

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



«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ  
«ХАЛЫҚ» ЖҚ

# Х А Б А Р Л А Р Ы

**ИЗВЕСТИЯ**

РОО «НАЦИОНАЛЬНОЙ  
АКАДЕМИИ НАУК РЕСПУБЛИКИ  
КАЗАХСТАН»  
ЧФ «Халық»

**N E W S**

OF THE ACADEMY OF SCIENCES  
OF THE REPUBLIC OF  
KAZAKHSTAN  
«Halyk» Private Foundation

**SERIES  
PHYSICS AND INFORMATION TECHNOLOGY**

**2 (350)**

**APRIL – JUNE 2024**

PUBLISHED SINCE JANUARY 1963  
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK



## ЧФ «ХАЛЫҚ»

В 2016 году для развития и улучшения качества жизни казахстанцев был создан частный Благотворительный фонд «Халык». За годы своей деятельности на реализацию благотворительных проектов в областях образования и науки, социальной защиты, культуры, здравоохранения и спорта, Фонд выделил более 45 миллиардов тенге.

Особое внимание Благотворительный фонд «Халык» уделяет образовательным программам, считая это направление одним из ключевых в своей деятельности. Оказывая поддержку отечественному образованию, Фонд вносит свой посильный вклад в развитие качественного образования в Казахстане. Тем самым способствуя росту числа людей, способных менять жизнь в стране к лучшему – профессионалов в различных сферах, потенциальных лидеров и «великих умов». Одной из значимых инициатив фонда «Халык» в образовательной сфере стал проект *Ozgeris powered by Halyk Fund* – первый в стране бизнес-инкубатор для учащихся 9-11 классов, который помогает развивать необходимые в современном мире предпринимательские навыки. Так, на содействие малому бизнесу школьников было выделено более 200 грантов. Для поддержки талантливых и мотивированных детей Фонд неоднократно выделял гранты на обучение в Международной школе «Мирас» и в *Astana IT University*, а также помог казахстанским школьникам принять участие в престижном конкурсе «*USTEM Robotics*» в США. Авторские работы в рамках проекта «Тәлімгер», которому Фонд оказал поддержку, легли в основу учебной программы, учебников и учебно-методических книг по предмету «Основы предпринимательства и бизнеса», преподаваемого в 10-11 классах казахстанских школ и колледжей.

Помимо помощи школьникам, учащимся колледжей и студентам Фонд считает важным внести свой вклад в повышение квалификации педагогов, совершенствование их знаний и навыков, поскольку именно они являются проводниками знаний будущих поколений казахстанцев. При поддержке Фонда «Халык» в южной столице был организован ежегодный городской конкурс педагогов «*Almaty Digital Ustaz*».

Важной инициативой стал реализуемый проект по обучению основам финансовой грамотности преподавателей из восьми областей Казахстана, что должно оказать существенное влияние на воспитание финансовой грамотности и предпринимательского мышления у нового поколения граждан страны.

Необходимую помощь Фонд «Халык» оказывает и тем, кто особенно остро в ней нуждается. В рамках социальной защиты населения активно проводится работа по поддержке детей, оставшихся без родителей, детей и взрослых из социально уязвимых слоев населения, людей с ограниченными возможностями, а также обеспечению нуждающихся социальным жильем, строительству социально важных объектов, таких как детские сады, детские площадки и физкультурно-оздоровительные комплексы.

В копилку добрых дел Фонда «Халык» можно добавить оказание помощи детскому спорту, куда относится поддержка в развитии детского футбола и карате в нашей стране. Жизненно важную помощь Благотворительный фонд «Халык» оказал нашим соотечественникам во время недавней пандемии COVID-19. Тогда, в разгар тяжелой борьбы с коронавирусной инфекцией Фонд выделил свыше 11 миллиардов тенге на приобретение необходимого медицинского оборудования и дорогостоящих медицинских препаратов, автомобилей скорой медицинской помощи и средств защиты, адресную материальную помощь социально уязвимым слоям населения и денежные выплаты медицинским работникам.

В 2023 году наряду с другими проектами, нацеленными на повышение благосостояния казахстанских граждан Фонд решил уделить особое внимание науке, поскольку она является частью общественной культуры, а уровень ее развития определяет уровень развития государства.

Поддержка Фондом выпуска журналов Национальной Академии наук Республики Казахстан, которые входят в международные фонды Scopus и Wos и в которых публикуются статьи отечественных ученых, докторантов и магистрантов, а также научных сотрудников высших учебных заведений и научно-исследовательских институтов нашей страны является не менее значимым вкладом Фонда в развитие казахстанского общества.

**С уважением,  
Благотворительный Фонд «Халык»!**

#### **БАС РЕДАКТОР:**

**МУТАНОВ Ғалымқайыр Мұтанұлы**, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), **Н=5**

#### **БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:**

**МАМЫРБАЕВ Өркен Жұмажанұлы**, ақпараттық жүйелер мамандығы бойынша философия докторы (Ph.D), ҚР БҒМ Ғылым комитеті «Ақпараттық және есептеуші технологиялар институты» РМК жауапты хатшысы (Алматы, Қазақстан), **Н=5**

#### **РЕДАКЦИЯ АЛҚАСЫ:**

**ҚАЛИМОЛДАЕВ Мақсат Нұрәділұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі (Алматы, Қазақстан), **Н=7**

**БАЙГУНЧЕКОВ Жұмаділ Жанабайұлы**, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Сатпаев университетінің Қолданбалы механика және инженерлік графика кафедрасы, (Алматы, Қазақстан), **Н=3**

**ВОЙЧИК Вальдемар**, техника ғылымдарының докторы (физика), Люблин технологиялық университетінің профессоры (Люблин, Польша), **Н=23**

**БОШКАЕВ Қуантай Авғазыұлы**, Ph.D. Теориялық және ядролық физика кафедрасының доценті, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=10**

**QUEVEDO Nemando**, профессор, Ядролық ғылымдар институты (Мехико, Мексика), **Н=28**

**ЖҮСІПОВ Марат Абжанұлы**, физика-математика ғылымдарының докторы, теориялық және ядролық физика кафедрасының профессоры, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=7**

**КОВАЛЕВ Александр Михайлович**, физика-математика ғылымдарының докторы, Украина ҰҒА академигі, Қолданбалы математика және механика институты (Донецк, Украина), **Н=5**

**РАМАЗАНОВ Тілекқабұл Сәбитұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университетінің ғылыми-инновациялық қызмет жөніндегі проректоры, (Алматы, Қазақстан), **Н=26**

**ТАКИБАЕВ Нұрғали Жабағаұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=5**

**ТИГИНЯНУ Ион Михайлович**, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), **Н=42**

**ХАРИН Станислав Николаевич**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Қазақстан-Британ техникалық университеті (Алматы, Қазақстан), **Н=10**

