUSTRIAL INTERNET-OF-THINGS	453
N.E. Karabayev, S.K. Serikbayeva, Y.M. Mardenov, B. Tassuov, M. Fajkus DETECTION OF CYBER ATTACKS IN TRANSPORT NETWORKS BASED ON MACHINE LEARNING METHODS	475
B.A. Kumalakov, A.O. Dargulova A HYBRID FRAMEWORK FOR RESUME-JOB MATCHING SYSTEM	494
V. Makhatova, B. Dzhugembayeva, A. Gabdulova, L. Nurgaliyeva, A. Abdigaliyeva MATHEMATICAL MODEL FOR OPTIMAL SENSOR SELECTION IN SIEM SYSTEMS USING THE ANALYTIC HIERARCHY PROCESS	509

МАЗМУНЫ

ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

А.Б. Жалғас, Е.Н. Калпаков, Б.Е. Амиргалиев АҚЫЛДЫ ҚАЛАЛАРДАҒЫ ЛОГИСТИКАНЫ МАШИНАЛЫҚ ОҚЫТУҒА НЕГІЗДЕЛГЕН ОҢТАЙЛАНДЫРУ: АСТАНАНЫҢ ЖАҒДАЙЫН ЗЕРТТЕУ.....	11
Л.Қурманғазиева, Ш. Қоданова, М. Уразғалиева, О. Findik, С. Искакова ЖАСАНДЫ ИНТЕЛЛЕКТ ПЕН АЙҚЫН ЕМЕС ЛОГИКАНЫ БІРІКТІРУ АРҚЫЛЫ БИЗНЕС- ПРОЦЕСТЕРДІ АВТОМАТТАНДЫРУ ШЕШІМДЕРІН ОҢТАЙЛАНДЫРУ	26
Е. Майлыбаев, У. Адилбаева, Р. Аманова ҰЙЫМДАСТЫРЫЛҒАН ОНЛАЙН САУАЛНАМА АРҚЫЛЫ БІЛІМ БЕРУ ПРОЦЕСІНЕ ҚАТЫСУШЫЛАРДЫҢ ПІКІРЛЕРІН ЖИНАУ ЖӘНЕ НӘТИЖЕЛЕРІН МОДИФИКАЦИЯЛАНҒАН ДЕЛЬФИ ӘДІСІ НЕГІЗІНДЕ ТАЛДАУ	48
В.А. Такижанов, А.Ж. Ибрагимов, А. Шалахметов	

МОДЕЛЬДЕУ НЕГІЗІНДЕ АСТАНАНЫҢ АВТОБУС ЖЕЛІСІНІҢ ТҮРАҚТЫЛЫҒЫН БАҒАЛАУ: КЕЗДЕЙСОҚ ЖӘНЕ МАҚСАТТЫ ІСТЕН ШЫҒУЛАР ЖАҒДАЙЫНДА	63
Г. Хамада, М. Хамада ЦИФРЛЫҚ БИЗНЕС-ҰЙЫМДАРЫНДА ҚҰРЫЛЫМДАЛМАҒАН ШЕШІМДЕР МЕН ҚҰРЫЛЫМДАЛ- МАҒАН ДЕРЕКТЕРМЕН ЖҰМЫС ІСТЕУ ТИІМДІЛІГІН АРТТЫРУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ МОДЕЛЬДЕРІН ПАЙДАЛАҢУ	78

АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

М.Ж. Айтимов, Г.К. Муратова, Ж.К. Бисенбаева, И.М. Бапиев, М. Кассим ОНТОЛОГИЯ ЖӘНЕ ІЗДЕУ МЕХАНИЗМДЕРІ АРҚЫЛЫ ҚАЗАҚ ТІЛІНДЕГІ ЭКСТРАКЦИЯЛЫҚ ҚА-ДАҒЫ СЕМАНТИКАЛЫҚ ТОЛЫҚТЫҚ	96
Ә. Әйтгім, Ә. Әуезова LLM ӨНДЕУІ ҮШІН NP-ТОЛЫҚ ЕСЕПТЕРДІҢ ЭКЗЕМПЛЯРЛАРЫН ТАБИҒИ ТІЛДЕ КОДТАУ	109
О.Н. Ақылбеков, Е.Т. Даулетбек, А.Н. Молдагулова, Г.С. Закария, Д.А. Гура ҚАЗАҚСТАННЫҢ АУМАҚТЫҚ ЖОСПАРЛАУЫНДАҒЫ ҮШ ӨЛШЕМДІ КЕҢІСТІКТІК МӨЛІМЕТТЕРДІ ТАЛДАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІ	126
С.Ж. Алиаскаров, Р.К. Ускенбаева, А. Разак, А.Б. Касымов, А.М. Анартаева АЙМАҚТЫҚ ЖҮЙЕЛЕРДЕГІ ҮЛКЕН ДЕРЕКТЕРДІ ТИІМДІ ТАЛДАУҒА ҚАРАЙ: ГИБРИДТІ АРХИТЕКТУРАНЫ ЕНГІЗУДІҢ ПРАКТИКАЛЫҚ ТҮСІНІКТЕР.....	146
Ж. Билялова, Н. Шаждекеева, К. Өтеулиева, Т. Ахмурзина, Л. Нұрғалиева АЙҚЫН ЕМЕС ЛОГИКА ЖӘНЕ ГЕНЕРАТИВТІ ЖАСАНДЫ ИНТЕЛЛЕКТ НЕГІЗІНДЕ БЕЛГІСІЗДІК ЖАҒДАЙЫНДА ШЕШІМ ҚАБЫЛДАУДЫҢ ГИБРИДТІ МАТЕМАТИКАЛЫҚ МОДЕЛІ	165
А.А. Исмаилова, Г.Р. Есенбаева, К.К. Кадиркулов, Р.Н. Молдашева, А. Амангелді РОСКОПИЯЛЫҚ БЕЙНЕЛЕРІН КӨПКЛАССТЫ ЖІКТЕУГЕ АРНАЛҒАН ГИБРИДТІ ТЕРЕҢ ОҚЫТУ МОДЕЛІН ӨЗІРЛЕУ	186
Г. Қалман, К. Ярослав, А.Н. Исмуханова, Н.М. Аусилова, В.Е. Махатова ПӘНДІК САЛА БІЛІМ НЕГІЗІНДЕ РЕУСРСТАРЫ АЗ ТІЛДЕРДЕГІ РЕФЕРЕНЦИЯНЫ ШЕШУДІҢ МОДЕЛІ.....	199
Е.Г. Кәмен, Ж.Т. Есендаулетова, Л.С. Фазылова, М.Б. Рахимжанова, А.М. Недзьведь ЭЭГ ДЕРЕКТЕРІ БОЙЫНША БАЛАЛАРДЫҢ ЗЕЙІНІН ОБЪЕКТИВТІ БАҒАЛАУ ҮШІН НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ҚОЛДАНУ	216
А.Е. Кулакаева, Е.А. Бахтиярова, Г.Т. Джаканова, Ш. Нурсултан ҰЯЛЫ БАЙЛАНЫС ЖЕЛІЛЕРІНІҢ ҚАМТУ АЙМАҒЫН БОЛЖАУҒА АРНАЛҒАН ӨРТҮРЛІ РАДИОТӨЛҚЫН ТАРАЛУ МОДЕЛЬДЕРІНІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ	231
А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова ИТ ТҰҚЫМДАРЫ МЕН АУРУЛАРЫ АРАСЫНДАҒЫ БАЙЛАНЫСТЫ ТАЛДАУҒА АРНАЛҒАН БІРЫҢҒАЙ ДЕРЕКТЕР ЖИЙНТЫҒЫН ҚАЛЫПТАСТЫРУ	247
Д. Кантаева, А. Сейсенбекқызы, С. Билощницкая ҒЫЛЫМИ МӘТІНДЕРДІ АБСТРАКТИВТІ РЕФЕРАТТАУҒА АРНАЛҒАН ЭКСПЕРТТІК ЖҮЙЕЛЕРГЕ НЕГІЗДЕЛГЕН САЛЫСТЫРМАЛЫ БАҒАЛАУ МОДЕЛІ	269
Н. Мордвинцев, Ж. Оралбекова, А. Бакиева ТЕРЕҢ ОҚЫТУҒА НЕГІЗДЕЛГЕН КӨЛІК ҚҰРАЛДАРЫНЫҢ МАРКАСЫ МЕН МОДЕЛІН ТАҢУ: ҚА- ЛАЛЫҚ КӨЛІК ҚОЗҒАЛЫСЫНДАҒЫ ДЕРЕКТЕР ЖИЙНТЫҒЫНЫҢ МӘСЕЛЕЛЕРІ ЖӘНЕ ЫҚПАЛ ЕТУШІ ФАКТОРЛАР	285
М.Б. Нұрпейісова, Ш.Қ. Айтқазина, А.М. Абенев, Н.С. Дөненбаева СПУТНИКТІК КООРДИНАТТАРДЫ ТОПОЦЕНТРЛІК ТІК БҰРЫШТЫ КООРДИНАТТАР ЖҮЙЕСІНЕ ТҮРЛЕНДІРУДІҢ ӘДІСТЕМЕСІ	315
Л. Найзабаева, Г. Түркен, Р.А. Друба БАҚЫЛАНБАЙТЫН ОҚЫТУДАҒЫ ОҢТАЙЛЫ КЛАСТЕРЛЕУДІҢ ТИІМДІ ӘДІСІ	328
С. Нұрғалиева, Ж. Сенткалиев, А. Сахипов, М. Кали, Р. Грмо ҮЛКЕН КӨЛЕМДЕГІ АГРОКЛИМАТТЫҚ ДЕРЕКТЕР НЕГІЗІНДЕ RANDOM FOREST ӘДІСІН ПАЙДАЛАНА ОТЫРЫП, ЖЕМІС-КӨКӨНІС ДАҚЫЛДАРЫНЫҢ ӨНІМДІЛІГІН БОЛЖАУ	339
А. Нұрғалықов, С. Сапакова, Д. Жалмұханбетов, Г. Нұрғазы, А. Мырзабаев ҮЛКЕН ТІЛДІК МОДЕЛЬДЕРДЕГІ ТІЛДІК АСИММЕТРИЯ: АҒЫЛШЫН ТІЛІНЕН БАСҚА ДЕРЕКТЕРДІҢ ДӘЛДІК ПЕН СЕНІМДІЛІККЕ ӘСЕРІ	355
А. Оспанов, П. Алонсо-Хорда, А. Жұмаділлаева	

