

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



**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ  
КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР  
ЖУРНАЛЫ**

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ  
ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

**INTERNATIONAL JOURNAL OF INFORMATION  
AND COMMUNICATION TECHNOLOGIES**

**2025 (21) 1**

*ақпан - наурыз*

ISSN 2708–2032 (print)  
ISSN 2708–2040 (online)

## БАС РЕДАКТОР:

**Исахов Асылбек Абдишимович** — есептеу теориясы саласында математика бойынша PhD доктор, "Компьютерлік ғылымдар және информатика" бағыты бойынша қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университетінің Басқарма Төрағасы – Ректор (Қазақстан)

## БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:

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

## ҒАЛЫМ ХАТШЫ:

**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, КР 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 Токсужаевич** — PhD, қауымдастырылған профессор, Халықаралық ақпараттық технологиялар университеті цифрландыру және инновациялар жөніндегі проректор (Қазақстан)

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

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

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

**Абдикаликова 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)

**Дамелия 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 Азизбековна** — 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 Сахиябай** — техника ғылымдарының кандидаты, профессор, Халықаралық ақпараттық технологиялар университеті компьютерлік инженерия кафедрасының профессор-зерттеушісі (Қазақстан)

**Алимереб 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)

**Талеуш Валдас** — PhD, Адам Мицкевич атындағы (Польша) университеттің проректоры

**Мамырбаев Оркен Жұмажанович** — PhD, КР 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 сәулет университеті жобаларды басқару кафедрасының меңгерушісі (Украина)

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

## 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)

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

Қазақстан Республикасы Ақпарат және қоғамдық даму министрлігіне мерзімді баспасөз басылымын есепке қою туралы куәлік № KZ82VPYU00020475, 20.02.2020 ж. берілген

Тақырып бағыты: ақпараттық технологиялар, ақпараттық қауіпсіздік және коммуникациялық технологиялар, әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технология.

Мерзімділігі: жылына 4 рет.

Тираж: 100 дана.

Редакция мекенжайы: 050040 Алматы қ., Манас к., 34/1, каб. 709, тел: +7 (727) 244-51-09.

E-mail: [ijict@iitu.edu.kz](mailto:ijict@iitu.edu.kz)

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

© Халықаралық ақпараттық технологиялар университеті АҚ, 2025

Журнал сайты: <https://journal.iitu.edu.kz> © Авторлар ұжымы, 2025

---

## ГЛАВНЫЙ РЕДАКТОР

**Исахов Асылбек Абдиашимович** — доктор PhD по математике в области теории вычислимости, ассоциированный профессор по направлению "Компьютерные науки и информатика", Председатель Правления – Ректор Международного университета информационных технологий (Казахстан)

## ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

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

## УЧЕНЫЙ СЕКРЕТАРЬ:

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

**Рысбайұлы 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)

**Дамеля Максумовна 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)

**Айтмагамбетов 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)

**Алмисреб 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)

**Талеуш 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)

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

## 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)

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

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

Тематическая направленность: информационные технологии, информационная безопасность и коммуникационные технологии, цифровые технологии в развитии социально-экономических систем.

Периодичность: 4 раза в год.

Тираж: 100 экземпляров.

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

E-mail: [ijict@iitu.edu.kz](mailto:ijict@iitu.edu.kz)

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

© АО Международный университет информационных технологий, 2025

© Коллектив авторов, 2025

#### EDITOR-IN-CHIEF

**Assylbek Issakhov** — PhD in Mathematics in Computability Theory, associate professor in “Computer Science and Informatics,” Chairman of the Board – Rector of the International Information Technology University (Kazakhstan)

#### DEPUTY EDITOR-IN-CHIEF

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

#### ACADEMIC SECRETARY

**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)

**Nurzhhan 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 Yeskendiroya** — 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)

#### EDITOR

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

«International Journal of Information and Communication Technologies»

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Owner: International Information Technology University JSC (Almaty).

The certificate of registration of a periodical printed publication in the Ministry of Information and Social Development of the Republic of Kazakhstan, Information Committee No. KZ8ZVPY00020475, issued on 20.02.2020.

Thematic focus: information technology, digital technologies in the development of socio-economic systems, information security and communication technologies

Periodicity: 4 times a year.

Circulation: 100 copies.

Editorial address: 050040. Manas st. 34/1, Almaty. +7 (727) 244-51-09. E-mail: [ijict@iitu.edu.kz](mailto:ijict@iitu.edu.kz)

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

© International Information Technology University JSC, 2025

© Group of authors, 2025

## МАЗМҰНЫ

### ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖҮЙЕЛЕРДІ ДАМУДАҒЫ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

**М.М. Жалгасова, К.В. Колесникова**

ІСВ ҚҰЗЫРЕТТІЛІК МОДЕЛІН ЖОБАЛАРДЫ БАСҚАРУДЫҢ САЛАЛЫҚ ҚАЖЕТТІЛІКТЕРІНЕ  
БЕЙІМДЕУ.....8

**Г. Мауина, А. Найзағараева, Э. Тулегенова, Б. Жүсіпбек, М.У. Худойберганов**

АУЫЛШАРУАШЫЛЫҚ РЕСУРСТАРЫН БАСҚАРУДЫ ОҢТАЙЛАНДЫРУ ҮШІН SNAR ЖӘНЕ PCA  
ҚОЛДАНУ АРҚЫЛЫ ФАКТОРЛАРДЫҢ МАҢЫЗДЫЛЫҒЫН ТАЛДАУ.....21

**Б. Тасуов, А.Н. Аманбаева, С. Сактиото**

БІЛІМ АЛУШЫЛАРДЫҢ ЦИФРЛЫҚ САУАТТЫЛЫҒЫН ҚАЛЫПТАСТЫРУ  
МАҚСАТЫНДА БҮЛТТЫ ТЕХНОЛОГИЯЛАРДЫ ПАЙДАЛАНУ.....40

**Д.А. Шрымбай, Э.Т. Адылбекова, Х.И. Бұлбұл**

БОЛАШАҚ МҰҒАЛІМДЕРДІҢ КӘСІБИ ДАЙЫНДЫҒЫН ДАМУ МӘСЕЛЕЛЕРІ.....58

### АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

**А.А. Быков, А. Нұрланұлы, Н.А. Дауренбаева**

ТЕМІР ЖОЛДЫҢ ЖЕР ТӨСЕМІНДЕГІ ГЕОФИЗИКАЛЫҚ ОҚИҒАЛАРДЫ БОЛЖАУ  
ӘДІСТЕМЕСІ.....71

**К.Б. Багитова, Ш.Ж. Мусиралиева, Л. Курмангазиева, Ж. Молдашева, И.Терейковский**  
ӘЛЕУМЕТТІК ЖЕЛІЛЕРДЕГІ ГРАФИКАЛЫҚ РЕСУРСТАРДЫ ӨНДЕУ МОДЕЛІ.....82

**А.Б. Касекева, А.К. Адилова, А.А. Шекербек, А.С. Баегизова, К.О. Рахимов**

БАЛА ДАМУЫНА ӘСЕР ЕТЕТІН ҚАЗАҚ ЛИНГВИСТИКАСЫНЫҢ ӘЛЕУМЕТТІК-МӘДЕНИ  
ДӘСТҮРЛЕРІН ЗЕРТТЕУГЕ АРНАЛҒАН АҚПАРАТТЫҚ ЖҮЙЕ ҚҰРУ».....98

**С.Б. Муханов, Д.М. Амрин, С.Ж. Жакыпбеков**

ТАРТЫЛҒАН ЖЕЛІЛЕРДЕГІ КЕЗЕКТІК ЖҮЙЕЛЕРДІҢ ӨЗАРА ӘРЕКЕТТЕРІН ЖҮЙЕ АРҚАЛЫҚ  
ТАЛДАУ.....113

**А. Оспанов, П. Алонсо-Жорда, А. Жумадилаева**

ІОТ ДАТЧИКТЕРІ МЕН МАШИНАЛЫҚ ОҚУ ӘДІСТЕРІН ПАЙДАЛАНУАРҚЫЛЫ  
ҚАЗЫРҒЫ ҚОЛДАНЫЛАТЫН ҚҰРМА ҚОЛДАУДЫҢ ОПТИМИЗАЦИЯСЫ: ЭМПИРИКАЛЫҚ  
ЗЕРТТЕУ.....127

**А.Т. Турсынова, Б.С. Омаров**

МИ ИНСУЛЬТІНІҢ КТ КЕСКІНІН КЛАССИФИКАЦИЯЛАУҒА АРНАЛҒАН КӨРУ  
ТРАНСФОРМАТОРЛАРЫ.....144

**А.Г. Шаушенова, М.Ж. Базарова, Ж.Ж. Ажибекова, С. Шадинова, К.С. Бакенова**

ӘРТҮРЛІ МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІ АРҚЫЛЫ ҚҰЖАТТАРДЫ АВТОМАТТЫ  
ТАЛДАУ МОДЕЛІН ҚҰРУ.....156