**ДАВЛЕТОВ Асқар Ербуланович**, физика-математика ғылымдарының докторы, профессор, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=12**

**КАЛАНДРА Пьетро**, Ph.D (физика), Нанокұрылымды материалдарды зерттеу институтының профессоры (Рим, Италия), **Н=26**

**«ҚР ҰҒА Хабарлары. Физика және информатика сериясы».**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **№ 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *физика және ақпараттық коммуникациялық технологиялар сериясы*. Қазіргі уақытта: *«ақпараттық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*  
*http://www.physico-mathematical.kz/index.php/en/*

## ГЛАВНЫЙ РЕДАКТОР:

**МУТАНОВ Галимжаир Мутанович**, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МОН РК (Алматы, Казахстан), **Н=5**

## ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

**МАМЫРБАЕВ Оркен Жумажанович**, доктор философии (PhD) по специальности Информационные системы, ответственный секретарь РГП «Института информационных и вычислительных технологий» Комитета науки МОН РК (Алматы, Казахстан), **Н=5**

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

**КАЛИМОЛДАЕВ Максат Нурадилович**, доктор физико-математических наук, профессор, академик НАН РК (Алматы, Казахстан), **Н=7**

**БАЙГУНЧЕКОВ Жумадил Жанабаевич**, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сагпаева (Алматы, Казахстан), **Н=3**

**ВОЙЧИК Вальдемар**, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), **Н=23**

**БОШКАЕВ Куантай Авгазыевич**, доктор Ph.D, преподаватель, доцент кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=10**

**QUEVEDO Hemando**, профессор, Национальный автономный университет Мексики (UNAM), Институт ядерных наук (Мехико, Мексика), **Н=28**

**ЖУСУПОВ Марат Абжанович**, доктор физико-математических наук, профессор кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=7**

**КОВАЛЕВ Александр Михайлович**, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), **Н=5**

**РАМАЗАНОВ Тлексабул Сабитович**, доктор физико-математических наук, профессор, академик НАН РК, проректор по научно-инновационной деятельности, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=26**

**ТАКИБАЕВ Нурғали Жабағевич**, доктор физико-математических наук, профессор, академик НАН РК, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=5**

**ТИГИНЯНУ Ион Михайлович**, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), **Н=42**

**ХАРИН Станислав Николаевич**, доктор физико-математических наук, профессор, академик НАН РК, Казахстанско-Британский технический университет (Алматы, Казахстан), **Н=10**

**ДАВЛЕТОВ Аскар Ербуланович**, доктор физико-математических наук, профессор, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=12**

**КАЛАНДРА Пьетро**, доктор философии (Ph.D, физика), профессор Института по изучению наноструктурированных материалов (Рим, Италия), **Н=26**

**«Известия НАН РК. Серия физика и информатики».**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ 16906-Ж** выданное 14.02.2018 г.

Тематическая направленность: *серия физика и информационные коммуникационные технологии.* В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*

*<http://www.physico-mathematical.kz/index.php/en/>*

#### **EDITOR IN CHIEF:**

**MUTANOV Galimkair Mutanovich**, doctor of technical Sciences, Professor, Academician of NAS RK, acting director of the Institute of Information and Computing Technologies of SC MES RK (Almaty, Kazakhstan), **H=5**

#### **DEPUTY EDITOR-IN-CHIEF**

**MAMYRBAYEV Orken Zhumazhanovich**, Ph.D. in the specialty "Information systems, executive secretary of the RSE "Institute of Information and Computational Technologies", Committee of Science MES RK (Almaty, Kazakhstan) **H=5**

#### **EDITORIAL BOARD:**

**KALIMOLDAYEV Maksat Nuradilovich**, doctor in Physics and Mathematics, Professor, Academician of NAS RK (Almaty, Kazakhstan), **H=7**

**BAYGUNCHEKOV Zhumadil Zhanabayevich**, doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), **H=3**

**WOICIK Waldemar**, Doctor of Phys.-Math. Sciences, Professor, Lublin University of Technology (Lublin, Poland), **H=23**

**BOSHKAYEV Kuantai Avgazievich**, PhD, Lecturer, Associate Professor of the Department of Theoretical and Nuclear Physics, Al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=10**

**QUEVEDO Hemando**, Professor, National Autonomous University of Mexico (UNAM), Institute of Nuclear Sciences (Mexico City, Mexico), **H=28**

**ZHUSSUPOV Marat Abzhanovich**, Doctor in Physics and Mathematics, Professor of the Department of Theoretical and Nuclear Physics, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=7**

**KOVALEV Alexander Mikhailovich**, Doctor in Physics and Mathematics, Academician of NAS of Ukraine, Director of the State Institution «Institute of Applied Mathematics and Mechanics» DPR (Donetsk, Ukraine), **H=5**

**RAMAZANOV Tlekkabul Sabitovich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Vice-Rector for Scientific and Innovative Activity, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=26**

**TAKIBAYEV Nurgali Zhabagaevich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=5**

**TIGHINEANU Ion Mikhailovich**, Doctor in Physics and Mathematics, Academician, Full Member of the Academy of Sciences of Moldova, President of the AS of Moldova, Technical University of Moldova (Chisinau, Moldova), **H=42**

**KHARIN Stanislav Nikolayevich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Kazakh-British Technical University (Almaty, Kazakhstan), **H=10**

**DAVLETOV Askar Erbulanovich**, Doctor in Physics and Mathematics, Professor, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=12**

**CALANDRA Pietro**, PhD in Physics, Professor at the Institute of Nanostructured Materials (Monterotondo Station Rome, Italy), **H=26**

#### **News of the National Academy of Sciences of the Republic of Kazakhstan.**

**Series of physics and informatics.**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. 16906-ЖК**, issued 14.02.2018  
Thematic scope: *series physics and information technology.*

Currently: *included in the list of journals recommended by the CCSES MES RK in the direction of «information and communication technologies».*

Periodicity: *4 times a year.*

Circulation: *300 copies.*

Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*

*<http://www.physico-mathematical.kz/index.php/en/>*

NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN

PHYSICO-MATHEMATICAL SERIES

ISSN 1991-346X

Volume 2. Number 350 (2024). 108–119

<https://doi.org/10.32014/2024.2518-1726.270>

ӨОЖ 004.931

© A.M. Bissengaliyeva<sup>1</sup>, A.U. Issembayeva<sup>1\*</sup>, T.K. Dushayeva<sup>1</sup>, N.M. Almabayeva<sup>2</sup>,  
G.O. Ilyassova<sup>2</sup>, 2024

<sup>1</sup>Zhangir Khan West Kazakhstan agrarian Technical University;

<sup>2</sup>Kazakh National University named after S.D. Asfendiyarov, Almaty, Kazakhstan.