БЛОКЧЕЙН-ТЕХНОЛОГИЯСЫМЕН ЫҚПАЛДАС ERP ҚОЙМА ЖҮЙЕСІН ІОТ ДИМЕНСИОНЕРЛЕР ЖӘНЕ МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ ОПТИМИЗАЦИЯЛАНҒАН ӨЛШЕМДІ САЛМАҚ ЕСЕПТЕУМЕН ИНТЕГРАЦИЯЛАУ	371
А.А. Сахипов, Р.Б. Сейітбек ОҚИҒАҒА БАҒДАРЛАНҒАН МИКРОҚЫЗМЕТТЕР ЖҮЙЕСІ АРҚЫЛЫ АҚЫЛДЫ ТРАФИК ЖҮЙЕЛЕРІНДЕ ОҚИҒАЛАРДЫ АНЫҚТАУ ЖӘНЕ ШАРАЛАР ҚОЛДАНУ	387
А. Турсынова, Б. Омаров ИНСУЛЬТТИ ДИАГНОСТИКАЛАУ КЕЗІНДЕ МИ КТ-БЕЙНЕЛЕРІН СИНТЕЗДЕУГЕ АРНАЛҒАН ШАРТТЫ ҒАН ТИІМДІЛІГІНІҢ ЭВОЛЮЦИЯСЫ	400
Г.М. Юсупова, К.С. Шадинова, Д.И. Усипбекова, Ж.Ж. Ажибекова, Р. Schmidt ТОПЫРАҚ ПРОФИЛІНІҢ 0–200 СМ ТЕРЕНДІКТЕГІ СТРАТИФИКАЦИЯСЫН КӨПДЕНГЕЙЛІ СТЕКИНГ-МОДЕЛІМЕН АНЫҚТАУ.....	414

АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН

С.А. Адилжанова, М.Ж. Сақыпбекова, Л.Ш. Черикбаева, Г.А. Түлепбердинова, Г.Т. Жубанышева АҚПАРАТТЫҚ ҚАУІПСІЗДІКТЕ ТӘУЕКЕЛДЕРДІ БАҒАЛАУ ӘДІСТЕРІ МЕН МОДЕЛЬДЕРІН ЖҮЙЕЛІ ТАЛДАУ.....	427
Т.К. Жукабаева, Д. Б. Баумуратова, Е. Бенкхелифа, Н.А. Ниетбаева ШЕКАРАЛЫҚ ЕСЕПТЕУЛЕРДІ ҚОЛДАНА ОТЫРЫП, ЗАТТАРДЫҢ ӨНЕРКӘСІПТІК ИНТЕРНЕТІНІҢ КИБЕРФИЗИКАЛЫҚ ЖҮЙЕЛЕРІНДЕГІ ШАБУЫЛДАРДЫ АНЫҚТАУ ҚҰРАЛДАРЫН ҚҰРУ ӘДІСТЕМЕСІ.....	453
Н.Е. Қарабаев, С.К. Серикбаева, Е.М. Марденов, Б. Тасуов, М. Файқус МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІНЕ НЕГІЗДЕЛГЕН КӨЛІК ЖЕЛІЛЕРІНДЕГІ КИБЕРШАБУЫЛДАРДЫ АНЫҚТАУ	475
Б.А. Кумалаков, А.О. Даргулова ТҮЙІНДЕМЕЛЕР МЕН ВАКАНСИЯЛАРДЫ АВТОМАТТАНДЫРЫЛҒАН СӘЙКЕСТЕНДІРУГЕ НЕГІЗДЕЛГЕН ГИБРИДТІ ҮМІТКЕРЛЕРДІ ІРІКТЕУ ЖҮЙЕСІ	494
В. Махатова, Б. Джугембаева, А. Габдулова, Л. Нургалиева, А. Абдигалиева ИЕРАРХИЯЛАРДЫ ТАЛДАУ ӘДІСІ НЕГІЗІНДЕ SIEM ЖҮЙЕЛЕРІНДЕ ОҢТАЙЛЫ СЕНСОРДЫ ТАҢДАУДЫҢ МАТЕМАТИКАЛЫҚ МОДЕЛІ	509

СОДЕРЖАНИЕ ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИ СОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ

А.Б. Жалғас, Е.Н. Калпаков, Б.Е. Амиргалиев ОПТИМИЗАЦИЯ ЛОГИСТИКИ В УМНЫХ ГОРОДАХ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ: НА ПРИМЕРЕ АСТАНЫ	11
Л. Курмангазиева, Ш. Коданова, М. Уразғалиева, О. Финдик, С. Исакова ИНТЕГРАЦИЯ НЕЧЕТКОЙ ЛОГИКИ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ ОПТИМИЗАЦИИ РЕШЕНИЙ ПО АВТОМАТИЗАЦИИ БИЗНЕС-ПРОЦЕССОВ	26
Е. Майлыбаев, У. Адилбаева, Р. Аманова СБОР МНЕНИЙ УЧАСТНИКОВ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПОСРЕДСТВОМ ОРГАНИЗОВАННОГО ОНЛАЙН-АНКЕТИРОВАНИЯ И АНАЛИЗ РЕЗУЛЬТАТОВ НА ОСНОВЕ МОДИФИЦИРОВАННОГО МЕТОДА ДЕЛЬФИ	48
В.А. Такижанов, А.Ж. Ибрагимов, А. Шалахметов ОЦЕНКА УСТОЙЧИВОСТИ АВТОБУСНОЙ СЕТИ АСТАНЫ НА ОСНОВЕ МОДЕЛИРОВАНИЯ ПРИ СЛУЧАЙНЫХ И ЦЕЛЕНАПРАВЛЕННЫХ ОТКАЗАХ	63
Г. Хамада, М. Хамада ИСПОЛЬЗОВАНИЕ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ РАБОТЫ С НЕСТРУКТУРИРОВАННЫМИ РЕШЕНИЯМИ И НЕСТРУКТУРИРОВАННЫМИ ДАННЫМИ В ЦИФРОВЫХ БИЗНЕС-ОРГАНИЗАЦИЯХ	78