### АҚПАРАТТЫҚ ҚАУІПСІЗДІК ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАРҒА АРНАЛҒАН

**А.З. Айтмагамбетов, С.Ж. Жұмағали, Е.К. Коньсыбаев, М.М. Онгарбаева, И.В. Мелешкина**

ҚАУІПСІЗ ЖӘНЕ ТИІМДІ ЛОГИСТИКА ҮШІН ИНТЕЛЛЕКТУАЛДЫ НАВИГАЦИЯЛЫҚ ПЛОМБАНЫ  
ӨЗІРЛЕУ.....170

**Б.А. Кумалаков, А.Б. Казиз**

ҒИМАРАТТАРДАҒЫ KUBERNETES АРҚЫЛЫ ОРКЕСТРЛЕНГЕН КӨПАГЕНТТІК  
ЖҮЙЕЛЕРДЕГІ АҚАУҒА ТӨЗІМДІЛІК ПЕН СЕНІМДІЛІК: УНИВЕРСИТЕТТІҢ КЕСТЕСІН  
ЖОСПАРЛАУ КЕЙСІ.....185

**Л. Рзаева, Д. Поголовкин, И. Шайя**

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫ МЕН ВЕКТОРЛЫҚ ДЕРЕКҚОРДЫ ПАЙДАЛАНА  
ОТЫРЫП, ХАТ АЛМАСУЛАРДЫ ТАЛДАУ ҚЫЗМЕТІН ЦИФРЛЫҚ КРИМИНАЛИСТИКА ҮШІН  
ӨЗІРЛЕУ.....201

**Е. Чуракова, О. Новиков, О. Барановский, Т.В. Бабенко, Н.Е. Аскарбекова**

ГЕНЕТИКАЛЫҚ БАҒДАРЛАМАЛУ ӘДІСІМЕН ШАБЫЛ ВЕКТОРЛАРЫН ҚАЙТА ҚҰРУ.....226



## СОДЕРЖАНИЕ

## ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИ СОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ

**М.М. Жалгасова, К.В. Колесникова**

АДАПТАЦИЯ МОДЕЛИ КОМПЕТЕНЦИЙ ИСВ К ОТРАСЛЕВЫМ ПОТРЕБНОСТЯМ УПРАВЛЕНИЯ ПРОЕКТАМИ.....8

**Г.М. Мауина, А.А. Найзагараева, Э.Н. Тулегенова, Б.К. Жусипбек, М.У. Худойберганов**  
АНАЛИЗ ЗНАЧИМОСТИ ФАКТОРОВ С ИСПОЛЬЗОВАНИЕМ SHAP И PCA ДЛЯ ОПТИМИЗАЦИИ УПРАВЛЕНИЯ АГРАРНЫМИ РЕСУРСАМИ.....21**Б. Тасуов, А.Н. Аманбаева, С. Сактино**

ИСПОЛЬЗОВАНИЕ ОБЛАЧНЫХ ТЕХНОЛОГИЙ В ЦЕЛЯХ ФОРМИРОВАНИЯ ЦИФРОВОЙ ГРАМОТНОСТИ ОБУЧАЮЩИХСЯ.....40

**Д.А. Шрымбай, Э.Т. Адылбекова, Х.И. Бюльбюль**

ПРОБЛЕМЫ РАЗВИТИЯ ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ БУДУЩИХ УЧИТЕЛЕЙ.....58

## ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

**А.А. Быков, А. Нурланулы, Н.А. Дауренбаева**

МЕТОДИКА ПРОГНОЗИРОВАНИЯ ГЕОФИЗИЧЕСКИХ СОБЫТИЙ В ЗЕМЛЯНОМ ПОЛОТНЕ ЖЕЛЕЗНОЙ ДОРОГИ.....71

**К.Б. Багитова, Ш.Ж. Мусиралиева, Л. Курмангазиева, Ж. Молдашева, И. Терейковский**  
МОДЕЛЬ ОБРАБОТКИ ГРАФИЧЕСКИХ РЕСУРСОВ СОЦИАЛЬНЫХ СЕТЕЙ.....82**А.Б. Касекеева, А.К. Адилова, А.А. Шекербек, А.С. Баегизова, К.О. Рахимов**

ПОСТРОЕНИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ ИЗУЧЕНИЯ СОЦИОКУЛЬТУРНЫХ ТРАДИЦИЙ КАЗАХСКОЙ ЛИНГВИСТИКИ, ВЛИЯЮЩИХ НА РАЗВИТИЕ РЕБЕНКА.....98

**С.Б. Муханов, Д.М. Амрин, С.Ж. Жакыпбеков**

МЕЖСИСТЕМНЫЙ АНАЛИЗ ВЗАИМОДЕЙСТВИЙ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ В РАСПРЕДЕЛЕННЫХ СЕТЯХ.....113

**А. Оспанов, П. Алонсо-Жорда, А. Жумадилаева**

ОПТИМИЗАЦИЯ МОНИТОРИНГА СКЛАДА С ИСПОЛЬЗОВАНИЕМ ДАТЧИКОВ ИОТ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ: ЭМПИРИЧЕСКОЕ ИССЛЕДОВАНИЕ.....127

**А.Т. Турсынова, Б.С. Омаров**

ТРАНСФОРМАТОРЫ ЗРЕНИЯ ДЛЯ КЛАССИФИКАЦИИ КТ-ИЗОБРАЖЕНИЙ ИНСУЛЬТА ГОЛОВНОГО МОЗГА.....144

**А.Г. Шаушенова, М.Ж. Базарова, Ж.Ж. Ажибекова, К.С. Шадинова, К.С. Бакенова**

СОЗДАНИЕ МОДЕЛИ АВТОМАТИЧЕСКОГО АНАЛИЗА ДОКУМЕНТОВ С ИСПОЛЬЗОВАНИЕМ РАЗЛИЧНЫХ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ.....156

## ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

**А.З. Айтмагамбетов, С.Ж. Жумагали, Е.К. Конысбаев, М.М. Онгарбаева, И.В. Мелешкина**

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ НАВИГАЦИОННОЙ ПЛОМБЫ ДЛЯ БЕЗОПАСНОЙ И ЭФФЕКТИВНОЙ ЛОГИСТИКИ.....170

**Б.А. Кумалаков, А.Б. Казиз**

ОТКАЗОУСТОЙЧИВОСТЬ И НАДЕЖНОСТЬ В МУЛЬТИАГЕНТНЫХ СИСТЕМАХ, ОРКЕСТРИРУЕМЫХ KUBERNETES: КЕЙС РАСПИСАНИЯ УНИВЕРСИТЕТА.....185

**Л. Рзаева, Д. Поголовкин, И. Шайя**

РАЗРАБОТКА СЕРВИСА АНАЛИЗА ПЕРЕПИСОК С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ВЕКТОРНОЙ БАЗЫ ДАННЫХ ДЛЯ ЦИФРОВОЙ КРИМИНАЛИСТИКИ.....201

**Е. Чуракова, О. Новиков, О. Барановский, Т.В. Бабенко, Н.Е. Аскарбекова**

РЕКОНСТРУКЦИЯ ВЕКТОРОВ АТАК МЕТОДОМ ГЕНЕТИЧЕСКОГО ПРОГРАММИРОВАНИЯ.....226

## CONTENT

## DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF SOCIO-ECONOMIC SYSTEMS

**M.M. Zhalgassova, K.V. Kolesnikova**ADAPTING THE ICB COMPETENCY MODEL TO INDUSTRY-SPECIFIC PROJECT  
MANAGEMENT NEEDS.....8**G. Mauina, A. Naizagarayeva, E. Tulegenova, B. Zhussipbek, M.U. Khudoyberganov**  
FACTOR IMPORTANCE ANALYSIS USING SHAP AND PCA FOR OPTIMIZING AGRICULTURAL  
RESOURCE MANAGEMENT.....21**B. Tassuov, A.N. Amanbayeva, S.Saktioto**  
USING CLOUD TECHNOLOGY FOR THE FORMATION OF DIGITAL  
LITERACY OF STUDENTS.....40**D. Shrymbay, E. Adylbekova, H.I. Bulbul**  
PROBLEMS OF DEVELOPMENT OF FUTURE TEACHERS PROFESSIONAL TRAINING.....58

## INFORMATION TECHNOLOGY

**A.A. Bykov, A. Nurlanuly, N.A. Daurenbayeva**  
METHOD OF FORECASTING GEOPHYSICAL EVENTS IN THE RAILWAY GROUND BED.....71**K. Bagitova, Sh. Mussiraliyeva, L. Kurmangazyeva, Zh. Moldasheva, I. Tereikovskiy**  
THE MODEL FOR PROCESSING GRAPHIC RESOURCES OF SOCIAL NETWORKS.....82**A.B. Kassekeyeva, A.K. Adilova, A.A. Shekerbek, A. Bayegizova, K.O. Rahimov**  
CREATION OF AN INFORMATION SYSTEM FOR THE STUDY OF SOCIO-CULTURAL TRADITIONS OF  
KAZAKH LINGUISTICS THAT INFLUENCE CHILD DEVELOPMENT.....98**S.B. Mukhanov, D.M. Amrin, S.Zh. Zhakybpekov**  
CROSS-SYSTEM ANALYSIS OF QUEUEING SYSTEMS INTERACTIONS IN DISTRIBUTED  
NETWORKS OPTIMIZING .....113**A. Ospanov, Alonso-Jord Pedro, A. Zhumadillayeva**  
WAREHOUSE MONITORING WITH IOT SENSORS AND MACHINE  
LEARNING: AN EMPIRICAL STUDY.....127**A.T. Tursynova, B.S. Omarov**  
VISION TRANSFORMERS FOR CLASSIFICATION OF CT IMAGES OF BRAIN STROKE.....144**A.G. Shaushenova, M.Zh. Bazarova, Zh.Zh. Azhibekova, K.S. Shadinova, K.S. Bakenova**  
CREATION OF AUTOMATIC DOCUMENT ANALYSIS MODEL USING DIFFERENT MACHINE  
LEARNING ALGORITHMS.....156