E-mail: aida\_bz@mail.ru

### KEYWORD COVERAGE USING SEMANTIC DATA ANALYSIS

**Bissengaliyeva Assyl** — Master of Technical Sciences in specialty , Senior lecturer, Zhangir Khan West Kazakhstan Agrarian-Technical University NCJSC, Uralsk, Kazakhstan

E-mail: B.a.m69@mail.ru, <https://orcid.org/0000-0002-6914-2352>; **Issembayeva Aida** — «Zhangir Khan West Kazakhstan agrarian Technical University», Senior lecturer, Master of Technical Sciences, Uralsk, Kazakhstan E-mail: aida\_bz@mail.ru, <https://orcid.org/0009-0005-8440-9106>;

**Dushayeva Tursyngul Kanatovna** — «Zhangir Khan West Kazakhstan agrarian Technical University», Senior lecturer, Master of Technical Sciences, Uralsk, Kazakhstan

E-mail: sultanna2011@mail.ru, <https://orcid.org/0000-0002-1244-8203>;

**Almabayeva Nurgaisa Makhmetshariyevna** — Lecturer of the department of «Normal physiology with biophysics course», Almaty, Kazakhstan

E-mail: almabaeva@kaznmu.kz, <https://orcid.org/0009-0005-7976-009X>;

**Ilyassova Gulzhan Ospanbekovna** — Lecturer of the department of «Normal physiology with biophysics course», Almaty, Kazakhstan

E-mail: ilyasova.g@kaznmu.kz, <https://orcid.org/0009-0008-6205-4959>.

**Abstract.** This article presents a thorough comparative analysis of two modern hybrid approaches in machine learning: Bidirectional Encoder Representations from Transformers (BERT) combined with an autoencoder (AE), and Term Frequency-Inverse Document Frequency (TF-IDF) combined with an autoencoder. The research focuses on the task of keyword extraction using semantic analysis of text data. The main goal of the work is to evaluate the effectiveness of these methods in adequately covering keywords in large text corpora spanning various thematic areas. The authors provide a detailed examination of the architecture and operating principles of each method. Special attention is given to the integration of these methods with autoencoders, which significantly improves the semantic integrity and relevance of the extracted keywords. The experimental part of the study includes a detailed analysis of the effectiveness of both methods on various text datasets, demonstrating how the structure and semantic richness of the original data affect the results of each method. The methodology used for evaluating the quality of keyword extraction, including metrics such as precision, recall, and F1 score, is also thoroughly described. The advantages and disadvantages of each



approach, as well as their suitability for specific types of text tasks, are analyzed. The results of the study provide valuable data for the scientific community and can be used to select the most appropriate text processing method for various applications requiring deep semantic understanding and high accuracy in information extraction.

**Keywords:** Machine learning, semantic data analysis, keyword extraction, TF-IDF, BERT, autoencoder, scientific texts, contextual analysis, semantic analysis, text data processing

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

© А.М. Бисенгалиева<sup>1</sup>, А.У. Исембаева<sup>1\*</sup>, Т.К. Душаева<sup>1</sup>, Н.М. Алмабаева<sup>2</sup>,  
Г.О. Ильясова<sup>2</sup>, 2024

<sup>1</sup>Жәңгір хан атындағы Батыс Қазақстан аграрлық-техникалық университеті, Орал,  
Қазақстан;

<sup>2</sup>С.Ж. Асфендияров атындағы Қазақ ұлттық медициналық университеті, Алматы,  
Қазақстан.

E-mail: aida\_bz@mail.ru

## СЕМАНТИКАЛЫҚ ДЕРЕКТЕРДІ ТАЛДАУ АРҚЫЛЫ КІЛТ СӨЗДЕРДІ ҚАМТУ

**Бисенгалиева Асыл Макымовна** — техника ғылымдарының магистрі, аға оқытушы, Жәңгір хан атындағы Батыс Қазақстан аграрлық-техникалық университеті» КЕАК, Орал, Қазақстан

E-mail: B.a.m69@mail.ru, <https://orcid.org/0000-0002-6914-2352>;

**Исембаева Аида Уалихановна** — «Жәңгір хан атындағы Батыс Қазақстан аграрлық-техникалық университеті» аға оқытушы, техника ғылымының магистрі, Орал қаласы, Қазақстан

E-mail: aida\_bz@mail.ru, <https://orcid.org/0009-0005-8440-9106>;

**Душаева Турсынгуль Канатовна** — «Жәңгір хан атындағы Батыс Қазақстан аграрлық-техникалық университеті» аға оқытушы, техника ғылымының магистрі, Орал қаласы, Қазақстан

E-mail: sultanna2011@mail.ru, <https://orcid.org/0000-0002-1244-8203>;

**Алмабаева Нургайша Махаметшариевна** — «Биофизика курсымен қалыпты физиология» кафедрасының лекторы, С.Ж. Асфендияров атындағы Қазақ Ұлттық Медициналық Университеті, Алматы, Қазақстан

E-mail: almabaeva@kaznmu.kz, <https://orcid.org/0009-0005-7976-009X>;

**Ильясова Гүлжан Оспанбековна** — «Биофизика курсымен қалыпты физиология» кафедрасының лекторы, С.Ж. Асфендияров атындағы Қазақ ұлттық медициналық университеті, Алматы, Қазақстан

E-mail: ilyasova.g@kaznmu.kz, <https://orcid.org/0009-0008-6205-4959>.

**Аннотация.** Бұл мақалада машиналық оқытудың екі заманауи гибриді тәсілдері, яғни Bidirectional Encoder Representations from Transformers (BERT) автоэнкодрмен (Autoencoder, AE) және Term Frequency-Inverse Document Frequency, TF-IDF автоэнкодрмен үйлесімде салыстырмалы талдау жасалады. Зерттеу мәтіндік деректердің семантикалық талдауын қолдану арқылы кілт сөздерді шығару міндетіне бағытталған. Жұмыстың негізгі мақсаты – түрлі тақырыптық салаларды қамтитын үлкен мәтіндерден кілт сөздерді тиімді қамту үшін осы әдістердің



тиімділігін бағалау. Авторлар әрбір қарастырылатын әдістің архитектурасын және жұмыс принциптерін егжей-тегжейлі зерттейді. Бұл әдістерді автоэнкодерлермен біріктірудің ерекшеліктеріне назар аударылады, бұл бөлінген кілт сөздердің семантикалық тұтастығын және маңыздылығын айтарлықтай жақсартуға мүмкіндік береді. Зерттеудің эксперименттік бөлімі әртүрлі мәтіндік деректер жинақтарында екі әдістің тиімділігін егжей-тегжейлі талдауды қамтиды, бастапқы деректердің құрылымы мен семантикалық байлығы әр әдістің нәтижелеріне қалай әсер ететінін көрсетеді. Жұмыста сондай-ақ кілт сөздерді шығару сапасын бағалау әдіснамасы, оның ішінде дәлдік, толықтық және F1 өлшемі сияқты көрсеткіштер туралы егжей-тегжейлі сипатталады. Әрбір тәсілдің артықшылықтары мен кемшіліктері, сондай-ақ олардың нақты мәтіндік тапсырмаларға жарамдылығы талданады. Зерттеу нәтижелері ғылыми қауымдастық үшін құнды мәліметтер береді және мәтіндердің семантикалық мазмұнын терең түсінуді және ақпаратты жоғары дәлдікпен шығаруды талап ететін әртүрлі қосымшаларда ең қолайлы өңдеу әдісін таңдауға пайдаланылуы мүмкін.