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

М.Ж. Айтимов, Г.К. Муратова, Ж.К. Бисенбаева, И.М. Бапиев, М. Кассим

СЕМАНТИЧЕСКАЯ ПОЛНОТА В КАЗАХСКОЯЗЫЧНОМ EXTRACTIVE QA ЧЕРЕЗ ОНТОЛОГИЮ И RETRIEVAL-МЕХАНИЗМЫ	96
А. Айтим, А. Ауезова КОДИРОВАНИЕ ЭКЗЕМПЛЯРОВ NP-ПОЛНЫХ ЗАДАЧ В ЕСТЕСТВЕННОМ ЯЗЫКЕ ДЛЯ ОБРАБОТКИ LLM	109
О.Н. Акылбеков, Е.Т. Даулетбек, А.Н. Молдагулова, Г.С. Закария, Д.А. Гура МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА ТРЁХМЕРНЫХ ПРОСТРАНСТВЕННЫХ ДАННЫХ В ТЕРРИТОРИАЛЬНОМ ПЛАНИРОВАНИИ КАЗАХСТАНА	126
С.Ж. Алиаскаров, Р.К. Ускенбаева, А. Разак, А.Б. Касымова, А.М. Анартаева НА ПУТИ К ЭФФЕКТИВНОЙ АНАЛИТИКЕ БОЛЬШИХ ДАННЫХ В РЕГИОНАЛЬНЫХ СИСТЕМАХ: ПРАКТИЧЕСКИЕ ВЫВОДЫ ИЗ ВНЕДРЕНИЯ ГИБРИДНОЙ АРХИТЕКТУРЫ	146
Ж. Билялова, Н. Шаждекеева, К. Утеулиева, Т. Ахмурзина, Л. Нурғалиева ГИБРИДНАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ НА ОСНОВЕ НЕЧЕТКОЙ ЛОГИКИ И ГЕНЕРАТИВНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	165
А.А. Исмаилова, Г.Р. Есенбаева, К.К. Кадиркулов, Р.Н. Молдашева, А. Амангелды РАЗРАБОТКА ГИБРИДНОЙ МОДЕЛИ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ МНОГОКЛАССОВОЙ КЛАССИФИКАЦИИ МИКРОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ БАКТЕРИЙ	186
Г. Калман, К. Ярослав, А.Н. Исмуханова, Н.М. Аусилова, В.Е. Махатова МОДЕЛЬ НА ОСНОВЕ ЗНАНИЙ ПРЕДМЕТНОЙ ОБЛАСТИ ДЛЯ РАЗРЕШЕНИЯ КОРЕФЕРЕНЦИИ В МАЛОРЕСУРСНЫХ ЯЗЫКАХ	199
Е.Г. Камен, Ж.Т. Есендаулетова, Л.С. Фазылова, М.Б. Рахимжанова, А.М. Недзведь ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБЪЕКТИВНОЙ ОЦЕНКИ ВНИМАНИЯ У ДЕТЕЙ ПО ДАННЫМ ЭЭГ	216
А.Е. Кулакаева, Е.А. Бахтиярова, Г.Т. Джаканова, Ш. Нурсултан СРАВНИТЕЛЬНЫЙ АНАЛИЗ РАЗЛИЧНЫХ МОДЕЛЕЙ РАСПРОСТРАНЕНИЯ РАДИОВОЛН ДЛЯ ПРОГНОЗИРОВАНИЯ ПОКРЫТИЯ СЕТЕЙ МОБИЛЬНОЙ СВЯЗИ	231
А. Калыкулова, К. Бакиров, А. Шоман, А. Нугуманова ФОРМИРОВАНИЕ ЕДИНОГО НАБОРА ДАННЫХ ДЛЯ АНАЛИЗА ВЗАИМОСВЯЗИ МЕЖДУ ПОРОДАМИ СОБАК И ЗАБОЛЕВАНИЯМИ	247
Д. Кантаева, А. Сейсенбеккызы, С. Билощицкая СРАВНИТЕЛЬНАЯ МОДЕЛЬ ОЦЕНКИ НА ОСНОВЕ ЭКСПЕРТНЫХ СИСТЕМ ДЛЯ АБСТРАКТИВНОГО РЕФЕРИРОВАНИЯ НАУЧНЫХ ТЕКСТОВ	269
Н. Мордвинцев, Ж. Оралбекова, А. Бакиева РАСПОЗНАВАНИЕ МАРКИ И МОДЕЛИ ТРАНСПОРТНЫХ СРЕДСТВ НА ОСНОВЕ ГЛУБОКОГО ОБУЧЕНИЯ: ПРОБЛЕМЫ НАБОРОВ ДАННЫХ И ФАКТОРЫ, ВЛИЯЮЩИЕ НА РАБОТУ СИСТЕМЫ В УСЛОВИЯХ ГОРОДСКОГО ДВИЖЕНИЯ	285
М.Б. Нурпеисова, Ш.К. Айтказинова, А.М. Абенов, Н.С. Доненбаева МЕТОДИКА ПРЕОБРАЗОВАНИЯ СПУТНИКОВЫХ КООРДИНАТ В ТОПОЦЕНТРИЧЕСКУЮ ПРЯМОУГОЛЬНУЮ СИСТЕМУ КООРДИНАТ	315
Л. Найзабаева, Г. Туркен, Р.А. Друба ЭФФЕКТИВНЫЙ МЕТОД ОПТИМАЛЬНОЙ КЛАСТЕРИЗАЦИИ В ОБУЧЕНИИ БЕЗ УЧИТЕЛЯ	328
С. Нурғалиева, Ж. Сенткалиев, А. Сахипов, М. Кали, Р. Грмо ПРОГНОЗИРОВАНИЕ УРОЖАЙНОСТИ ФРУКТОВ И ОВОЩЕЙ С ИСПОЛЬЗОВАНИЕМ RANDOM FOREST НА ОСНОВЕ КРУПНОМАСШТАБНЫХ АГРОКЛИМАТИЧЕСКИХ ДАННЫХ	339
А. Нурғалыков, С. Сапакова, Д. Жалмуханбетов, Г. Нургазы, А. Мырзабаев ЯЗЫКОВАЯ АСИММЕТРИЯ В БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЯХ: ВЛИЯНИЕ ДАННЫХ НА НЕАНГЛИЙСКИХ ЯЗЫКАХ НА ТОЧНОСТЬ И НАДЕЖНОСТЬ	355
А. Оспанов, П. Алонсо-Хорда, А. Жумадилаева ИНТЕГРАЦИЯ СКЛАДСКИХ МОДУЛЕЙ ERP-СИСТЕМ С ИСПОЛЬЗОВАНИЕМ БЛОКЧЕЙНА, ИОТ-ДИМЕНСИОНЕРОВ И ОПТИМИЗИРОВАННОГО МАШИНЫМ ОБУЧЕНИЕМ РАСЧЁТА ГАБАРИТНОГО ВЕСА	371
А.А. Сахипов, Р.Б. Сейитбек СОБЫТИЯ-ОРИЕНТИРОВАННЫЕ МИКРОСЕРВИСЫ ДЛЯ ОБНАРУЖЕНИЯ И РЕАГИРОВАНИЯ НА ИНЦИДЕНТЫ В ИНТЕЛЛЕКТУАЛЬНЫХ ТРАНСПОРТНЫХ СИСТЕМАХ	387
А. Тұрсынова, Б. Омаров ЭВОЛЮЦИЯ ЭФФЕКТИВНОСТИ УСЛОВНОГО GAN ДЛЯ СИНТЕЗА КТ-ИЗОБРАЖЕНИЙ ГОЛОВНО-	

ГО МОЗГА ПРИ ДИАГНОСТИКЕ ИНСУЛЬТА	400
Г.М. Юсупова, К.С. Шадинова, Д.И. Усипбекова, Ж.Ж. Ажибекова, П. Шмидт	
ОПРЕДЕЛЕНИЕ СТРАТИФИКАЦИИ ПОЧВЕННОГО ПРОФИЛЯ НА ГЛУБИНЕ 0–200 СМ С	
ИСПОЛЬЗОВАНИЕМ МОДЕЛИ МНОГОУРОВНЕВОГО НАЛОЖЕНИЯ	414

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

С.А. Адилжанова, М.Ж. Сакыпбекова, Л.Ш. Черикбаева, Г.А. Тюлепбердинова, Г.Т. Жубанышева	
СИСТЕМАТИЧЕСКИЙ АНАЛИЗ МЕТОДОВ И МОДЕЛЕЙ ОЦЕНКИ РИСКОВ ИНФОРМАЦИОННОЙ	
БЕЗОПАСНОСТИ	427
Т.К. Жукабаева, Д.Б. Баумуратова, Е. Бенкхелифа, Н.А. Ниембаева	
МЕТОДИКА ПОСТРОЕНИЯ СРЕДСТВ ОБНАРУЖЕНИЯ АТАК В КИБЕРФИЗИЧЕСКИХ СИСТЕМАХ	
ПРОМЫШЛЕННОГО ИНТЕРНЕТА ВЕЩЕЙ С ИСПОЛЬЗОВАНИЕМ ГРАНИЧНЫХ	
ВЫЧИСЛЕНИЙ	453
Н.Е. Карабаев, С.К. Серикбаева, Е.М. Марденов, Б. Тасуов, М. Файкус	
ОБНАРУЖЕНИЕ КИБЕРАТАК В ТРАНСПОРТНЫХ СЕТЯХ НА ОСНОВЕ МЕТОДОВ МАШИННОГО	
ОБУЧЕНИЯ	475
Б.А. Кумалаков, А.О. Даргулова	
ГИБРИДНЫЙ ПОДХОД К АВТОМАТИЗИРОВАННОМУ ПОДБОРУ КАНДИДАТОВ НА ОСНОВЕ	
СОПОСТАВЛЕНИЯ РЕЗЮМЕ И ВАКАНСИЙ	494
В. Махатова, Б. Джугембаева, А. Габдулова, Л. Нургалиева, А. Абдигалиева	
МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ВЫБОРА ОПТИМАЛЬНОГО СЕНСОРА В SIEM-СИСТЕМАХ	
СРЕДСТВАМИ МЕТОДА АНАЛИЗА ИЕРАРХИЙ	509

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Vol. 7. Is.1. Number 25 (2026). Pp. 339–354

Journal homepage: <https://journal.iitu.edu.kz><https://doi.org/10.54309/IJICT.2026.25.1.020>

YIELD FORECASTING FOR FRUITS AND VEGETABLES USING RANDOM FOREST ON LARGE-SCALE AGROCLIMATIC DATA

S. Nurgaliyeva¹, Zh. Seitkaliyev^{1*}, A. Sakhipov¹, M. Kali¹, R. Hrmo²

¹Astana IT University, Astana, Kazakhstan;

²Dubnica Institute of Technology in Dubnica nad Vahom, Dubnica nad Vahom, Slovakia.