## INFORMATION SECURITY AND COMMUNICATION TECHNOLOGIES

**A.Z. Aitmagambetov, S.Zh. Zhumagali, Ye.K. Konysbayev, M.M. Ongarbayeva, I.V. Meleshkina**  
DEVELOPMENT OF AN INTELLIGENT NAVIGATION SEAL FOR SAFE AND  
EFFICIENT LOGISTICS.....170**B. Kumalakov, A. Kaziz**  
FAULT TOLERANCE AND RELIABILITY IN KUBERNETES-ORCHESTRATED MULTI-AGENT  
SYSTEMS: UNIVERSITY SCHEDULING CASE STUDY.....185**L. Rzayeva, D. Pogolovkin, I. Shayea**  
DEVELOPMENT OF A CORRESPONDENCE ANALYSIS SERVICE USING ARTIFICIAL INTELLIGENCE  
TECHNOLOGY AND A VECTOR DATABASE FOR DIGITAL FORENSICS.....201**Y. Churakova, O. Novikov, O. Baranovskyi, T.V. Babenko, N.Y. Askarbekova**  
RECONSTRUCTING ATTACK VECTORS USING GENETIC PROGRAMMING.....226

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

Vol. 6. Is. 1. Number 21 (2025). Pp. 156–169

Journal homepage: <https://journal.iitu.edu.kz><https://doi.org/10.54309/IJICT.2025.21.1.011>

УДК 004.931

## CREATION OF AUTOMATIC DOCUMENT ANALYSIS MODEL USING DIFFERENT MACHINE LEARNING ALGORITHMS

*A.G. Shaushenova<sup>1</sup>, M.Zh. Bazarova<sup>2</sup>, Zh.Zh. Azhibekova<sup>3</sup>, K.S. Shadinova<sup>3</sup>,  
K.S. Bakenova<sup>4</sup>*

<sup>1</sup>S. Seifullin Kazakh Agro Technical Research University, Astana, Kazakhstan;

<sup>2</sup>Sarsen Amanzholov East Kazakhstan university, Ust-Kamenogorsk, Kazakhstan;

<sup>3</sup>Kazakh National University named after S.D. Asfendiyarova, Almaty, Kazakhstan;

<sup>4</sup>Astana International University, Astana, Kazakhstan.

E-mail: [madina\\_vkgtu@mail.ru](mailto:madina_vkgtu@mail.ru)

**Shaushenova A.** — S. Seifullin Kazakh Agro Technical Research University, Candidate of Technical Sci-ences, associate professor, Head of the Department of Information Systems, Astana, Kazakhstan

E-mail: [shaushenova\\_78@mail.ru](mailto:shaushenova_78@mail.ru), <https://orcid.org/0000-0002-3164-3688>;

**Bazarova M.** — Sarsen Amanzholov East Kazakhstan university, associate professor of the Department of computer modeling and information technology, PhD, Ust-Kamenogorsk, Kazakhstan

E-mail: [madina\\_vkgtu@mail.ru](mailto:madina_vkgtu@mail.ru), <https://orcid.org/0000-0003-2580-6580>;

**Azhibekova Zh.** — Candidate of sciences in pedagogical, head of the department of information communication technologies, Kazakh National University named after S.D. Asfendiyarova, Almaty, Kazakhstan

E-mail: [Azhibekova.z@kaznmu.kz](mailto:Azhibekova.z@kaznmu.kz), <https://orcid.org/0000-0002-4396-1261>;

**Shadinova K.** — Asfendiyarov Kazakh National Medical University, associate professor of the Department of information and communication technology, Almaty, Kazakhstan E-mail:

[shadinova.ks@mail.ru](mailto:shadinova.ks@mail.ru), <https://orcid.org/0009-0006-5534-7927>;

**Bakenova K.** — Astana International University, lecturer at the Pedagogical Institute, L.N.Gumilyov Eur-asian National University, doctoral student of the Department of Information Security, Astana, Kazakhstan

E-mail: [bakenova.k@bk.ru](mailto:bakenova.k@bk.ru), <https://orcid.org/0009-0004-2567-173X>.

© A.G. Shaushenova, M.Zh. Bazarova, Zh.Zh. Azhibekova, K.S. Shadinova, K.S. Bakenova, 2025

**Abstract.** This article presents a model for automatic document analysis.

It is based on the TF-IDF (Term Frequency-Inverse Document Frequency) method combined with various machine learning algorithms, including SVM (Support Vector Machine), Random Forest, and Word2Vec+SVM. The aim of the study is to compare the effectiveness of these methods in text classification tasks and identify the most efficient approach. Experimental results showed that the hybrid model using a combination of TF-IDF and Word2Vec with SVM achieved the highest accuracy (90.2%) and F1-score (82.52%). The TF-IDF method allows evaluating the importance of terms in the text, while Word2Vec converts words into vector representations, improving semantic matching. The SVM algorithm effectively divides data into

classes using hyperplanes, while Random Forest enhances classification quality through the use of decision tree ensembles. Additionally, the study highlighted the importance of text preprocessing (tokenization, normalization, stopword removal, and lemmatization), which significantly improves classification performance. The proposed model can be effectively applied in areas such as information retrieval, topic modeling, and automatic document summarization. Such hybrid approaches improve the accuracy and reliability of automatic text analysis, offering opportunities for adaptation to multilingual environments and the integration of new data sources. Experimental results confirmed the effectiveness of this approach for solving complex tasks such as sentiment analysis, document categorization, and topic modeling. This study is a significant step toward developing new solutions in the field of automatic text analysis.

**Key words:** Automated document analysis, Machine learning, Term Frequency-Inverse Document Frequency, Natural language processing, Text summarization

**For citation:** A.G. Shaushenova, M.Zh. Bazarova, Zh.Zh. Azhibekova, K.S. Shadinova, K.S. Bakenova. CREATION OF AUTOMATIC DOCUMENT ANALYSIS MODEL USING DIFFERENT MACHINE LEARNING ALGORITHMS// INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES. 2025. Vol. 6. No. 21. Pp. 156–169 (In Kaz.). <https://doi.org/10.54309/IJICT.2025.21.1.011>.

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

## ӘРТҮРЛІ МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІ АРҚЫЛЫ ҚҰЖАТТАРДЫ АВТОМАТТЫ ТАЛДАУ МОДЕЛІН ҚҰРУ

**А.Г. Шаушенова<sup>1</sup>, М.Ж. Базарова<sup>2</sup>, Ж.Ж. Ажибекова<sup>3</sup>, С. Шадинова<sup>3</sup>,  
К.С. Бакенова<sup>4</sup>**

<sup>1</sup>С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, Астана, Қазақстан;

<sup>2</sup>С. Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен, Қазақстан;

<sup>3</sup>С.Ж. Асфендияров атындағы Қазақ Ұлттық Медициналық Университеті, Алматы, Қазақстан;

<sup>4</sup>Астана халықаралық университеті, Астана, Қазақстан.  
E-mail: [madina\\_vkgtu@mail.ru](mailto:madina_vkgtu@mail.ru)

**Шаушенова А.** — С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, «Ақпараттық жүйелер» кафедрасының меңгерушісі, техника ғылымдарының кандидаты, қауымдастырылған профессор, Астана, Қазақстан  
E-mail: [shaushenova\\_78@mail.ru](mailto:shaushenova_78@mail.ru), <https://orcid.org/0000-0002-3164-3688>;

**Базарова М.** — С. Аманжолов атындағы Шығыс Қазақстан университеті, «Компьютерлік үлгілеу және ақпараттық технологиялар» кафедрасының қауымдастырылған профессоры, PhD, Өскемен қ., Қазақстан  
E-mail: [madina\\_vkgtu@mail.ru](mailto:madina_vkgtu@mail.ru), <https://orcid.org/0000-0003-2580-6580>;

**Ажибекова Ж.** — «Ақпараттық коммуникациялық технологиялар» кафедрасының меңгерушісі, педагогика ғылымдарының кандидаты, С.Ж. Асфендияров атындағы Қазақ Ұлттық

Медициналық Университеті, Алматы, Қазақстан

E-mail: Azhibekova.z@kaznmu.kz, <https://orcid.org/0000-0002-4396-1261>;

**Шадинова К.** — С.Ж.Асфендияров атындағы Қазақ ұлттық медицина университеті, ақпараттық-коммуникациялық технологиялар кафедрасының қауымдастырылған профессоры, Алматы, Қазақстан

E-mail: shadinova.ks@mail.ru, <https://orcid.org/0009-0006-5534-7927>;

**Бакенова К.** — Astana International University, Педагогикалық институтының оқытушысы, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Ақпараттық қауіпсіздік кафедрасының докторанты, Астана, Қазақстан

E-mail: bakenova.k@bk.ru, <https://orcid.org/0009-0004-2567-173X>.