**Түйін сөздер:** машиналық оқыту, семантикалық деректерді талдау, кілт сөздерді шығару, TF-IDF, BERT, автоэнкодер, ғылыми мәтіндер, контекстік талдау, семантикалық талдау, мәтіндік деректерді өңдеу

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

© А.М. Бисенгалиева<sup>1</sup>, А.У. Исембаева<sup>1\*</sup>, Т.К. Душаева<sup>1</sup>, Н.М. Алмабаева<sup>2</sup>,  
Г.О. Ильясова<sup>2</sup>, 2024

<sup>1</sup>Западно-Казахстанский аграрно-технический университет имени Жангир хана,  
Уральск, Казахстан;

<sup>2</sup>Казахский национальный университет имени С.Д. Асфендиярова, Алматы,  
Казахстан.

E-mail: aida\_bz@mail.ru

## ОХВАТ КЛЮЧЕВЫХ СЛОВ С ПРИМЕНЕНИЕМ СЕМАНТИЧЕСКОГО АНАЛИЗА ДАННЫХ

**Бисенгалиева Асыл Макымовна** — магистр технических наук, старший преподаватель, Западно-Казахстанский аграрно-технический университет имени Жангир хана, Уральск, Казахстан  
E-mail: B.a.m69@mail.ru, <https://orcid.org/0000-0002-6914-2352>;

**Исембаева Аида Уалихановна** — Западно-Казахстанский аграрно-технический университет имени Жангир хана, старший преподаватель, магистр технических наук, Уральск, Казахстан  
E-mail: aida\_bz@mail.ru, <https://orcid.org/0009-0005-8440-9106>;

**Душаева Турсынгуль Канатовна** — Западно-Казахстанский аграрно-технический университет имени Жангир хана, старший преподаватель, магистр технических наук, Уральск, Казахстан  
E-mail: sultanna2011@mail.ru, <https://orcid.org/0000-0002-1244-8203>;

**Алмабаева Нургайша Махаметшариевна** — лектор кафедры «Нормальная физиология с курсом биофизики», Казахский национальный университет имени С.Д. Асфендиярова, Алматы, Казахстан  
E-mail: almabaeva@kaznmu.kz, <https://orcid.org/0009-0005-7976-009X>;

**Ильясова Гулжан Оспанбековна** — лектор кафедры «Нормальная физиология с курсом биофизики», Казахский национальный университет имени С.Д. Асфендиярова, Алматы, Казахстан

**Аннотация.** В статье осуществляется сравнительный анализ двух современных гибридных подходов в машинном обучении, такие как Bidirectional Encoder Representations from Transformers (BERT) в сочетании с автокодировщиком (Autoencoder, AE) и Термино-Частотное Обратное Документное Частотное (Term Frequency-Inverse Document Frequency, TF-IDF) в сочетании с автокодировщиком. Исследование фокусируется на задаче извлечения ключевых слов с применением методов семантического анализа текстовых данных. Основная цель работы заключается в оценке эффективности данных методов для обеспечения адекватного охвата ключевых слов в больших текстовых корпусах, охватывающих различные тематические области. Авторы подробно изучают архитектуру и принципы работы каждого из рассматриваемых методов. Особое внимание уделяется особенностям интеграции этих методов с автоэнкодерами, что позволяет значительно улучшить семантическую целостность и релевантность выделенных ключевых слов. Экспериментальная часть исследования включает в себя детальный анализ эффективности обоих методов на различных наборах текстовых данных, демонстрируя, как структура и семантическая насыщенность исходных данных влияют на результаты работы каждого из методов. В работе также подробно описывается примененная методология оценки качества извлечения ключевых слов, включая такие показатели, как точность, полнота и мера F1. Анализируются преимущества и недостатки каждого подхода, а также их пригодность для конкретных типов текстовых задач. Результаты исследования предоставляют ценные данные для научного сообщества и могут быть использованы для выбора наиболее подходящего метода обработки текстов в различных приложениях, где требуется глубокое понимание семантического содержания и высокая точность извлечения информации.

**Ключевые слова:** машинное обучение, семантический анализ данных, извлечение ключевых слов, TF-IDF, BERT, автокодировщик, научные тексты, контекстный анализ, семантический анализ, обработка текстовых данных

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

### **Кіріспе**

Қазіргі таңда мәтінді өңдеуде кілт сөздерді шығару (Хапкс, 2023) міндеті үлкен маңызға ие, өйткені бұл ақпаратты ұйымдастыру, іздеу және талдауда (Чанг, 2023) маңызды рөл атқарады (Маниас, 2023). Кілт сөздерді тиімді шығару іздеу (Фу, 2023) жүйелерінің, ұсыныс жүйелерінің, сондай-ақ аналитикалық және білім беру құралдарының жұмысын жақсартады. Осыған байланысты, осы процесті автоматтандырып, оңтайландыруға қабілетті машиналық оқыту (Эдара, 2023; Баладжи, 2021) әдістерін әзірлеу және салыстыру өзекті ғылыми-зерттеу міндетіне айналууда. Бұл жұмыста екі гибриді тәсілдің терең талдауы жүргізіледі: BERT+Автоэнкодер (AE) және TF-IDF+Автоэнкодер. Екі әдіс те үлкен көлемдегі мәтіндік деректерді өңдеу үшін машиналық оқыту және табиғи тілді өңдеу саласындағы заманауи жетістіктерді пайдаланады. Google әзірлеген BERT (Bidirectional Encoder Representations from Transformers) моделі мәтіндегі сөздердің семантикасын түсінуге арналған озық технология болып табылады, ал TF-IDF

(Term Frequency-Inverse Document Frequency) сөздің құжаттағы және жинақтағы маңыздылығын бағалаудың дәстүрлі статистикалық әдісі болып табылады (Саранья, 2023). Бұл тәсілдерді деректерді тиімді түрде сығымдап және қалпына келтіретін автоэнкодерлермен біріктіру алынған кілт сөздердің сапасы мен семантикалық байлығын жақсартуға мүмкіндік береді (Маккитрик, 2023).