E-mail: 242753@astanait.edu.kz

Symbat Nurgaliyeva — PhD, assistant professor of School of Software Engineering, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0009-0009-5880-6166>;

Zhandos Seitkaliyev — 2-year Master's student, Astana IT University, Astana, Kazakhstan

E-mail: 242753@astanait.edu.kz, <https://orcid.org/0009-0000-3427-9714>;

Aivar Sakhipov — PhD in Computer Science, Assistant Professor of the School of «Computer Engineering», Astana IT University, Astana, Kazakhstan

<https://orcid.org/0000-0003-1045-4199>;

Marzhan Kali — master of science in Engineering, senior-lecturer of School of Cybersecurity, Astana IT University, Astana, Kazakhstan

<https://orcid.org/0000-0001-5588-8164>;

Roman Hrmo — Dubnica Institute of Technology in Dubnica nad Vahom, Dubnica nad Vahom, Slovakia

<https://orcid.org/0000-0003-2458-2706>.

© S. Nurgaliyeva, Zh. Seitkaliyev, A. Sakhipov, M. Kali, R. Hrmo

Abstract. Accurate yield forecasting is essential for sustainable agriculture and climate adaptation, especially in regions exposed to extreme weather conditions such as Central Asia. This study investigates the use of machine learning models, such as Linear Regression (LR), Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB), for predicting fruit and vegetable yields based on large-scale agroclimatic data. After preprocessing and feature encoding, all models achieved high accuracy ($R^2 > 0.90$), with LR and GB slightly outperforming RF by less than 1%. Despite this small difference, RF demonstrated greater robustness and interpretability, remaining stable under data variability and missing values. Feature-importance analysis identified rainfall as the dominant predictor, followed by fertilizer use and irrigation. Temperature exhibited a nonlinear effect on yield, with productivity plateauing above 32–34 °C, highlighting the

risk of heat stress in arid climates. These findings confirm that ensemble-based methods, particularly Random Forest, provide a practical balance between accuracy, stability, and explainability for climate-smart yield forecasting in real agricultural environments.

Keywords: Yield forecasting, Random Forest, agroclimatic data, robustness, Central Asia, nonlinear effects

For citation: Nurgaliyeva S., Seitkaliyev Zh., Sakhipov A., Kali M., Hrmo R. (2026). Yield forecasting for fruits and vegetables using random forest on large-scale agroclimatic data // International journal of information and communication technologies. Vol. 7. No. 25. Pp. 339-354. <https://doi.org/10.54309/IJICT.2026.25.1.020>. (In Eng.).

Conflict of interest: The authors declare that there is no conflict of interest.

ҮЛКЕН КӨЛЕМДЕГІ АГРОКЛИМАТТЫҚ ДЕРЕКТЕР НЕГІЗІНДЕ RANDOM FOREST ӘДІСІН ПАЙДАЛАНА ОТЫРЫП, ЖЕМІС-КӨКӨНІС ДАҚЫЛДАРЫНЫҢ ӨНІМДІЛІГІН БОЛЖАУ

С. Нургалиева¹, Ж. Сеиткалиев¹, А. Сахипов¹, М. Кали¹, Р. Грмо^{2*}

¹Astana IT University, Астана, Қазақстан;

²Dubnica Institute of Technology in Dubnica nad Vahom, Дубница-над-Вагом, Словакия.

E-mail: 242753@astanait.edu.kz

Нургалиева Сымбат — PhD, Бағдарламалық инженерия мектебінің қауымдастырылған профессоры, Astana IT University, Астана, Қазақстан
<https://orcid.org/0009-0009-5880-6166>;

Сеиткалиев Жандос — магистратураның 2-курс студенті, Astana IT University, Астана, Қазақстан

E-mail: 242753@astanait.edu.kz, <https://orcid.org/0009-0000-3427-9714>;

Сахипов Айвар — компьютерлік ғылымдар саласындағы PhD, Компьютерлік инженерия мектебінің қауымдастырылған профессоры, Astana IT University, Астана, Қазақстан

<https://orcid.org/0000-0003-1045-4199>;

Кали Маржан — инженерия ғылымдарының магистрі, Киберқауіпсіздік мектебінің аға оқытушысы, Astana IT University, Астана, Қазақстан

<https://orcid.org/0000-0001-5588-8164>;

Грмо Роман — Dubnica Institute of Technology in Dubnica nad Vahom, Дубница-над-Вагом, Словакия

<https://orcid.org/0000-0003-2458-2706>.

© С. Нургалиева, Ж. Сеиткалиев, А. Сахипов, М. Кали, Р. Грмо

Аннотация. Өнімділікті дәл болжау ауыл шаруашылығын тұрақты дамыту және климаттың өзгеруіне бейімделу үшін аса маңызды, әсіресе Орталық

Азия сияқты экстремалды ауа райы жағдайларына бейім өңірлерде. Бұл зерттеуде ауқымды агроклиматтық деректер негізінде жемістер мен көкөністердің өнімділігін болжау үшін сызықтық регрессия (LR), шешімдер ағашы (DT), Random Forest (RF) және градиенттік бустинг (GB) сияқты машиналық оқыту модельдерін қолдану қарастырылады. Деректерді алдын ала өңдеу және белгілерді кодтаудан кейін барлық модельдер жоғары дәлдік көрсетті ($R^2 > 0,90$). Бұл ретте LR және GB модельдерінің нәтижелері RF моделінен 1 %-дан аз ғана көрсеткішке жоғары болды. Нәтижелер арасындағы айырмашылықтың аз болғанына қарамастан, RF моделі жоғары тұрақтылық пен интерпретациялану мүмкіндігін көрсетті және деректердің өзгермелілігі мен бос мәндердің болуына қарамастан тұрақты нәтижелерді сақтады. Белгілердің маңыздылығын талдау жауын-шашын мөлшерінің негізгі болжамдық фактор екенін көрсетті. Одан кейін тыңайтқыштарды қолдану және суару көрсеткіштері маңызды факторлар ретінде анықталды. Температура өнімділікке сызықтық емес әсер етеді: температура 32–34 °C-тан жоғары болған кезде өнімділік тұрақтанады. Бұл құрғақ климат жағдайында жылу стрессінің туындау қаупін көрсетеді. Алынған нәтижелер ансамбльдік әдістердің, әсіресе Random Forest моделінің, нақты ауыл шаруашылығы жағдайларында климаттық факторларды ескере отырып өнімділікті болжау кезінде дәлдік, тұрақтылық және түсіндірмелілік арасында практикалық тепе-теңдікті қамтамасыз ететінін көрсетеді.

Түйінді сөздер: өнімділікті болжау, Random Forest, агроклиматтық деректер, тұрақтылық, Орталық Азия, сызықтық емес әсерлер

Дәйексөздер үшін: Нургалиева С., Сеиткалиев Ж., Сахипов А., Кали М., Грмо Р. (2026). Үлкен көлемдегі агроклиматтық деректер негізінде random forest әдісін пайдалана отырып, жеміс-көкөніс дақылдарының өнімділігін болжау // Халықаралық ақпараттық және коммуникациялық технологиялар журналы. Т. 7. No 25. Б. 339–354. <https://doi.org/10.54309/IJICT.2026.25.1.020>. (Ағыл. тіл.).

Мүдделер қақтығысы: Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдейді.

ПРОГНОЗИРОВАНИЕ УРОЖАЙНОСТИ ФРУКТОВ И ОВОЩЕЙ С ИСПОЛЬЗОВАНИЕМ RANDOM FOREST НА ОСНОВЕ КРУПНОМАСШТАБНЫХ АГРОКЛИМАТИЧЕСКИХ ДАННЫХ

С. Нургалиева¹, Ж. Сеиткалиев¹, А. Сахипов¹, М. Кали¹, Р. Грмо^{2*}

¹Astana IT University, Астана, Казахстан;

²Dubnica Institute of Technology in Dubnica nad Vahom, Дубница-над-Вагом, Словакия.