© А.Г. Шаушенова, М.Ж. Базарова, Ж.Ж. Ажибекова, К.С. Шадинова, К.С. Бакенова, 2025

**Аннотация.** Бұл мақалада құжаттарды автоматты түрде талдауға арналған модель ұсынылған. Ол TF-IDF (Term Frequency-Inverse Document Frequency) әдісін әртүрлі машина оқыту алгоритмдерімен, соның ішінде SVM (Support Vector Machine), Random Forest және Word2Vec+SVM комбинациясымен үйлестіру негізінде құрылған. Зерттеудің мақсаты – мәтіндерді жіктеу міндеттерінде осы әдістердің тиімділігін салыстыру және ең нәтижелі тәсілді анықтау. Эксперименттік нәтижелер көрсеткендей, TF-IDF және Word2Vec комбинациясын SVM-пен бірге қолданатын гибриді модель ең жоғары дәлдік (90,2 %) және F1-score (82,52 %) көрсеткіштеріне қол жеткізді. TF-IDF әдісі мәтіндегі терминдердің маңыздылығын бағалауға мүмкіндік береді, ал Word2Vec сөздерді векторлық көрініске айналдырып, семантикалық сәйкестікті жақсартады. SVM алгоритмі деректерді тиімді сыныптарға бөлу үшін гипержазықтықтарды пайдаланады, ал Random Forest шешім ағаштары арқылы жіктеу сапасын арттырады. Сонымен қатар, зерттеу мәтіндерді алдын ала өңдеудің (токенизация, нормализация, стоп-сөздерді жою және лемматизация) маңыздылығын көрсетті. Бұл процестер жіктеу сапасын айтарлықтай жақсартуға ықпал етті. Ұсынылған модель ақпараттық іздеу, тақырыптық модельдеу және құжаттарды автоматты түрде рефераттау сияқты салаларда тиімділігі жоғары құрал ретінде қолданылуы мүмкін. Мұндай гибриді тәсілдер автоматты мәтін талдауының дәлдігі мен сенімділігін арттырып, көптілді ортаға бейімдеу және жаңа деректер көздерін қосу сияқты болашақ зерттеулер үшін перспективаларды ашады. Эксперименттік нәтижелер бұл тәсілдің тональдылықты талдау, құжаттарды санаттау және тақырыптық модельдеу сияқты күрделі тапсырмаларда тиімді жұмыс істейтінін дәлелдеді. Бұл зерттеу нәтижелері автоматты мәтін талдауы саласында жаңа шешімдерді іздеу және оларды тиімді қолдану бойынша маңызды қадам болып табылады.

**Түйін сөздер:** Мәтіндерді автоматты талдау, Машиналық оқыту, Term Frequency-Inverse Document Frequency, Табиғи тілдерді өңдеу, Мәтіндерді қысқаша сипаттау

**Дәйексөздер үшін:** А.Г. Шаушенова, М.Ж. Базарова, Ж.Ж. Ажибекова, К.С. Шадинова, К.С. Бакенова. ӘРТҮРЛІ МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІ АРҚЫЛЫ ҚҰЖАТТАРДЫ АВТОМАТТЫ ТАЛДАУ МОДЕЛІН ҚҰРУ//ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ. 2025. Т. 6. No. 21. 156–



169 бет. (қазақ тілінде). <https://doi.org/10.54309/IJICT.2025.21.1.011>.

**Мүдделер қақтығысы:** Авторлар осы мақалада мүдделер қақтығысы жоқ деп мәлімдемейді.

## СОЗДАНИЕ МОДЕЛИ АВТОМАТИЧЕСКОГО АНАЛИЗА ДОКУМЕНТОВ С ИСПОЛЬЗОВАНИЕМ РАЗЛИЧНЫХ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

**А.Г. Шаушенова<sup>1</sup>, М.Ж. Базарова, Ж.Ж. Ажибекова<sup>3</sup>, К.С. Шадинова<sup>3</sup>,  
К.С. Бакенова<sup>4</sup>**

<sup>1</sup>Казахский агротехнический исследовательский университет имени С. Сейфуллина, Астана, Казахстан;

<sup>2</sup>Восточно-Казахстанский университет имени С. Аманжолова, Усть-Каменогорск, Казахстан;

<sup>3</sup>Казахский Национальный Университет имени С.Д. Асфендиярова, Алматы, Казахстан;

<sup>4</sup>Международный университет Астана, Астана, Казахстан.

E-mail: [madina\\_vkgtu@mail.ru](mailto:madina_vkgtu@mail.ru)

**Шаушенова А.** — Казахский агротехнический исследовательский университет имени С. Сейфуллина, заведующая кафедры «Информационные системы», кандидат технических наук, ассоциированный профессор, Астана, Казахстан

E-mail: [shaushenova\\_78@mail.ru](mailto:shaushenova_78@mail.ru), <https://orcid.org/0000-0002-3164-3688>;

**Базарова М.** — Восточно-Казахстанский университет имени С. Аманжолова, ассоциированный профессор кафедры «Компьютерное моделирование и информационные технологии», PhD, г. Усть-Каменогорск, Казахстан

E-mail: [madina\\_vkgtu@mail.ru](mailto:madina_vkgtu@mail.ru), <https://orcid.org/0000-0003-2580-6580>;

**Ажибекова Ж.** — Кандидат педагогических наук, заведующий кафедрой «Информационно коммуникационные технологии», Казахский Национальный Университет имени С.Д. Асфендиярова, Алматы, Казахстан

E-mail: [Azhibekova.z@kaznmu.kz](mailto:Azhibekova.z@kaznmu.kz), <https://orcid.org/0000-0002-4396-1261>;

**Шадинова К.** — Казахский национальный медицинский университет им. С.Д. Асфендиярова, ассоциированный профессор кафедры информационно-коммуникационных технологий, Алматы, Казахстан

E-mail: [shadinova.ks@mail.ru](mailto:shadinova.ks@mail.ru), <https://orcid.org/0009-0006-5534-7927>;

**Бакенова К.** — Международный университет Астана, преподаватель Педагогического института, докторант кафедры информационной безопасности Евразийского национального университета им. Л.Н. Гумилева, Астана, Казахстан

E-mail: [bakenova.k@bk.ru](mailto:bakenova.k@bk.ru), <https://orcid.org/0009-0004-2567-173X>.

© А.Г. Шаушенова, М.Ж. Базарова, Ж.Ж. Ажибекова, К.С. Шадинова, К.С. Бакенова, 2025

**Аннотация.** В данной статье представлена модель для автоматического анализа документов. Она основана на методе TF-IDF (Term Frequency-Inverse Document Frequency), комбинированном с различными алгоритмами машинного обучения, включая SVM (Support Vector Machine), Random Forest и Word2Vec+SVM. Цель исследования – сравнить эффективность данных методов в задачах классификации текстов и выявить наиболее результативный подход. Экспериментальные результаты показали, что гибридная модель, использующая комбинацию TF-IDF и Word2Vec с SVM, достигла наивысшей точности (90,2 %) и F1-score (82,52 %). Метод TF-IDF позволяет оценивать значимость терминов в тексте, в то время как Word2Vec преобразует слова в векторное представление, улучшая семантическое соответствие. Алгоритм SVM эффективно разделяет данные на классы с помощью гиперплоскостей, а Random Forest улучшает качество классификации за счёт использования ансамбля решающих деревьев. Кроме того, исследование подтвердило важность предварительной обработки текста (токенизация, нормализация, удаление стоп-слов и лемматизация), которая значительно повышает качество классификации. Предложенная модель может быть применена в таких областях, как информационный поиск, тематическое моделирование и автоматическое реферирование документов. Такие гибридные подходы повышают точность и надёжность автоматического анализа текста, открывая перспективы для адаптации в многоязычной среде и добавления новых источников данных. Экспериментальные результаты доказали эффективность данного подхода для решения сложных задач, таких как анализ тональности, категоризация документов и тематическое моделирование. Данное исследование является важным шагом на пути к созданию новых решений в области автоматического анализа текста..

**Ключевые слова:** автоматический анализ документов, машинное обучение, term frequency-inverse document frequency, обработка естественного языка, краткое описание текста

**Для цитирования:** А.Г. Шаушенова, М.Ж. Базарова, Ж.Ж. Ажибекова, К.С. Шадинова, К.С. Бакенова. СОЗДАНИЕ МОДЕЛИ АВТОМАТИЧЕСКОГО АНАЛИЗА ДОКУМЕНТОВ С ИСПОЛЬЗОВАНИЕМ РАЗЛИЧНЫХ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ//МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ. 2025. Т. 6. №. 21. Стр. 156–169. (На каз.). <https://doi.org/10.54309/IJICT.2025.21.1.011>.

**Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

### Кіріспе

Мәтіндерді автоматты түрде талдау (Тұрғынова, 2023; Гао, 2024) қазіргі ақпараттық қоғамда маңызды рөл атқарады, онда деректердің көлемі керемет жылдамдықпен өсуде. Мәтіндік ақпаратты тиімді өңдеу және талдау (Лю, 2023) ақпараттық іздеу (Милинцевич, 2023), тақырыптық модельдеу, құжаттарды автоматты түрде рефераттау (Дель Валье-Кано), 2023 және басқа да салаларда (Асудани, 2023) барған сайын өзекті болып келеді.

Осыған байланысты, мәтіндік деректерді автоматты және дәл жіктеуге қабілетті модельдерді әзірлеу және енгізу ерекше маңызға ие. Мәтіндерді автоматты түрде талдау саласындағы перспективалы бағыттардың бірі (Маттас, 2023) терминдердің кездесу жиілігін ғана емес, сонымен қатар олардың мәтіннің контекстіндегі маңызын ескеруге мүмкіндік беретін семантикалық сәйкестендіру әдістерін қолдану болып табылады.

Term Frequency-Inverse Document Frequency (TF-IDF) әдісі (Мехриш, 2023; Қайбасова, 2022) мәтіндегі терминдердің маңыздылығын бағалауда тиімді құрал ретінде өзін баяғыда дәлелдеген. TF-IDF жиі кездесетін, бірақ ақпараттық мәні аз сөздердің әсерін азайтып, құжаттағы ең маңызды сөздерді анықтауға мүмкіндік береді (Яворский, 2022). Дегенмен, мәтіндерді жіктеу міндеттерінде (Танг, 2023; Хуан, 2024) ең жақсы нәтижелерге қол жеткізу үшін оны қуатты машина оқыту алгоритмдерімен біріктіру орынды. Бұл мақалада TF-IDF-тің Support Vector Machine (SVM), Random Forest және Word2Vec+SVM алгоритмдерімен комбинациялары қарастырылады. SVM әдісі гипержазықтықты құру арқылы деректерді тиімді жіктеуге мүмкіндік береді (Фанк, 2024; Досент, 2020), ал Random Forest шешім ағаштарының ансамблін пайдаланып, жіктеу дәлдігін арттырады. Word2Vec (Word to Vector) сөздерді векторлық көрініске айналдырып, семантикалық сәйкестіктің сапасын едәуір жақсарттады. Үнемі өсіп келе жатқан мәтіндік ақпарат көлемі және оны автомат-тандырылған өңдеудің қажеттілігі жағдайында мәтіндерді семантикалық сәй-кестендіру үшін тиімді модельдерді әзірлеу өте өзекті мәселе болып табылады. Заманауи әдістер, мысалы, TF-IDF, SVM, Random Forest және Word2Vec, мәтін-дерді талдау үшін қуатты құралдар ұсынады, алайда олардың комбинациялары жіктеудің дәлдігі мен тиімділігін едәуір арттыруы мүмкін. Мұндай біріктірілген модельдерді зерттеу және әзірлеу ақпараттық мәтіндерді автоматты түрде тал-даудың жоғары дәлдіктегі жүйелерін құру жолындағы маңызды қадам болып табылады. Осы зерттеу аясында мәтіндерді жіктеу сапасын біріктірілген тәсіл-дерді пайдалану арқылы жақсарту мүмкіндігі қарастырылады.

Осы зерттеудің мақсаты – TF-IDF әдістерін SVM, Random Forest және Word2Vec+SVM алгоритмдерімен біріктіру негізінде құжаттарды автоматты түрде талдау үшін семантикалық сәйкестендіру моделін әзірлеу және бағалау. Осы мақсатқа жету үшін келесі міндеттер қойылды: семантикалық сәйкестендіру және мәтіндерді жіктеудің қолданыстағы әдістеріне шолу жүргізу; TF-IDF және әртүрлі машина оқыту алгоритмдері негізінде біріктірілген модельдерді әзірлеу; ұсынылған модельдердің мәтіндерді жіктеу міндеттеріндегі тиімділігін бағалау үшін бірқатар эксперименттер өткізу; эксперимент нәтижелерін салыстырып, әдістердің ең жақсы комбинациясын анықтау; мәтіндерді автоматты түрде талдау үшін ұсынылған тәсілдердің болашағы туралы қорытынды жасау. Сонымен қатар, осы саланы одан әрі дамыту мәтіндерді талдаудың өнімділігі мен дәлдігін арттыру үшін семантикалық сәйкестендіру әдістерімен машиналық



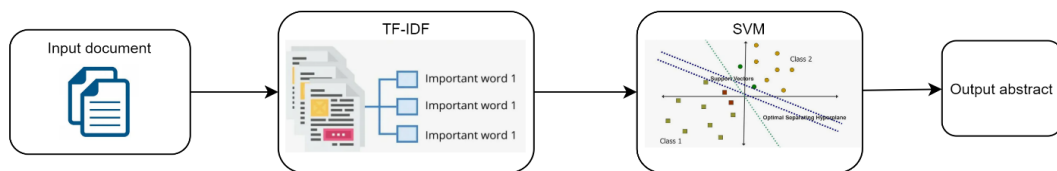
оқытудың басқа алгоритмдерін және олардың комбинацияларын зерттеуді қамтуы мүмкін. Әлеуметтік желілер мен жаңалықтар сияқты жаңа деректер көздерін қосу, сондай-ақ модельдерді көптілді ортаға бейімдеу ұсынылған тәсілдің қолдану аясын едәуір кеңейте алады. Бұл жақын болашақта мәтіндерді автоматты түрде талдауға арналған неғұрлым әмбебап және тиімді құралдар жасауға үміт береді.

### Әдістер мен материалдар

Осы жұмыста құжаттарды автоматты түрде талдау үшін семантикалық сәйкестендіру моделін жасау және бағалау үшін қолданылған әдістер мен алгоритмдер сипатталған. Негізгі назар TF-IDF (Term Frequency-Inverse Document Frequency) әдісін машина оқыту алгоритмдерімен, атап айтқанда SVM (Support Vector Machine), Random Forest және Word2Vec+SVM комбинацияларына аударылады. Деректерді алдын ала өңдеу, модельдерді құру және эксперименттерді жүргізу процестері сипатталады. Модельдерді құруды бастамас бұрын барлық мәтіндік деректер бірнеше кезеңді қамтитын алдын ала өңдеуден өтті. Алдымен мәтінді жеке сөздерге (токендерге) бөлу, яғни токенизация орындалды. Содан кейін барлық сөздер жоғарғы және төменгі регистрдегі айырмашылықтарды жою үшін төменгі регистрге келтірілетін нормализация жүргізілді. Бұдан кейін жиі кездесетін, бірақ маңызды ақпаратты қамтымайтын сөздерді, мысалы, шылаулар мен жалғаулықтарды мәтіннен алып тастау арқылы стоп-сөздерді жою жүзеге асырылды. Өңдеу лемматизациямен аяқталды, онда сөздер бірегей токендердің санын азайту үшін олардың бастапқы түріне (лемма) келтірілді. TF-IDF әдісі мәтіндерді сандық векторлық көріністерге айналдыру үшін қолданылды. TF-IDF әрбір терминнің құжаттағы маңыздылығын бүкіл құжаттар корпусы бойынша бағалауға мүмкіндік береді. TF-IDF формуласы келесідей ұсынылған (1):

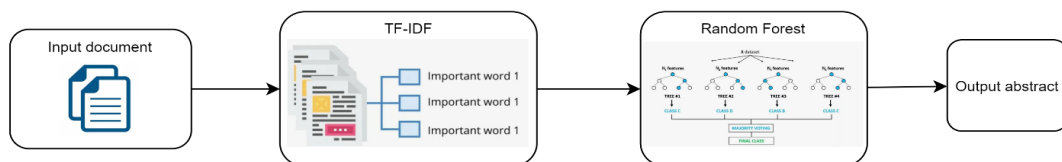
$$TF - IDF(t, d) = TF(t, d) * IDF(t) \quad (1)$$

$TF(t, d)$  –  $t$  терминінің  $d$  құжаттағы жиілігі, ал  $IDF(t)$  –  $t$  терминін қамтитын құжаттардың кері жиілігі. Осылайша, TF-IDF әрбір құжат үшін ең маңызды терминдерді бөліп көрсетуге мүмкіндік береді. Support Vector Machine (SVM) — бұл классификация және регрессия міндеттеріне арналған машина оқыту әдісі. Ол жоғары өлшемді кеңістікте гипержазықтық немесе бірнеше гипержазықтықтар құрастырады, оларды классификация, регрессия немесе басқа міндеттер үшін қолдануға болады. Осы жұмыста SVM мәтіндерді TF-IDF көрінісі негізінде жіктеу үшін пайдаланылады (1-сурет).