Зерттеудің мақсаты – кілт сөздерді шығару бойынша осы екі әдісті салыстыру, олардың әртүрлі мәтіндерде кілт сөздерді жеткілікті түрде қамту қабілетін бағалау. Біз әрбір әдістің архитектурасын, жұмыс принциптерін және автоэнкодерлермен интеграциялау ерекшеліктерін талдаймыз, сондай-ақ бірнеше мәтіндік деректер жиынтығында эксперименттік салыстыру жүргіземіз. Бұл бастапқы деректердің құрылымы мен семантикасына байланысты олардың тиімділігіндегі айырмашылықтарды анықтауға мүмкіндік береді. Жұмыста кілт сөздерді шығару сапасын бағалау әдіснамасына, оның ішінде дәлдік, толықтық және F1 өлшемі сияқты көрсеткіштерге ерекше назар аударылады. Біздің зерттеу нәтижелері ғылыми қауымдастық үшін маңызды бастапқы деректерді ұсынады және терең семантикалық талдау мен мәтінді өңдеуде жоғары дәлдікті талап ететін нақты қосымшалар үшін ең қолайлы әдісті таңдауға көмектесуі мүмкін.

Сонымен қатар, бұл жұмыс үлкен көлемді деректері бар құжаттарды жіктеуде ерекше өзектілікке ие болады. Талдауға алынған әдістер ауқымды мәтіндік жиынтықтарды өңдеудің дәлдігі мен тиімділігін арттыру үшін қолданылуы мүмкін, бұл ақпараттық ағындардың тұрақты өсуі және оларды жедел өңдеу қажеттілігі жағдайында (Син, 2023) өте маңызды.

#### Әдістер мен материалдар

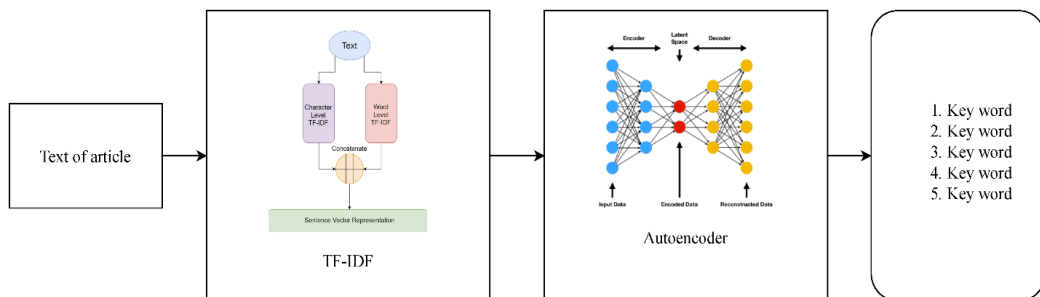
Бұл зерттеуде біз машиналық оқытудың екі гибриді әдісіне назар аудардық: BERT+Автоэнкодер және TF-IDF+Автоэнкодер, семантикалық деректерді талдау арқылы кең көлемді мәтіндік корпустардан кілт сөздерді шығару тиімділігін бағалау мақсатында. Екі әдіс те терең оқыту және автоэнкодер тұжырымдамаларын біріктіреді, бұл мәтіндерді тек үстірт деңгейде ғана емес, сонымен қатар сөздер мен фразалар арасындағы терең семантикалық және контекстік байланыстарды талдауға мүмкіндік береді, осылайша кілт сөздерді шығару процесін байытады.

Зерттеуіміздегі бірінші қадамымыз TF-IDF әдісін қолдану болды, ол құжаттағы әр сөздің барлық мәтіндік корпусқа қатысты маңыздылығын бағалауға мүмкіндік береді. TF-IDF сөздердің құжаттағы жиілігі мен олардың кездесетін корпустағы құжаттардың кері жиілігі негізінде сөздердің салмақтарын есептейді. Бұл тәсіл жиі қолданылатын, бірақ ақпараттық мазмұны аз сөздердің әсерін азайтуға көмектеседі және нақты мәтіндер үшін бірегей болып табылатын сөздерді бөлектейді. Алынған сөздердің векторлық көріністері автоэнкодерде одан әрі өңдеу үшін кіріс деректері ретінде қызмет етті.

Бұл жұмыста қолданылған автоэнкодер екі негізгі компоненттен тұрады: кодер және декодер. Кодер TF-IDF көмегімен алынған мәтіннің векторлық көрінісін тығыз әрі ақпараттық ішкі көрініске сығады. Кейін декодер осы сығылған көріністен бастапқы векторды қалпына келтіруге тырысады, ақпарат жоғалтуды барынша азайтуға ұмтылады. Бұл кезеңнің мақсаты – модельді бастапқы мәтіннен ең маңызды семантикалық белгілерді бөліп, сақтай алатындай етіп оқыту, бұл кілт сөздерді шығару сапасы мен дәлдігін арттырады.

TF-IDF пен автоэнкодерлерді біріктіру алынған кілт сөздердің семантика-

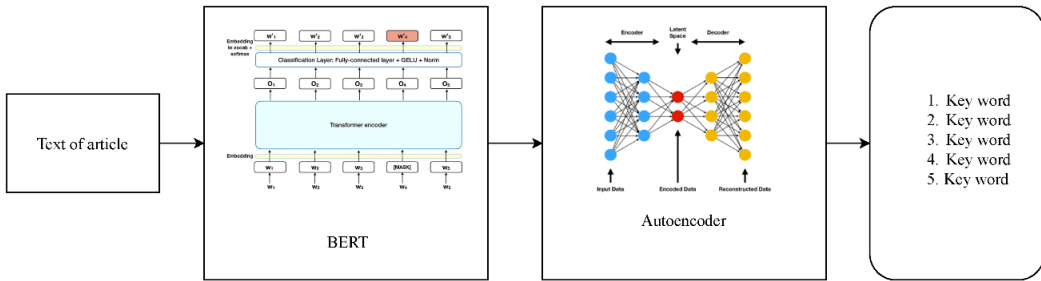
лық тұтастығын айтарлықтай жақсартуға мүмкіндік берді. Бұл әдіс мәтіндік деректердің тереңірек талдауын қамтамасыз етеді, сөздердің кең мағынадағы жасырын байланыстары мен маңыздылығын анықтайды. Осылайша, бұл тәсілдерді біріктіру мәтінді талдаудың қуатты және дәл құралдарын жасауға ықпал етеді, бұл кілт сөздерді шығаруды автоматтандыру және үлкен деректерді ақпараттық іздеу мен өңдеуді жақсарту бағытында маңызды қадам болып табылады.