E-mail: 242753@astanait.edu.kz

Нургалиева Сымбат — PhD, ассистент профессор школы Программной инженерии, Astana IT University, Астана, Казахстан
<https://orcid.org/0009-0009-5880-6166>;

Сеиткалиев Жандос — магистрант 2 курса, Astana IT University, Астана, Казахстан

E-mail: 242753@astanait.edu.kz, <https://orcid.org/0009-0000-3427-9714>;

Сахипов Айвар — PhD в области компьютерных наук, ассоциированный профессор Школы компьютерной инженерии, Astana IT University, Астана, Казахстан
<https://orcid.org/0000-0003-1045-4199>;

Кали Маржан — магистр технических наук, сеньор-лектор школы Кибербезопасности, Астана, Казахстан
<https://orcid.org/0000-0001-5588-8164>;

Грмо Роман — Dubnica Institute of Technology in Dubnica nad Vahom, Дубница-над-Вагом, Словакия
<https://orcid.org/0000-0003-2458-2706>.

© С. Нургалиева, Ж. Сеиткалиев, А. Сахипов, М. Кали, Р. Грмо

Аннотация. Точное прогнозирование урожайности имеет важное значение для обеспечения устойчивого развития сельского хозяйства и адаптации к климатическим изменениям, особенно в регионах, подверженных экстремальным погодным условиям, таких как Центральная Азия. В настоящем исследовании рассматривается применение моделей машинного обучения, таких как линейная регрессия (LR), дерево решений (DT), Random Forest (RF) и градиентный бустинг (GB), для прогнозирования урожайности фруктов и овощей на основе крупномасштабных агроклиматических данных. После предварительной обработки данных и кодирования признаков все модели продемонстрировали высокую точность ($R^2 > 0,90$), при этом LR и GB незначительно превосходили RF менее чем на 1 %. Несмотря на небольшую разницу в показателях, модель RF продемонстрировала более высокую устойчивость и интерпретируемость, сохраняя стабильность при изменчивости данных и наличии пропущенных значений. Анализ важности признаков показал, что количество осадков является наиболее значимым предиктором, за которым следуют использование удобрений и орошение. Температура оказывает нелинейное влияние на урожайность: при температуре выше 32–34 °C производительность стабилизируется, что указывает на риск теплового стресса в засушливом климате. Полученные результаты подтверждают, что ансамблевые методы, в частности Random Forest, обеспечивают практический баланс между точностью, стабильностью и объяснимостью при прогнозировании урожайности с учетом климатических факторов в реальных условиях сельскохозяйственного производства.

Ключевые слова: прогнозирование урожайности, Random Forest, агроклиматические

Для цитирования: Нургалиева С., Сеиткалиев Ж., Сахипов А., Кали М., Грмо Р. (2026). Прогнозирование урожайности фруктов и овощей с использованием random forest на основе крупномасштабных агроклиматических данных //

Международный журнал информационных и коммуникационных технологий. Т. 7. No. 25. Стр. 339-354. <https://doi.org/10.54309/IJICT.2026.25.1.020>. (На англ.).

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Introduction.

Agricultural production is increasingly affected by climate variability, especially in arid and semi-arid regions such as Central Asia. Irregular rainfall, rising temperatures, and frequent droughts reduce crop stability and complicate long-term planning for irrigation and fertilizer management. Under these conditions, reliable yield forecasting becomes a key component of climate-smart agriculture and food security strategies.

Traditional statistical models often fail to capture the nonlinear and multivariate nature of agroclimatic systems. In contrast, machine learning (ML) methods can process large heterogeneous datasets and identify complex relationships between environmental and agronomic variables. Recent studies have successfully applied ML algorithms such as Random Forest (RF), Gradient Boosting (GB), and neural networks to predict yields of cereals, fruits, and vegetables using weather, soil, and management features. However, most of these works emphasize numerical accuracy while overlooking model robustness, interpretability, and practical deployment in real-world agricultural conditions.

This study addresses these gaps by evaluating and comparing four widely used algorithms – Linear Regression (LR), Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) on a large-scale agroclimatic dataset of fruit and vegetable production. The main objectives are:

1. to identify the most influential environmental factors affecting yield;
2. to quantify how temperature impacts productivity under varying climatic conditions;
3. to assess which model provides the best balance between accuracy, stability, and interpretability.

The study hypothesizes that ensemble-based models, particularly Random Forest, can deliver more robust and interpretable yield predictions than single-model approaches, effectively capturing nonlinear temperature effects that are critical for forecasting in climate-stressed regions.

In addition to its methodological contribution, this research also supports agricultural digitalization and sustainability education initiatives by demonstrating how interpretable AI can enhance understanding of complex agroclimatic systems. The proposed framework illustrates how data-driven approaches can assist in decision-making for precision farming, encouraging the integration of machine learning into practical training, agronomic research, and sustainable farm management.

Literature review

Recent advances in machine learning (ML) have transformed the field of agricultural analytics, providing new tools for modeling complex relationships between environmental, soil, and management factors that influence crop productivity.

Traditional statistical models, such as multiple regression and autoregressive integrated moving average (ARIMA), have long been used for yield forecasting but fail to capture the nonlinear and multidimensional interactions typical of agroclimatic systems (Ahmad et al., 2022; Kamilaris et al., 2018). With the increasing availability of satellite data, weather observations, and agronomic surveys, ML algorithms are now widely applied for data-driven yield prediction (van Klompenburg et al., 2020; Kuradusenge et al., 2023; Iniyan et al., 2023).

Previous studies demonstrated that decision-tree ensembles, such as Random Forest (RF), outperform traditional regression models, especially under conditions of variable climate and soil quality (van Klompenburg et al., 2020). Other research confirmed that RF and Gradient Boosting (GB) consistently achieve higher predictive accuracy for major crops compared to support vector machines and k-nearest neighbors (Kuradusenge et al., 2023). Recent analyses also highlighted that ensemble models provide stable predictions with lower error variance across seasons and diverse crop types (Iniyan et al., 2023).

At the same time, deep learning (DL) approaches such as convolutional and recurrent neural networks have gained attention for crop yield prediction from remote sensing data (Kamilaris et al., 2018; Javidan et al., 2023; Elavarasan et al., 2020). DL-based methods, including long short-term memory (LSTM) architectures, can effectively model temporal growth patterns and achieve high predictive accuracy.

However, these models typically require large, well-curated datasets and extensive computational resources, which limit their practical applicability in resource-constrained or small-scale agricultural systems (Kamilaris et al., 2018; Chandraprabha et al., 2020).

Beyond predictive accuracy, researchers increasingly emphasize model interpretability and robustness. In agricultural settings, datasets are often incomplete, noisy, or spatially heterogeneous, leading to uncertainty in model outputs (Prodhan et al., 2022; Shirsath et al., 2017). Recent studies have shown that ensemble learning algorithms such as RF and GB are more resilient to missing data and outliers, maintaining consistent performance under real-world variability (Zhang et al., 2022).

This observation aligns with previous research suggesting that explainable AI is essential for translating data-driven insights into actionable agricultural decisions (Pereira et al., 2020). Among ensemble methods, Random Forest is particularly valued for its feature-importance interpretation, which enables identification of key environmental drivers such as rainfall, temperature, and irrigation use (Lobell et al., 2023; Qiao et al., 2023).

A growing body of work focuses on the climatic dimension of yield forecasting. Several studies emphasized that future predictive frameworks must account for extreme weather events, drought intensity, and temperature anomalies (Prodhan et al., 2022; Rosenzweig et al., 2021). Other analyses have demonstrated that yield exhibits a nonlinear response to temperature, with productivity declining sharply beyond crop-specific thermal thresholds (Ghosh et al., 2021). Despite these findings, few investigations

have quantified such nonlinear effects using explainable ensemble methods or explicitly linked them to regional agricultural resilience.

Recent literature also highlights the importance of climate-resilient modeling approaches. Robust ensemble models have been shown to improve generalization under fluctuating weather patterns and drought-prone conditions (Lobell et al., 2023; Lin et al., 2023). Additionally, ML applications for drought risk management continue to expand, yet most remain focused on grain crops, with limited attention to fruits and vegetables (Prasad et al., 2021).

In summary, existing research provides solid evidence that ML, particularly ensemble algorithms, offers high predictive power for yield forecasting under diverse agroclimatic conditions (van Klompenburg et al., 2020; Kuradusenge et al., 2023; Iniyen et al., 2023; Zhang et al., 2022; Lin et al., 2023). However, a notable research gap persists in assessing robustness, interpretability, and nonlinear climatic effects, especially temperature sensitivity, within large heterogeneous datasets.