Сур. 1. TF-IDF+SVM моделінің архитектурасы

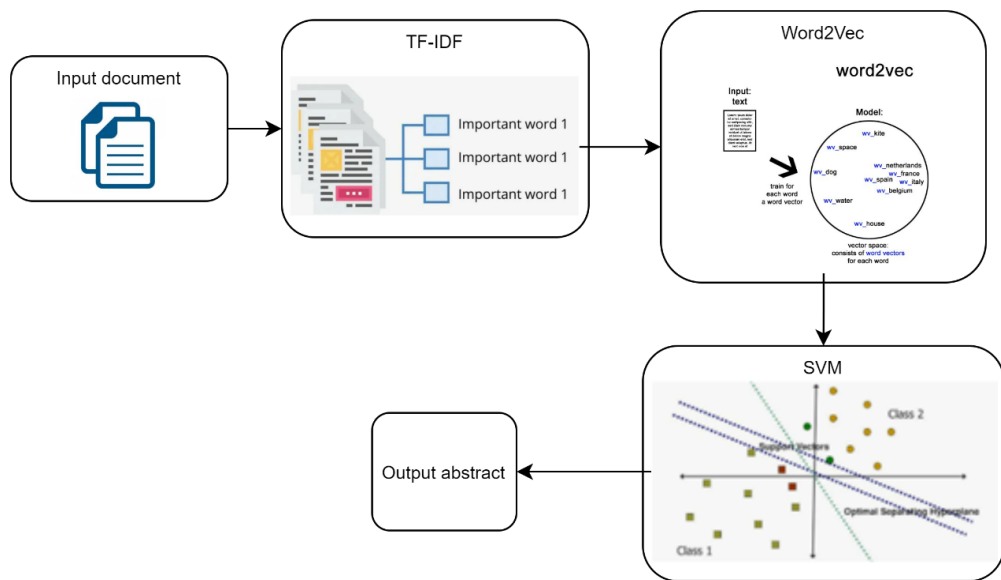
Random Forest алгоритмі - шешімдер ағаштарының ансамблі болып табылады, мұнда әрбіркласификатор деректердің кездейсоқ қосалқы жиынында құрылады. Бұл әдіс модельдің шамадан тыс үйренуге тұрақтылығын арттырып, жалпы дәлдігін жақсартуға мүмкіндік береді. Осы жұмыста Random Forest мәтіндерді олардың TF-IDF көріністері негізінде жіктеу үшін қолданылды (2-сурет).



Сур. 2. TF-IDF + Random Forest моделінің архитектурасы

Осы зерттеуде Word2Vec+SVM моделін жасау үшін контекстік сөздерді ұсынуды қамтамасыз ететін Word2Vec әдісі қолданылды. Атап айтқанда, Word2Vec моделін оқыту үшін Skip-gram архитектурасы тандалды, себебі ол сөздердің контекстін тиімді болжайды. Бұл әдіс сөздерді жоғары өлшемді векторларға айналдыруға мүмкіндік береді, мұнда әрбір сөз белгілі бір ұзындықтағы сандық вектор түрінде ұсынылады. Корпустағы әрбір сөз үшін Word2Vec моделін оқытқаннан кейін, келесі қадам әрбір құжат үшін векторлық көрініс алуды қамтыды. Бұл үшін құжаттағы барлық сөз векторлары орташаланды, осылайша құжатты толығымен бейнелейтін біртұтас вектор құрылды. Сөз векторларын орташалау құжаттағы барлық сөздердің семантикалық ақпаратын ескеретін мәтіннің жалпыланған көрінісін алуға мүмкіндік береді. Алынған құжат векторлары SVM (Support Vector Machine) алгоритміне кіріс деректері ретінде пайдаланылды. SVM деректерді тиімді жіктеу қабілеті арқасында тандалды, өйткені ол сыныптарды барынша бөлуге мүмкіндік беретін гипержазықтықты құрайды. SVM моделін оқыту үшін алдыңғы қадамда алынған құжаттардың векторлық көріністері пайдаланылды. SVM моделін оқыту процесі классификация үшін оңтайлы параметрлерді табу үшін кросс-валидацияны пайдалана отырып, гиперпараметрлерді баптауды қамтыды. SVM моделін баптап, оны оқыту жиынтығында оқытқаннан кейін

модель өнімділігін бағалау үшін тест жиынтығында тексерілді. Бағалау үшін негізгі көрсеткіштер дәлдік (accuracy) және F1-score болды, бұл модельдің классификация сапасын өлшеуге мүмкіндік берді. Осылайша, Word2Vec және SVM әдістерінің үйлесімі сөздердің контекстік векторлық көріністерін пайдаланып, семантикалық сәйкестендіру мен мәтіндерді жіктеу сапасын жақсартуға мүмкіндік берді. Эксперименттердің нәтижелері Word2Vec+SVM моделі құжаттарды автоматты түрде талдау міндеттерінде жоғары дәлдік пен тиімділікті қамтамасыз ететінін көрсетті (3-сурет).



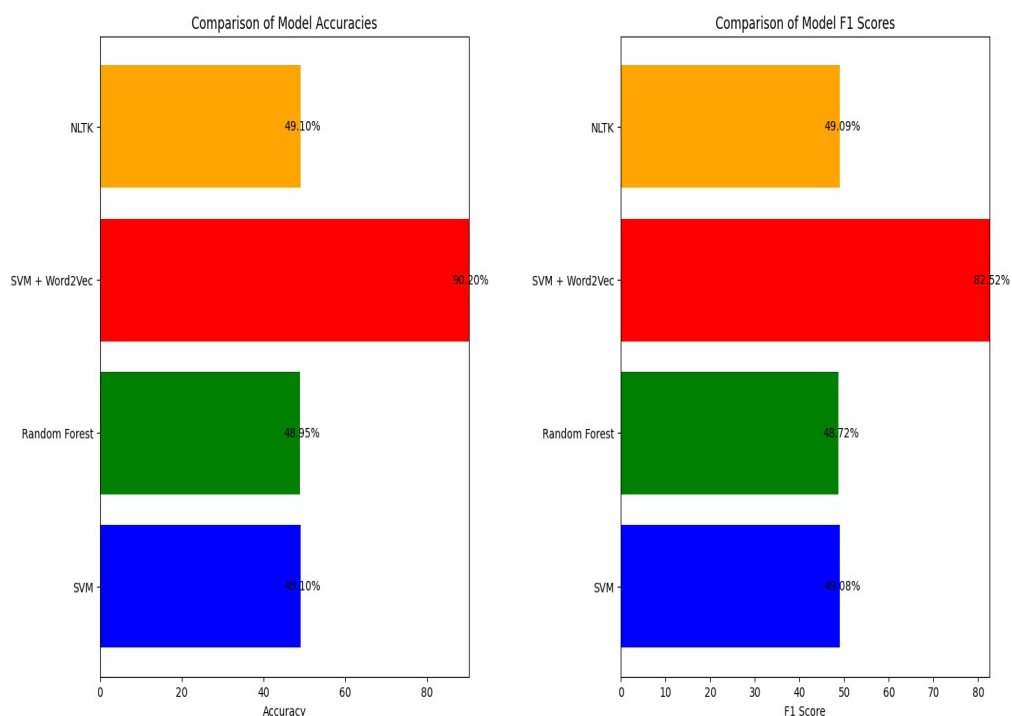
Сур. 3. Word2Vec+ SVM моделінің архитектурасы

Өткізілген эксперименттер барысында TF-IDF және Word2Vec сияқты біріктірілген тәсілдерді SVM және Random Forest сияқты машина оқыту алгоритмдерімен қолдану мәтіндерді жіктеудің дәлдігі мен тиімділігін едәуір арттыруға мүмкіндік беретіні анықталды. Бұл нәтижелер осындай модельдерді мәтіндік ақпаратты автоматты түрде талдау үшін одан әрі зерттеу мен әзірлеудің маңыздылығын растайды.

### Нәтижелер және оларды талқылау

Зерттеу барысында TF-IDF әдістері мен SVM, Random Forest және Word2Vec+SVM сияқты машина оқыту алгоритмдеріне негізделген семантикалық сәйкестендіру модельдерінің өнімділігіне салыстырмалы талдау жүргізілді. Экспериментте қолданылған модельдердің бірі TF-IDF пен SVM комбинациясы болды. Бұл тәсіл мәтіндерді жіктеу міндеттерінде 49,10 % дәлдікпен және 49,08 % F1 көрсеткішімен жеткілікті тиімді екенін көрсетті. TF-IDF+SVM моделінің негізгі артықшылығы - оның сызықтық бөлінетін

деректермен жұмыс істеу қабілеті, бұл гиперпараметрлерді дұрыс реттеген жағдайда мәтін сыныптарын дәл анықтауға мүмкіндік береді. Дегенмен, TF-IDF+SVM моделінің бірқатар шектеулері бар. Негізгі мәселелердің бірі — үлкен деректер жиынтығын масштабтау мүмкіндігі. Деректер көлемі ұлғайған сайын, модельдің өнімділігі төмендей бастайды, бұл SVM алгоритмінің үлкен көлемдегі мәтіндік ақпаратты өңдеудегі шектеулерімен байланысты. Сонымен қатар, гиперпараметрлерді таңдауға сезімталдығы оңтайлы нәтижелерге қол жеткізу үшін мұқият баптауды талап етеді. Бұл әсіресе деректер сипаттамалары айтарлықтай өзгеруі мүмкін әртүрлі мәтіндік корпустармен жұмыс істегенде уақытты қажет ететін процесс болуы мүмкін. Модель негізделген SVM алгоритмі спам-фильтрация, құжаттарды санаттау және тональдылықты талдау сияқты мәтіндік жіктеу міндеттерінде кеңінен қолданылады. Ол деректер сыныптарын тиімді бөлетін гипержазықтықтарды құруға мүмкіндік береді, жоғары жіктеу дәлдігін қамтамасыз етеді. Алайда, масштабтаудың және параметрлерге сезімталдықтың кемшіліктері үлкен деректер жағдайында оның қолданылуын шектеуі мүмкін. Сондықтан, әсіресе деректер көлемі үлкен және жіктеу дәлдігі жоғары болған жағдайда, мәтіндерді автоматты түрде талдау үшін модельді таңдағанда осы мүмкіндіктерді ескеру маңызды (4-сурет).