**Сур. 1.** TF-IDF + Автоэнкодер гибридік әдісінің архитектурасы

(Fig. 1. Architecture of the hybrid method TF-IDF + Autoencoder)

Мәтіндік деректерді талдауды жақсарту және кілт сөздерді дәлірек шығару үшін бұл зерттеуде BERT (Bidirectional Encoder Representations from Transformers) моделі қолданылды. BERT табиғи тілді өңдеу саласындағы озық әдістердің бірі болып табылады, ол мәтіндерді талдау үшін трансформерлердің назар аудару механизмдерін пайдаланады. BERT-тің ерекшелігі - мәтінді екі бағытта өңдеу қабілеті, бұл модельге сөздердің мағынасын солдан оңға және оңнан солға бір уақытта талдауға мүмкіндік береді. Бұл екі бағытты мәтінді түсіну модельдің сөздер арасындағы семантикалық байланыстарды түсіну қабілетін айтарлықтай арттырады, контекстік нюанстарды терең түсінуді қамтамасыз етеді және кілт сөздерді шығару сапасын жақсартады. BERT көмегімен мәтінді өңдегеннен кейін келесі қадам алынған векторлық көріністерді сығу үшін автоэнкодерді қолдану болды. Кодер мен декодерден тұратын автоэнкодерлер ақпараттың жоғалуын азайту қағидаты бойынша бастапқы вектордан сығылған көрініске және кері ауысуға жұмыс істейді. Біздің жағдайда кодер BERT-тің көп өлшемді векторларын тығыз векторларға сығып, содан кейін декодер оларды қайта қалпына келтіруге тырысты. Бұл процедураның мақсаты - мәтінде бар ең маңызды ақпаратты бөліп көрсету және сақтау, бұл кілт сөздерді шығару кезінде жоғары дәлдік пен сапаны қамтамасыз етті. Бұл сығу деректердің үлкен көлемін басқаруды және талдауды жақсартуға, мәтіннің ең маңызды элементтерін анықтауға көмектесті, оларды ақпаратты өңдеудің әртүрлі қосымшаларында кейіннен қолдануға мүмкіндік берді.



Сур. 2. BERT + Автоэнкодер гибридік әдісінің архитектурасы

(Fig. 2. Architecture of the BERT + Autoencoder Hybrid Method)

BERT және TF-IDF әдістерін автоэнкодерлермен біріктіруді қолдану зерттеуімізде кілт сөздерді шығару үшін мәтінді контекстуалды өңдеудің маңыздылығын толық түсінуге мүмкіндік берді. Әсіресе, BERT арқылы қамтамасыз етілетін мағынаны екі бағытты түсіну өте маңызды болды. Бұл мәтіннің реті ғана емес, сонымен қатар көптеген контекстуалды тәуелділіктерді де ескеруге мүмкіндік береді, бұл мәтінді қабылдауды айтарлықтай байытады. Бұл технологияларды ақпаратты тиімді түрде сығатын автоэнкодерлермен бірге қолдану бұл әсерді күшейтеді, ең маңызды семантикалық атрибуттарды бөліп көрсетіп, алынған кілт сөздердің сапасын жақсартады. Эксперименттердің нәтижелері осындай біріктірілген әдістерді қолдану тек кілт сөздерді анықтау дәлдігін арттырып қана қоймай, мәтіндерді тереңірек талдауға да ықпал ететінін растады. Ақпаратты сығу үшін автоэнкодерлерді қолдану деректердің шамадан тыс жүктелуінен сақтайды және деректерді өңдеуді басқаруға оңай әрі тиімді етеді. Бұл әсіресе, әрбір сөз бен оның мағынасы талдау нәтижесіне айтарлықтай әсер етуі мүмкін үлкен мәтіндік корпустармен жұмыс істегенде өте маңызды. Нәтижесінде, мұндай тәсіл кілт сөздерді шығару дәлдігін жақсартып қана қоймай, мәтіннің құрылымы мен мағынасын түсінуді байытады, бұл көптеген салаларда, соның ішінде ғылыми зерттеулерде, білімді басқаруда және ақпараттық іздеуде өте маңызды.

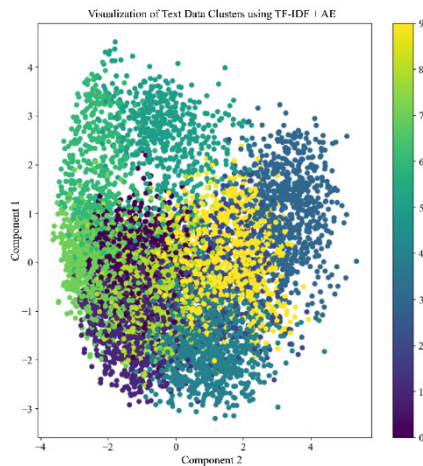
Нәтижелер және оларды талқылау

Зерттеу үшін әртүрлі ғылыми бағыттарды қамтитын 182 ғылыми мақала пайдаланылды. Бұл бағыттар биологияны, компьютерлік ғылымдарды, физиканы, химияны, психологияны және лингвистиканы қамтиды. Әрбір категория белгілі бір білім саласын білдіреді және шамамен 20-25 құжаттан тұрады. Мұндай сан алуан категориялар жан-жақты талдауды қамтамасыз етеді және әртүрлі ғылыми пәндерде кілт сөздерді шығару әдістерінің тиімділігін бағалауға мүмкіндік береді. Мақалалар әрбір ғылыми бағытты біркелкі көрсету үшін таңдалған, бұл таңдалған әдістердің әртүрлі бағыттарда қолданылуын жалпылау туралы қорытынды жасауға мүмкіндік береді.

Эксперименттер жүргізу үшін біз әртүрлі тақырыптар мен жазу стильдерін қамтитын мәтіндік деректер жиынтықтарын пайдаландық. Бағалау әдістері кілт сөздерді шығару сапасын бағалау үшін дәлдік, толықтық және F1 өлшемін талдауды қамтыды. Екі әдіспен алынған кілт сөздер KMeans және PCA әдістерін пайдаланып, сөздерді векторлық кеңістікте кластерлеу және өзара орналасуын

талдау үшін визуализацияланды. Бұл әдістер кілт сөздердің семантикалық жақындығы мен тақырыптық маңыздылығына қарай қалай топтастырылатынын бағалауға көмектесті.