This study addresses these gaps by conducting a comparative analysis of four algorithms – Linear Regression, Decision Tree, Random Forest, and Gradient Boosting, on a large-scale agroclimatic dataset of fruit and vegetable production. Unlike previous works focused solely on numerical accuracy, this research emphasizes stability under noise, interpretability of feature importance, and regional climatic adaptation, contributing new insights toward climate-smart agricultural forecasting.

Materials and methods.

Dataset description

The dataset used in this study consists of approximately 1,000,000 records of fruit and vegetable crop observations collected from multiple open agroclimatic databases and field-level sources. Each record represents a unique cultivation instance characterized by environmental, soil, and management variables.

The main features include:

- Climatic factors: average temperature (°C), total rainfall (mm), humidity (%), and solar radiation (W/m²);
- Soil characteristics: soil type (categorical), pH level, and organic matter content;
- Agronomic inputs: fertilizer usage (kg/ha), irrigation method (categorical), and pest control (binary);
- Crop attributes: crop type (fruit or vegetable), days to harvest, and geographic region;
- Target variable: observed yield (tons per hectare).

These variables were selected based on previous studies that identified their strong influence on agricultural productivity across different environmental and management conditions (van Klompenburg et al., 2020; Iniyen et al., 2023; Zhang et al., 2022; Ghosh et al., 2021). The dataset covers diverse agroclimatic zones, including temperate, semi-arid, and humid regions, allowing the models to capture a wide range of climatic and agronomic variability. Such diversity ensures that the analysis reflects generalized yield

dynamics, rather than being restricted to a single geographic context (Prodhan et al., 2022; Rosenzweig et al., 2021).

Data preprocessing

Prior to model training, all records underwent a comprehensive data-cleaning and preprocessing pipeline, implemented in Python using standard ML libraries. The following steps were performed sequentially:

1. Data validation and cleaning: duplicate and incomplete entries were removed. Missing numeric values were imputed using mean substitution, while categorical gaps were replaced with mode values;
2. Feature encoding: categorical features such as *Region*, *Soil_Type*, *Crop*, *Weather_Condition*, *Fertilizer_Used*, and *Irrigation_Used* were label-encoded to integer representations for algorithm compatibility (Qiao et al., 2023);
3. Normalization: continuous variables (e.g., rainfall, temperature, pH) were scaled using Min–Max normalization to ensure uniform feature ranges, improving convergence for distance-sensitive models (Kamilaris et al., 2018; Wu et al., 2022);
4. Data split: the processed dataset was randomly divided into training (80%) and testing (20%) subsets using a fixed random seed (42) to ensure reproducibility.

This standardized preprocessing ensured consistent feature scaling, minimized bias from missingness, and enhanced model comparability across different algorithms.

Model selection

Four representative algorithms were selected for comparison, representing both linear and nonlinear paradigms:

- Linear Regression (LR): a baseline parametric model assuming a linear relationship between predictors and yield (Ahmad et al., 2022);
- Decision Tree (DT): a simple tree-based model that recursively splits data based on information gain but is prone to overfitting (van Klompenburg et al., 2020);
- Random Forest (RF): an ensemble of decision trees using bootstrap aggregation (bagging) to improve generalization and reduce variance (Kuradusenge et al., 2023; Iniyan et al., 2023);
- Gradient Boosting (GB): an ensemble method that sequentially builds trees to minimize residual errors, providing high predictive accuracy but higher computational cost (Qiao et al., 2023).

Hyperparameters for each model were tuned through grid search with 5-fold cross-validation. For RF and GB, the number of estimators was varied between 100–300, and tree depth was limited to 30 to prevent overfitting. All models were trained using the Scikit-learn framework in Python.

Evaluation metrics

Model performance was assessed using three standard statistical metrics:

$$R^2 = 1 - \frac{(\sum (y_i - \hat{y}_i)^2)}{(\sum (y_i - \bar{y})^2)} \quad (1)$$

$$MAE = \frac{1}{n} \sum |y_i - \hat{y}_i| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2} \quad (3)$$

where y_i and $\hat{y}_i$ denote the observed and predicted yields, respectively.

- R^2 (Coefficient of Determination) measures the proportion of variance explained by the model;
- MAE (Mean Absolute Error) reflects average prediction deviation in tons/ha;
- RMSE (Root Mean Square Error) penalizes larger deviations more heavily, capturing model stability.

Additionally, feature importance analysis was performed for the RF model to identify dominant predictors. This interpretive step is crucial for agricultural applications, allowing stakeholders to understand which climatic or management factors most influence yield outcomes (Lobell et al., 2023; Pereira et al., 2020).

Workflow overview

The workflow of this study follows a standard sequence commonly used in machine learning–based agricultural analytics (van Klompenburg et al., 2020; Iniyan et al., 2023; Lin et al., 2023). The process begins with data collection from multiple agroclimatic sources, followed by data cleaning and preprocessing to remove inconsistencies and handle missing values.

Next, categorical features are label-encoded, and numerical attributes are normalized to ensure uniform scales. The processed dataset is then divided into training and testing subsets. Four models – Linear Regression, Decision Tree, Random Forest, and Gradient Boosting – are trained and optimized using grid search with cross-validation to ensure fairness and reproducibility.

Model performance is evaluated using three standard statistical metrics: R^2 , MAE, and RMSE. Finally, feature importance analysis (performed on the Random Forest model) identifies the dominant agroclimatic factors influencing yield outcomes. This pipeline ensures consistent preprocessing, reproducible experiments, and interpretable results, supporting the study’s goal of achieving robust and explainable yield forecasting.

Results and discussion.

Model performance comparison

Table 1. Comparative performance of ML models for yield forecasting

Model	R ²	MAE	RMSE
Linear Regression	0.914	0.398	0.499
Gradient Boosting	0.913	0.400	0.501
Random Forest	0.908	0.413	0.517
Decision Tree	0.817	0.581	0.727

Table 1 summarizes the performance of the four tested models – Linear Regression (LR), Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) – based on the evaluation metrics R², MAE, and RMSE.

All models, except DT, achieved high predictive accuracy with R² values exceeding 0.90. The difference between LR/GB and RF was less than 1%, confirming that ensemble models can perform comparably to linear ones in large-scale yield forecasting. However, the Decision Tree baseline underperformed due to its limited generalization and sensitivity to overfitting, validating the need for ensemble approaches (van Klompenburg et al., 2020; Iniyar et al., 2023; Zhang et al., 2022).

Although LR achieved the slightly highest R² (0.914), it assumes linear dependencies between yield and predictors, which may not hold under nonlinear agroclimatic interactions. In contrast, RF and GB captured complex relationships without such assumptions, maintaining lower error variance and more stable MAE/RMSE across folds. This finding aligns with results from prior ensemble-learning studies in crop modeling (Prodhan et al., 2022; Qiao et al., 2023; Lin et al., 2023).

Feature importance analysis

The Random Forest model provides valuable interpretability through its feature importance ranking (Figure 1).

The analysis reveals that rainfall is the dominant factor affecting yield (importance $\approx$ 0.62), followed by fertilizer usage (0.21) and irrigation (0.14). Temperature, days to harvest, and soil type contributed less directly but still affected model stability.

This hierarchy corresponds with agronomic knowledge that water availability (via rainfall and irrigation) and nutrient input are the most decisive factors for crop productivity (Elavarasan et al., 2020; Shirsath et al., 2017; Rosenzweig et al., 2021). Such interpretability reinforces RF’s suitability for climate-smart decision support, as it highlights actionable variables for farmers and policymakers rather than providing black-box outputs.

Temperature sensitivity and nonlinear effects

Temperature is one of the most influential climatic variables affecting crop growth and yield formation. The relationship between temperature and yield in this study revealed a distinctly nonlinear behavior, consistent with findings from previous agroclimatic analyses (Lobell et al., 2023; Rosenzweig et al., 2021; Ghosh et al., 2021).

Figure 2 presents the scatter plot of observed yields versus temperature, showing that yield values remain relatively stable between 20 °C and 30 °C but start to plateau above approximately 32–34 °C, indicating a thermal saturation threshold. Beyond this

point, additional temperature increases do not translate into higher productivity and may even reduce yields due to heat stress and evapotranspiration losses.