Сур. 4. Үлгі бойынша дәлдік көрсеткіштері

TF-IDF-ті қолдану арқылы Random Forest алгоритмі SVM-мен салыстырғанда сәл жақсы нәтижелер көрсетті, оның дәлдігі 49,40 % және F1 Score 49,34 % болды. Random Forest жоғары өлшемді деректерді жақсы өңдейді және күрделі сызықтық емес тәуелділіктерді модельдей алады, сондай-ақ шешім ағаштары ансамблін қолдану арқылы шамадан тыс үйренуге төзімді. Дегенмен, бұл модельді оқыту және болжау, әсіресе үлкен деректер жиынтығында, баяу және ресурсты қажет етуі мүмкін. Бұл алгоритм мәтіндерді жіктеу және ауытқуларды анықтау сияқты классификация және регрессия міндеттерінде қолданылады. Ең жоғары өнімділікті TF-IDF және Word2Vec комбинациясын SVM-мен бірге қолданған модель көрсетті. Бұл гибриді модель 90,20 % дәлдікке және 82,52 % F1 Score-ға қол жеткізді. Бұл модель сөздер арасындағы семантикалық байланыстарды түсіреді, бұл болжамдардың сапасын айтарлықтай жақсарттады және SVM мен терең оқытудың артықшылықтарын біріктіреді. Дегенмен, Word2Vec және SVM моделін оқыту үшін айтарлықтай есептеу ресурстары, сондай-ақ нәтижелерді баптау және түсіндірудегі қиындықтар қажет. Бұл модель, әсіресе, мәтінді терең семантикалық өңдеуді қажет ететін тапсырмаларға, мысалы, тональдылықты талдау, сөз тіркестерін анықтау және тақырыптық модельдеу үшін өте қолайлы.

Әртүрлі модельдердің өнімділігін талдау TF-IDF және Word2Vec алгоритмін SVM алгоритмімен біріктіру сияқты әдістерді біріктіру мәтіндерді жіктеу нәтижелерін айтарлықтай жақсартатынын көрсетті. Мұндай әдістердің комбинациясы TF-IDF сияқты терминдердің жиілігін ескеріп қана қоймай, Word2Vec көмегімен сөздер арасындағы контекстік байланыстарды тереңірек талдауға мүмкіндік береді. Word2Vec көмегімен алынған сөздердің векторлық көріністері сөздер арасындағы жасырын семантикалық байланыстарды түсіре алады, бұл модельге мәтіннің мазмұнын дәлірек анықтауға және, тиісінше, жіктеу дәлдігін арттыруға көмектеседі. Бұл мүмкіндіктер біріктірілген модельдерді мәтіндік деректерді терең талдауды қажет ететін тапсырмалар үшін қуатты құралға айналдырады. Зерттеу нәтижелері мәтіндерді жіктеу сапасын жақсарту үшін сөздер арасындағы контекстік байланыстарды ескеретін әдістерді қолданудың маңыздылығын атап көрсетеді. TF-IDF сияқты дәстүрлі әдістер негізгі талдау үшін тиімді болғанымен, олар мәтіннің семантикалық мазмұнына байланысты маңызды аспектілерді елемейді мүмкін. TF-IDF пен Word2Vec-ті біріктіру бұл олқылықтың орнын толтырып, мәтіннің толығырақ көрінісін қамтамасыз етеді, бұл әсіресе тональдылықты талдау, тақырыптық модельдеу және құжаттарды автоматты түрде рефераттау сияқты күрделі тапсырмалар үшін маңызды. Контекстік талдауды қосу мәтіндердің мазмұнын жақсы түсінуге және оларды жіктеу үшін қолданылатын машина оқыту модельдерінің тиімділігін арттыруға көмектеседі. Бұл саладағы одан әрі зерттеулердің перспективалары әртүрлі машина оқыту алгоритмдерімен мәтінді векторизациялаудың басқа әдістерін біріктіру мүмкіндіктерін зерттеуді қамтиды. Мысалы, BERT немесе GPT сияқты жетілдірілген әдістерді дәстүрлі

классификация алгоритмдерімен бірге қолдану одан да дәл және әмбебап модельдерге әкелуі мүмкін. Сонымен қатар, мұндай модельдерді көптілді ортаға бейімдеу және әлеуметтік желілер мен жаңалықтар сияқты жаңа деректер көздерін қосу олардың қолдану аясын едәуір кеңейте алады. Бұл жақын болашақта мәтіндерді автоматты түрде талдауға арналған неғұрлым әмбебап және тиімді құралдарды жасауға мүмкіндік береді, бұл ақпараттық мәтінді талдаудың жоғары дәлдіктегі жүйелерін құру жолындағы маңызды қадам болып табылады.

### **Қорытынды**

Осы зерттеу аясында TF-IDF әдістерін әртүрлі машина оқыту алгоритмдерімен: SVM, Random Forest және Word2Vec+SVM біріктіру негізінде құжаттарды автоматты түрде талдау үшін семантикалық сәйкестендіру моделі әзірленіп, бағаланды. Негізгі мақсат – мәтіндерді жіктеу міндеттерінде осы әдістердің тиімділігін салыстырып, ең нәтижелі тәсілді анықтау болды. Эксперименттердің нәтижелері ең жоғары өнімділікті TF-IDF және Word2Vec комбинациясын SVM-мен бірге қолданатын модель көрсеткенін көрсетті. Бұл гибриді модель 90,20 % дәлдікке және 82,52 % F1 Score-ға қол жеткізді, бұл қарастырылған басқа модельдердің нәтижелерінен айтарлықтай жоғары. TF-IDF әдісі мәтіндегі ең маңызды терминдерді анықтауға мүмкіндік берді, ал Word2Vec сөздердің контекстік көрінісін қамтамасыз етті, бұл семантикалық сәйкестіктің және, тиісінше, мәтіндерді жіктеудің сапасын жақсартты. TF-IDF+SVM және Random Forest модельдері де шамамен 49% дәлдікпен қанағаттанарлық нәтижелер көрсетті, бірақ олар тиімділік жағынан Word2Vec+SVM моделінен төмен болды. Бұл сөздер арасындағы семантикалық байланыстарды ескерудің және жіктеу дәлдігін арттыру үшін мәтінді тереңірек талдау әдістерін пайдаланудың маңыздылығын растайды. TF-IDF және Word2Vec әдістерін машина оқыту алгоритмдерімен біріктіру мәтіндерді автоматты түрде талдаудың дәлдігі мен сенімділігін арттырудың жаңа мүмкіндіктерін ашады.

Зерттеу барысында деректерді алдын ала өңдеудің, соның ішінде токенизация, нормализация, стоп-сөздерді жою және лемматизация сияқты қадамдарды қамтитын мұқият өңдеудің жіктеу нәтижелерін жақсартуда маңызды кезең екені анықталды. Бұл мәтіндік деректерді өңдеуге кешенді көзқарастың қажеттілігін, соның ішінде алдын ала өңдеуді және заманауи машина оқыту алгоритмдерін қолдануды көрсетеді. Болашақ зерттеулердің перспективалары мәтіндерді векторизациялаудың басқа әдістерін әртүрлі машина оқыту алгоритмдерімен біріктіру мүмкіндіктерін зерттеуді қамтиды. Жаңа дереккөздерді, мысалы, әлеуметтік желілер мен жаңалықтар сияқты мәліметтерді қосу, сондай-ақ модельдерді көптілді ортаға бейімдеу ұсынылған тәсілдің қолдану аясын едәуір кеңейте алады. Бұл мәтіндерді автоматты түрде талдауға арналған әмбебап және тиімді құралдарды жасауға мүмкіндік береді, бұл ақпараттық мәтінді талдаудың жоғары дәлдіктегі жүйелерін құру жолындағы маңызды қадам болып табылады. Осылайша, осы зерттеудің



нәтижелері мәтіндерді автоматты түрде өңдеуге арналған қуатты құралдарды жасау үшін заманауи семантикалық сәйкестендіру және машина оқыту әдістерін қолдану әлеуетін көрсетеді. Мұндай модельдерді әзірлеу және енгізу ақпараттық іздеу, тақырыптық модельдеу және құжаттарды автоматты түрде рефераттау сияқты әртүрлі салаларда мәтіндік деректерді талдау сапасы мен тиімділігін айтарлықтай жақсарты алады.