3-суретте нүктелер арасындағы байланысты кілт сөздердің семантикалық немесе жақындық дәрежесі ретінде түсіндіруге болады. Жақын топтастырылған нүктелері бар кластерлер деректердегі неғұрлым нақты анықталған тақырыптар немесе тұжырымдамаларды көрсетуі мүмкін. Оң жақтағы түс шкаласы кілт сөздермен байланысты салмақ параметрін немесе метриkanı көрсетеді, бірақ қосымша сөздерсіз оның табиғатын анықтау қиын. Бұл метрика кілт сөздердің деректер жиынтығындағы маңыздылығы немесе жиілігі болуы мүмкін. Кілт сөздердің кластерленуі KMeans және PCA (Principal Component Analysis) әдістерін пайдалана отырып, екі графикте көрсетілген. BERT моделін визуализациялау кластерлердің неғұрлым нақты және айқын бөлінуін көрсетеді, бұл сөздердің неғұрлым мәнерлі және сараланған векторлық көрінісін білдіреді. Бұл BERT арқылы алынған векторлар кілт сөздердің семантикалық ұқсастығы мен айырмашылығын жақсырақ көрсететінін білдіреді, бұл сапалы кластерлеу үшін маңызды. Көрсетілген график TF-IDF және автоэнкодерлер комбинациясын пайдаланып, кілт сөздерді кластерлеу нәтижелерін көрсетеді. Өлшемділікті азайту және кейінгі визуализация үшін KMeans және PCA әдістері қолданылады. Нәтижелер кілт сөздердің семантикалық жақындығының негізгі тенденцияларын көрсететін екі өлшемді кеңістікте концентрация мен дисперсияның әртүрлі дәрежесі бар кластерлердің қалыптасуын көрсетеді. Түс шкаласы әр кілт сөзге қатысты қосымша көрсеткіштерді көрсете алады.



**Сур. 3.** TF-IDF + Автоэнкодер әдісі бойынша кілт сөздерді кластерлеу нәтижесі

(Fig. 3. The result of keyword clustering using the TF-IDF + Autoencoder method)

4-суретте кілт сөздер тізімі мен оларға сәйкес пайыздар көрсетілген мәтіндік нәтиже көрсетілген. Бұл деректер TF-IDF әдісі мен автоэнкодерді қолдану арқылы ғылыми мақалалар мәтінінен алынған, сипаттамаға сәйкес. TF-IDF, яғни «термин жиілігіне кері құжат жиілігі», сөздің құжаттағы маңыздылығын бағалау үшін қол-



данылатын статистикалық өлшемді білдіреді, ол коллекция немесе корпус бөлігі болып табылады. Автоэнкодерлер деректердің өлшемін азайту үшін қолданылуы мүмкін, бұл ең маңызды бөлшектерді анықтауға көмектеседі.

```
1/1 [=====] - 0s 112ms/step
Top keywords:
layer : 0.389
deep : 0.344
network : 0.311
convolutional : 0.292
lstm : 0.255
neural network : 0.219
training : 0.21
architecture : 0.204
recurrent : 0.201
rnn : 0.193
```

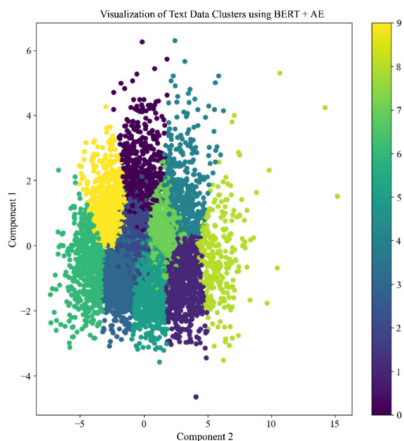
*Сур. 4.* TF-IDF + Автоэнкодер әдісі бойынша кілт сөздерді классификациялау нәтижесі

(Fig. 4. The result of keyword classification using the TF-IDF + Autoencoder method)

Бұл тізім негізінде бұл кілт сөздер алынған ғылыми мақаланың мәтіні Машиналық оқыту саласына, атап айтқанда терең оқыту (deep learning) және нейрондық желілерге (нейрондық желілерге) қатысты екенін болжауға болады. «Деректер», «маңызды», «алаң», «машина», «өңдеу» және «есептеу» сөздері деректерді өңдеу және машиналық оқытудың есептеу аспектілерімен байланысты тақырыптардың талқылануы мүмкін екенін растайды. «Оқыту» және «терең» сөздерінің ең жоғары пайызға ие болуы бұл ұғымдардың қарастырылып отырған мақалада орталық екенін көрсетуі мүмкін. Сонымен қатар, «машина», «өңдеу» және «есептеу» сияқты сөздердің тең пайызға ие болуы олардың зерттеліп жатқан тақырыптағы өзара байланысын көрсетуі мүмкін.

5-суретте автоэнкодермен бейімделген BERT моделін пайдалану арқылы тест деректер жиынын тексеру барысында алынған кілт сөздердің кластерлену нәтижелері көрсетілген. Нәтижелерді екі өлшемді кеңістікте визуализациялау үшін KMeans және PCA әдістері қолданылды. Графикте әрбір нүкте кілт сөзге сәйкес келеді, нүктелердің түстері кілт сөз тиесілі кластерді көрсетеді, ал олардың орны PCA қолдану нәтижесінде алынған алғашқы екі негізгі компоненттің мәндерімен анықталады.





**Сур. 5.** BERT + Автоэнкодер әдісі бойынша кілт сөздерді кластерлеу нәтижесі  
(Fig. 5. The result of keyword clustering using the BERT + Autoencoder method)

6-суретте BERT моделі мен автоэнкодер комбинациясын пайдалану арқылы ғылыми мақаланың мәтінінен алынған кілт сөздердің тізімі көрсетілген, әрбір кілт сөзге белгілі бір пайыз берілген. Ол өңделіп жатқан мәтінде әрбір кілт сөздің үлесін көрсетеді. Машиналық оқыту және жасанды интеллектпен байланысты сөздер тізімнің жоғарғы жағында орналасқан. «Оқыту» сөзі бірінші орында 5,78 % үлеспен, ал одан кейін «терең» сөзі 4,33 % үлеспен орналасқан, бұл талданатын мәтінде терең оқыту тақырыбына ерекше назар аударылатынын көрсетеді. Келесі терминдер «желі», «деректер» және «нейрондық» бұл назардың нейрондық желілер мен деректерге бағытталғанын растайды. «Маңызды», «алаң», «машина», «өңдеу» және «есептеу» сияқты кілт сөздер де мақаланың мазмұнында рөл атқара алады, бірақ олар негізгі тақырыппен салыстырғанда қосымша аспектілерді немесе сипаттамаларды білдіреді.

```

Top keywords with a certain percentage:
learning: 5.78%
deep: 4.33%
network: 3.61%
data: 2.53%
neural: 2.17%
significant: 1.81%
area: 1.44%
machine: 1.08%
processing: 1.08%
computing: 1.08%
    
```

**Сур. 6.** BERT + Автоэнкодер әдісі бойынша классификация нәтижесі  
(Fig. 6. The result of classification using the BERT + Autoencoder method)

Зерттеу нәтижелері Bidirectional Encoder Representations from Transformers (BERT) әдісі автокодировщикпен (Autoencoder) үйлесімде қолданғанда, TF-IDF әдісі автокодировщикпен үйлесімде қолданғаннан гөрі кілт сөздерді шығару