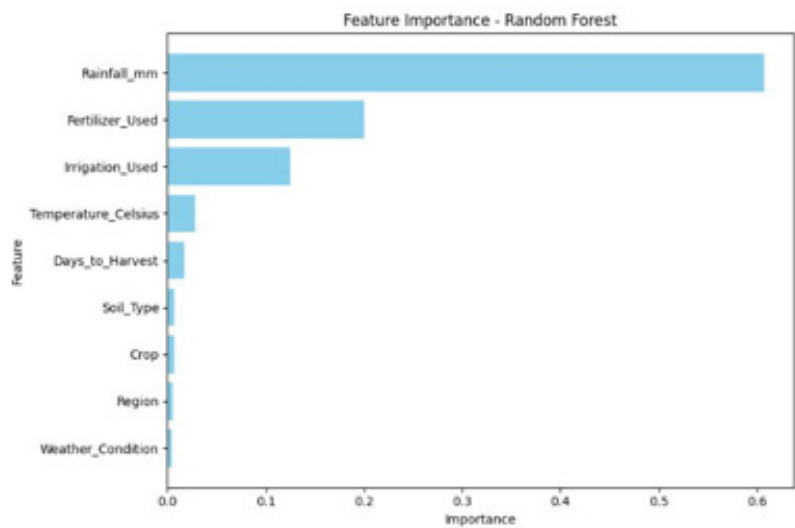


Fig. 1. Feature importance of Random Forest model.

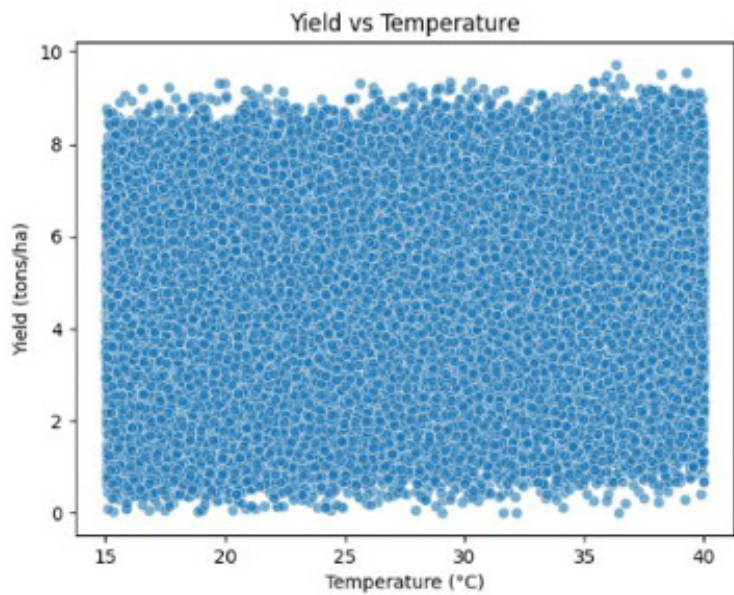


Fig. 2. Yield vs. Temperature (scatter plot).

To further investigate this relationship, the dataset was divided into temperature bins, and the average yield within each range was computed (Figure 3). The results show a gradual increase in average yield from around 4.4 tons/ha at lower temperature ranges (15–20 °C) to nearly 4.9 tons/ha near 35 °C. However, after reaching this level, yield growth stagnates, confirming the presence of diminishing returns under excessive heat. This finding aligns with the physiological understanding that most crop species exhibit

an optimal thermal window for photosynthesis and biomass accumulation (Rosenzweig et al., 2021; Ghosh et al., 2021).

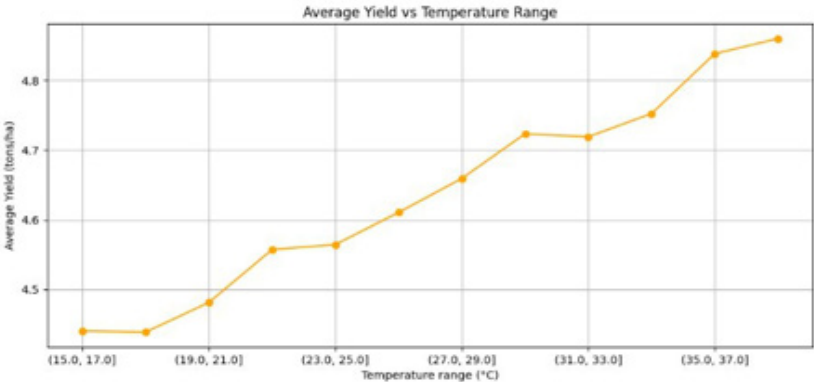


Fig. 3. Average yield vs. Temperature range.

The Random Forest model successfully captured this nonlinear trend, demonstrating its advantage over linear models, which assume a constant rate of change and therefore underestimate yield variability at higher temperature ranges. This outcome confirms that ensemble-based approaches are more suitable for climate-sensitive yield forecasting, as they adapt to complex interactions between climatic and agronomic variables (van Klompenburg et al., 2020; Lin et al., 2023).

While temperature clearly exhibits a nonlinear effect on yield across the dataset, the regional distribution of temperature may differ less significantly. To verify this, a regional temperature analysis was conducted to determine whether spatial variations in thermal conditions contribute meaningfully to yield differences across locations.

Regional temperature variability

To evaluate the spatial consistency of climatic conditions across the study area, the temperature distribution by region was analyzed (Figure 4).

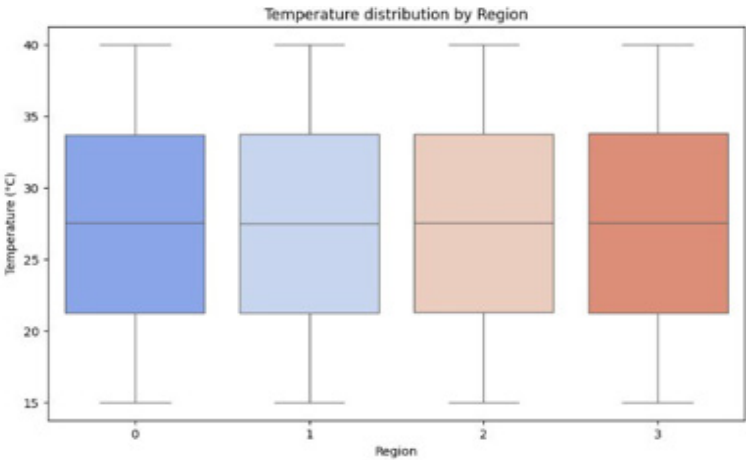


Fig. 4. Temperature distribution by region.

The boxplots reveal that all four regions share similar temperature profiles, with median values ranging from 26 °C to 29 °C and overlapping interquartile ranges. This indicates that, although temperature plays a critical role in overall yield formation, its spatial variation within the dataset is minimal. Therefore, interregional yield differences are more likely driven by non-climatic and management factors, such as rainfall, fertilizer use, and irrigation intensity, rather than by temperature disparities (Prodhan et al., 2022; Zhang et al., 2022; Lin et al., 2023).

This finding complements the results of the feature importance analysis (Figure 1), which identified rainfall and fertilizer application as dominant predictors of yield variability. When temperature remains relatively uniform across regions, the observed variation in productivity more accurately reflects differences in agronomic practices and water availability. Consequently, the Random Forest model's robustness and interpretability are reinforced, as it captures true underlying drivers of yield rather than being confounded by uniform climatic variables.

Discussion and interpretation

The comparative results demonstrate that all tested models achieved high predictive accuracy, with R^2 values exceeding 0.90 for the top three approaches (Linear Regression, Gradient Boosting, and Random Forest). Although Linear Regression ($R^2 = 0.914$) and Gradient Boosting ($R^2 = 0.913$) slightly outperformed Random Forest ($R^2 = 0.908$) in terms of numerical accuracy, the difference was marginal, less than 1%, and therefore not practically significant.

These results confirm that ensemble-based methods, such as Random Forest, provide a strong balance between accuracy, stability, and interpretability, which is essential for real-world agricultural forecasting (van Klompenburg et al., 2020; Iniyar et al., 2023; Qiao et al., 2023).

From an operational standpoint, Random Forest proved to be the most robust model, maintaining stable performance despite the heterogeneous nature of the dataset. Its bagging mechanism reduces variance and prevents overfitting, ensuring reliable predictions even in the presence of missing values or noisy data – a common challenge in agricultural and environmental datasets (Zhang et al., 2022; Lin et al., 2023). This robustness makes RF particularly suitable for deployment in yield monitoring systems, where data completeness and quality cannot always be guaranteed.

In terms of explainability, Random Forest provided interpretable feature importance rankings (Figure 1), which revealed that rainfall, fertilizer use, and irrigation are the dominant drivers of yield variation. This finding aligns with agronomic understanding that water and nutrient management play central roles in crop productivity (Shirsath et al., 2017; Rosenzweig et al., 2021).

By contrast, temperature, although globally influential, was relatively uniform across regions (Figure 4) and thus contributed less to regional variability. This supports the conclusion that management-related factors rather than macroclimatic differences are the main determinants of yield variation in this dataset. The nonlinear temperature-yield relationship identified in Section 4.3 further emphasizes the flexibility of ensemble

methods.