## ӘДЕБИЕТТЕР

- Асудани Д.С., Нагвани Н.К., Сингх П. (2023). Терең оқу ортасындағы мәтіндік аналитикаға сөздерді енгізу үлгілерінің әсері: шолу // Жасанды интеллект шолуы. — 2023. — Т. 56. — № 9. — Б. 10345–10425.
- дель Валле-Кано Г., Кижано-Санчес Л., Либерагоре Ф., Гомес Дж. (2023). SocialHaterBERT: Мәтіндік талдау және пайдаланушы профильдері арқылы Twitter-дегі өшпенділік сөзін автоматты түрде анықтауға арналған дихотомиялық тәсіл // Қолданбалары бар сараптамалық жүйелер. — 2023. — Т. 216. — 119446-6.
- Доцент Л.Н. (2020). Интеллектуалдық жүйені қалыптастыратын білім беру бағдарламалары үшін силлабустарды талдаудың әдістері мен алгоритмдері // Теориялық және қолданбалы ақпараттық технологиялар журналы. — 2020. — Т. 98. — № 05. — Б. 876–888.
- Funck A.S., Lau R.R. (2024). Эмоциялардың саяси ақпаратты іздеуге және шешім қабылдауға әсерін мета-аналитикалық бағалау // Американдық саясаттану журналы. — 2024. — Т. 68, № 3. — Б. 891–906.
- Гао Р., Мерцдорф Х.Е., Анвар С., Хипвелл М.С., Сриниваса А. (2024). Орта білімнен кейінгі білім берудегі мәтіндік жауаптарды автоматты түрде бағалау: Жүйелі шолу // Компьютерлер және білім: Жасанды интеллект. — 2024. — Т. 100206.
- Хуан З., Юань Л. (2024). Денсаулық сақтау саласындағы білімді басқару жүйесіндегі концепциялық семантикамен оқуды және зерттеуді іздеуді жақсарту: интерактивті білімді визуализациялау тәсілі // Қолданбалары бар сараптамалық жүйелер. — 2024. — Т. 237. — 121558 6.
- Қайбасова Д., Нұртай М. (2022). Мәтіндік академиялық жұмыстардың сапасын бағалау үшін машиналық оқыту үлгілерінің салыстырмалы талдауы // 2022 Халықаралық Smart ақпараттық жүйелер мен технологиялар конференциясы (SIST), IEEE. — 2022. — Б. 1–4.
- Лю Х., Чжоу Г., Конг М., Инь З., Ли Х., Ин Л., Чжэн В. (2023). Twitter қысқа мәтіндерінің көп таңбалы корпусын дамыту: жартылай автоматты әдіс // Жүйелер. — 2023. — Т. 11, No 8. — 390-6.
- Mattas P.S. (2023). ChatGPT: AI тілін өңдеуді және оның салдарын зерттеу // Журналдың басты беті: www.ijrpr.com, ISSN. — 2023. — Т. 2582. — № 7421. — Б. 7–8.
- Мехриш А., Маджумдер Н., Бхарадвадж Р., Михалча Р., Пория С. (2023). Сөйлеуді өңдеуге арналған терең оқыту әдістеріне шолу // Ақпаратты біріктіру. — 2023. — Т. 99. — 101869 6.
- Милинцевич К., Сирц К., Диас Г. (2023). Симптомдарды болжау арқылы депрессияны автоматты түрде мәтінге негізделген бағалауға қарай // Ми информатикасы. — 2023. — Т. 10, No 1. — 4-6.
- Тан З. және т.б. (2023). Әмбебап құжаттарды өңдеуге арналған көзқарасты, мәтінді және макетті біріктіру // Компьютерлік көру және үлгіні тану бойынша IEEE/CVF конференциясының материалдары. — 2023. — Б. 19254–19264 жж.
- Toselli A. H., Puigcerver J., Vidal E. (2024). Қолжазба мәтіндік кескіндердің үлкен жинақтарында ақпаратты іздеу және іздеу үшін ықтималдық индекстеу // Springer Nature. — 2024. — Т. 49.
- Тұрғынова Н., Тұрғынов Б., Умаралиев Ж. (2023). Мәтінді автоматты талдау: синтаксис және семантикалық талдау // Инженерлік мәселелер және инновациялар. — 2023 жыл.
- Яворский В., Кайбасова Д., Ключева Ю. (2022). Студенттердің білім беру жетістіктеріне арналған сақтау ортасын талдау бойынша шараларды әзірлеу мәселелері // 2022 IEEE 7-ші халықаралық энергетикалық конференциясы (ENERGYCON), IEEE. — 2022. — Б. 1–6.

## REFERENCES

- Asudani D.S., Nagwani N.K., Singh P. (2023). Impact of word embedding models on text analytics in deep learning environment: a review // *Artificial Intelligence Review*. — 2023. — Vol. 56, No. 9. — Pp. 10345–10425.
- del Valle-Cano G., Quijano-Sánchez L., Liberatore F., Gómez J. (2023). SocialHaterBERT: A dichotomous approach for automatically detecting hate speech on Twitter through textual analysis and user profiles // *Expert Systems with Applications*. — 2023. — Vol. 216. — P. 119446.
- Docent L.N. (2020). Methods and algorithms of analyzing syllabuses for educational programs forming intellectual system // *Journal of Theoretical and Applied Information Technology*. — 2020. — Vol. 98, No. 05. — Pp. 876–888.
- Funck A.S., Lau R.R. (2024). A Meta-Analytic Assessment of the Effects of Emotions on Political Information Search and Decision-Making // *American Journal of Political Science*. — 2024. — Vol. 68, No. 3. — Pp. 891–906.
- Gao R., Merzdorf H. E., Anwar S., Hipwell M.C., Srinivasa A. (2024). Automatic assessment of text-based responses in post-secondary education: A systematic review // *Computers and Education: Artificial Intelligence*. — 2024. — Vol. 100206.
- Huang Z., Yuan L. (2024). Enhancing learning and exploratory search with concept semantics in online healthcare knowledge management systems: An interactive knowledge visualization approach // *Expert Systems with Applications*. — 2024. — Vol. 237. — P. 121558.
- Kaibassova D., Nurtay M. (2022). The Comparative Analysis of Machine Learning Models for Quality Assessment of Textual Academic Works // *2022 International Conference on Smart Information Systems and Technologies (SIST)*, IEEE. — 2022. — Pp. 1–4.
- Liu X., Zhou G., Kong M., Yin Z., Li X., Yin L., Zheng W. Developing multi-labelled corpus of twitter short texts: a semi-automatic method // *Systems*. — 2023. — Vol. 11, No. 8. — P. 390.
- Mattas P.S. (2023). ChatGPT: A study of AI language processing and its implications // *Journal homepage: www.ijrpr.com*, ISSN. — 2023. — Vol. 2582, No. 7421. — Pp. 7–8.
- Mehrish A., Majumder N., Bharadwaj R., Mihalcea R., Poria S. (2023). A review of deep learning techniques for speech processing // *Information Fusion*. — 2023. — Vol. 99. — P. 101869.
- Milintsevich K., Sirts K., Dias G. (2023). Towards automatic text-based estimation of depression through symptom prediction // *Brain Informatics*. — 2023. — Vol. 10, No. 1. — P. 4.
- Tang Z. et al. (2023). Unifying vision, text, and layout for universal document processing // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. — 2023. — Pp. 19254–19264.
- Toselli A. H., Puigcerver J., Vidal E. (2024). Probabilistic indexing for information search and retrieval in large collections of handwritten text images // *Springer Nature*. — 2024. — Vol. 49.
- Turgunova N., Turgunov B., Umaraliyev J. (2023). Automatic text analysis: syntax and semantic analysis // *Engineering Problems and Innovations*. — 2023.
- Yavorskiy V., Kaibassova D., Klyuyeva Y. Issues (2022). Of Developing Measures To Analyze Storage Medium For Educational Achievements Of Students // *2022 IEEE 7th International Energy Conference (ENERGY-CON)*, IEEE. — 2022. — Pp. 1–6.



**ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ  
КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ**

**МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И  
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ**

**INTERNATIONAL JOURNAL OF INFORMATION AND  
COMMUNICATION TECHNOLOGIES**

Правила оформления статьи для публикации в журнале на сайте:

**<https://journal.iitu.edu.kz>**

ISSN 2708–2032 (print)

ISSN 2708–2040 (online)

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

**ОТВЕТСТВЕННЫЙ РЕДАКТОР**

Мрзабаева Раушан Жалиқызы

**НАУЧНЫЙ РЕДАКТОР**

Ермакова Вера Александровна

**ТЕХНИЧЕСКИЙ РЕДАКТОР**

Рашидинов Дамир Рашидинович

**КОМПЬЮТЕРНАЯ ВЕРСТКА**

Асанова Жадыра

Подписано в печать 15.03.2025.

Формат 60x881/8. Бумага офсетная. Печать - ризограф. 9,0 п.л. Тираж 100  
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).