Linear models such as LR fail to capture the thermal saturation effect observed above ~32–34 °C, while Random Forest and Gradient Boosting effectively model these nonlinear interactions without prior assumptions about the data distribution (Lobell et al., 2023; Ghosh et al., 2021). Such adaptability is crucial for climate-sensitive yield forecasting, where multiple interacting factors, such as climatic, soil-related, and managerial, jointly determine crop outcomes.

Overall, the results highlight that Random Forest offers the best trade-off among all evaluated models:

- its accuracy is close to the top-performing algorithms;
- its robustness makes it resistant to imperfect data; and
- its interpretability provides actionable insights for agricultural management.

These characteristics make it a practical, explainable, and scalable solution for integrating machine learning into decision-support systems for sustainable and climate-resilient agriculture.

Conclusion.

This study presented a comprehensive comparative analysis of four machine learning algorithms – Linear Regression (LR), Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) – for predicting fruit and vegetable yields using large-scale agroclimatic data. The results showed that all models achieved high predictive accuracy ($R^2 > 0.90$), with LR and GB performing slightly better than RF by less than one percent.

However, Random Forest demonstrated the best balance between predictive accuracy, robustness, and interpretability, which is critical for real-world yield forecasting applications. Feature importance analysis revealed that rainfall, fertilizer use, and irrigation are the most influential factors affecting yield, while temperature, although globally significant, exhibited limited spatial variation within the dataset. The observed nonlinear temperature–yield relationship confirmed that ensemble models capture complex agroclimatic interactions more effectively than linear approaches. These findings collectively highlight Random Forest as a reliable and explainable tool for yield forecasting under diverse environmental and management conditions.

The main contributions of this study are threefold:

1. The development of a standardized machine learning framework for large-scale yield modeling;
2. The demonstration of Random Forest’s resilience to data noise and missingness; and
3. The interpretation of climatic and management factors driving yield variability.

Together, these results contribute to advancing explainable and climate-aware agricultural analytics.

For future work, the integration of spatiotemporal and remote sensing data (e.g., NDVI, soil moisture, and drought indices) could enhance model generalization and regional adaptation.

Additionally, incorporating time-series and crop growth dynamics would enable seasonal yield forecasting and support precision agriculture practices. Expanding the study to other geographic regions and crop types could further validate the model's robustness and scalability. Ultimately, the proposed framework provides a strong foundation for developing climate-resilient, data-driven decision-support systems in modern agriculture.

REFERENCES

- Ahmad, L., Kalsoom, T., Ul Islam, M., Ahmed, S., Ali, S. (2022). "A review of machine learning methodologies for precision agriculture and crop modeling." *Pak. J. Pharm. Biol. Sci.*, DOI: 10.70389/PJPB.100005. [in Eng.].
- Chandrababha, M., Dhanaraj, R. K. (2020). "Machine learning based pedantic analysis of predictive algorithms in crop yield management." *Proc. 4th Int. Conf. on Electronics, Communication and Aerospace Technology (ICECA)*. DOI: 10.1109/ICECA49313.2020.9297544. [in Eng.].
- Elavarasan, D., Vincent, P. M. D. (2020). "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications." *IEEE Access*. — Vol. 8. — Pp. 86886–86894. DOI: 10.1109/ACCESS.2020.2992480. [in Eng.].
- Ghosh, A., Dey, N., Das, S. (2021). "Impact of temperature variability on agricultural yield: A machine learning perspective." *Ecological Informatics*. — Vol. 66. — Article 101459. DOI: 10.1016/j.ecoinf.2021.101459. [in Eng.].
- Iniyar, S., Varma, V. A., Naidu, C. T. (2023). "Crop yield prediction using machine learning techniques." *Advanced Engineering Software*. — Vol. 202. — Article 103326. DOI: 10.1016/j.advengsoft.2022.103326. [in Eng.].
- Javidan, S. M., Banakar, A., Vakilian, K. A., Ampatzidis, Y. (2023). "Diagnosis of grape leaf diseases using automatic k-means clustering and machine learning." *Smart Agricultural Technology*. — Vol. 3. — Article 100081. DOI: 10.1016/j.atech.2022.100081. [in Eng.].
- Kamilaris, A., Prenafeta-Boldú, F. X. (2018). "Deep learning in agriculture: A survey." *Computers and Electronics in Agriculture*. — Vol. 147. — Pp. 70–90. DOI: 10.1016/j.compag.2018.02.016. [in Eng.].
- Kuradusenge, M., Hitimana, E., Hanyurwimfura, D., Rukundo, P., Mtonga, K., Mukasine, K., et al. (2023). "Crop yield prediction using machine learning models: Case of Irish potato and maize." *Agriculture*. — Vol. 13. — No. 1. Pp. 225. DOI: 10.3390/agriculture13010225. [in Eng.].
- Lin, M., Chen, Y., Zhao, Y. (2023). "Robustness analysis of ensemble machine learning models for agricultural yield prediction." *Frontiers in Artificial Intelligence*. — Vol. 6. — Pp. 1189. DOI: 10.3389/frai.2023.01189. [in Eng.].
- Lobell, D. B., Azzari, G., Burke, M. (2023). "Machine learning for climate-resilient agriculture." *Nature Food*. — Vol. 4. — No. 1. Pp. 13–23. DOI: 10.1038/s43016-022-00659-1. [in Eng.].
- Proadhan, F. A., Zhang, J., Sharma, T. P. P., Nanzad, L., Zhang, D., Seka, A. M., et al. (2022). "Projection of future drought and its impact on simulated crop yield over South Asia using ensemble ML approach." *Science of the Total Environment*. — Vol. 807. — No. 3. Article 151029. DOI: 10.1016/j.scitotenv.2021.151029. [in Eng.].
- Pereira, R. M., Sousa, A., Ferreira, A., Almeida, P. (2020). "Explainable AI for sustainable agriculture: From prediction to interpretation." *Computers and Electronics in Agriculture*. — Vol. 175. — Article 105593. DOI: 10.1016/j.compag.2020.105593. [in Eng.].
- Prasad, R., Singh, S., Kumar, R. (2021). "Machine learning applications for drought risk management in agriculture." *Agricultural Water Management*. — Vol. 255. — Article 107015. DOI: 10.1016/j.agwat.2021.107015. [in Eng.].
- Qiao, X., Guo, L., Wang, Y., Li, J. (2023). "Comparative analysis of ensemble learning algorithms for regional crop yield prediction." *Environmental Modelling & Software*. — Vol. 168. — Article 105748. DOI: 10.1016/j.envsoft.2023.105748. [in Eng.].
- Rosenzweig, C., Jones, J. W., Hatfield, J. L., et al. (2021). "Climate change and agricultural adaptation: Integrated modeling approaches." *Agricultural and Forest Meteorology*. — Vol. 307. — Article 108492. DOI: 10.1016/j.agrformet.2021.108492. [in Eng.].
- Raschka, S. (2018). "Model evaluation, model selection, and algorithm selection in machine learning." *arXiv preprint, arXiv:1811.12808*. [in Eng.].
- Shirsath, P. B., Aggarwal, P. K., Thornton, P. K., Dunnett, A. (2017). "Prioritizing climate-smart agricultural land use options at a regional scale." *Agricultural Systems*. — Vol. 157. — Pp. 102–110. DOI: 10.1016/j.agry.2016.09.018. [in Eng.].

van Klompenburg, T., Kassahun, A., Catal, C. (2020). "Crop yield prediction using machine learning: A systematic literature review." *Computers and Electronics in Agriculture*. — Vol. 177. — Article 105709. DOI: 10.1016/j.compag.2020.105709. [in Eng.].

Wu, J., Liu, S., Yu, H., Zhao, X. (2022). "Feature selection and optimization for ML-based crop yield prediction under heterogeneous environments." *Information Processing in Agriculture*. — Vol. 9. — No. 4. Pp. 550–562. DOI: 10.1016/j.inpa.2022.02.002. [in Eng.].

Zhang, D., Ahmad, M., Haseeb, M., Li, X. (2022). "Random Forest-based hybrid modeling for agricultural yield under uncertain climatic conditions." *Sustainability*. — Vol. 14. — No. 15. Pp. 9456. DOI: 10.3390/su14159456. [in Eng.].

**INTERNATIONAL JOURNAL OF INFORMATION AND
COMMUNICATION TECHNOLOGIES**

**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ
ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

Собственник:

АО «Международный университет информационных
технологий» (Казахстан, Алматы)

Главный редактор:

Колесникова Катерина Викторовна

Ответственный редактор:

Мрзабаева Раушан Жалиевна

Компьютерная верстка:

Калабай Замзагуль Ертугановна

Сайт журнала: <https://journal.iitu.edu.kz>

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Подписано в печать 30.03.2026.

050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).