тапсырмасында жақсы нәтижелер көрсететінін көрсетеді. Бұл BERT+Autoencoder әдісінің табиғи тілді тереңірек түсіну қабілетінің бар екенін дәлелдейді. Сонымен қатар, ол күрделі мәтіндік деректерге жақсырақ бейімделеді. Модельдерді оқыту тарихы оқыту және тексеру барысында қателердің өзгеруін көрсететін графиктерде ұсынылған. BERT (Bidirectional Encoder Representations from Transformers) моделі графиктері қателіктің оқыту кезеңінде де, тексеру кезеңінде де барлық эпохаларда үздіксіз және дәйекті түрде азайғанын көрсетеді. Бұл BERT моделінің жоғары жалпылау қабілетін көрсетеді, ол айқын қайта оқыту белгілерінсіз ақпаратты сенімді түрде игере алады. Керісінше, TF-IDF (Term Frequency-Inverse Document Frequency) моделі бойынша деректер ұсынылмаған, бірақ машиналық оқытудың жалпы тенденцияларына сүйене отырып, TF-IDF белгілі бір деректер түрлерімен жақсы жұмыс істей алады, бірақ күрделі табиғи тілді өңдеу тапсырмаларында BERT моделінен төмен болуы мүмкін деп болжауға болады

#### Қорытынды

Бұл мақалада ғылыми мәтіндерден кілт сөздерді шығару үшін машиналық оқытудың екі заманауи гибриді тәсілдерінің мұқият салыстырмалы талдауы жүргізілген: BERT (Bidirectional Encoder Representations from Transformers) автокодировщикпен (Autoencoder) және TF-IDF (Term Frequency-Inverse Document Frequency) автокодировщикпен үйлесімде. Зерттеу негізінде келесі қорытындылар жасауға болады.

Біріншіден, нәтижелер BERT+Autoencoder әдісінің TF-IDF+Autoencoder әдісімен салыстырғанда кілт сөздерді шығару тапсырмасында жақсы нәтижелер көрсететінін көрсетті. Бұл BERT моделінің табиғи тілді тереңірек түсіну және күрделі мәтіндік деректерге жақсырақ бейімделу қабілетін дәлелдейді. BERT-тің мәтінді екі бағытта өңдеуі модельге сөздер арасындағы байланыстарды ескеруге мүмкіндік береді, бұл алынған кілт сөздердің дәлдігі мен өзектілігін айтарлықтай арттырады.

Екіншіден, графиктермен ұсынылған модельдерді оқыту тарихы BERT моделі үшін қателіктің оқыту кезеңінде де, тексеру кезеңінде де барлық дәуірлерде үздіксіз азайғанын көрсетеді, бұл оның қайта оқыту белгілерінсіз жоғары жалпылау қабілетін көрсетеді. TF-IDF моделі үшін мұндай тенденция анықталған жоқ, бұл оның күрделі мәтіндерді өңдеуде төмен тиімділігін көрсетуі мүмкін.

Үшіншіден, KMeans және PCA әдістерін пайдалана отырып, кілт сөздерді кластерлеу нәтижелерін визуализациялау BERT моделінің артықшылығын растады. BERT арқылы алынған сөздердің векторлық көріністері кілт сөздер арасындағы семантикалық ұқсастық пен айырмашылықты жақсырақ көрсетеді, бұл дәлірек кластерлеуге және мәтіндерді тереңірек талдауға ықпал етеді.

Осылайша, BERT әдісі автокодировщикпен үйлесімде ғылыми мәтіндерден кілт сөздерді шығару тапсырмаларында TF-IDF әдісінен автоэнкодермен үйлесімде асып түсетінін қорытындылауға болады. BERT тек кілт сөздерді жақсырақ кластерлеуді және шығаруды қамтамасыз етіп қана қоймай, сонымен қатар оқыту және тексеру процесінде қателердің тұрақты түрде азаюын көрсетеді. Сөздердің қуатты көріністері және терең семантикалық талдау мүмкіндіктерінің арқасында BERT әрбір терминнің нақты маңызы шешуші рөл атқаратын күрделі ғылыми материалдарды өңдеу үшін таңдаулы таңдау болып табылады.

ӘДЕБИЕТ

Баладжи Т.К., Аннаварапу К.С.Р. және Баблани А. (2021). Әлеуметтік медианы талдауға арналған машиналық оқыту алгоритмдері: сауалнама. *Информатикаға Шолу*, — 40, — 100395.

Маккистрик М.К., Шуурман Н. және Крукс В.А. (2023). Орналасқан жеріне негізделген әлеуметтік медиа деректерін жинау, талдау және визуализациялау: ГАЖ-дағы әдістерге шолу-әлеуметтік медианы талдау. *Геожурнал*, — 88 (1). — 1035–1057.

Маниас Г., Маврогиоргу А., Киуртис А., Симвулидис С. және Кириазис Д. (2023). Мәтінді көптілді санаттарға бөлу және көңіл-күйді талдау: twitter деректерін жіктеудің көптілді тәсілдерін қолданудың салыстырмалы талдауы. *Нейрондық Есептеу және Қолдану*, — 35(29). — 21415–21431.

Саранья С. және Уша Г. (2023). Twitter-дегі Көңіл-күйді Талдау Үшін IntelligentWordNet Lemmatize Көмегімен Машиналық оқытуға негізделген Әдіс. *Интеллектуалды Автоматтандыру Және Жұмсақ Есептеу*, — 36 (1).

Син К., Ол Ю., Пан Ю., Ван Ю. және Ду С. (2023). NLP алгоритміне негізделген ЖАСАНДЫ ИНТЕЛЛЕКТКЕ негізделген жарнамалық итеру жүйесін енгізу. *Халықаралық Информатика Және Ақпараттық Технологиялар Журналы*, — 1 (1), — 30–37.

Фу Г., Ли Б., Янг Ю. және Ли С. (2023). Мәтінді санаттарға бөлу үшін көп сатылы агрегациясы бар ансамбльді ҚАЙТА бағалау және ТОПСИС негізінде таңдау. *Үлгіні Тану Хаттары*, — 168. — Pp. 47–56.

Хапкс Д., Джулианелли М., Данкерс В., Артеткс М., Элазар Ю., Пиментель Т. & Джин З. (2023). Nlp-дегі жалпылама зерттеулердің таксономиясы және шолуы. *Табиғат Машинасының Интеллекстісі*, — 5 (10). — 1161–1174.

Чанг С., Мун С., Ким Дж., Ким Дж., Лим С. және Чи С. (2023). Құрылыстағы және информатикадағы табиғи тілдерді өңдеу (NLP) қолданбаларын жүйелі шолулар (PRISMA) үшін таңдаулы есеп беру элементтерін пайдалана отырып салыстыру. *Құрылыстағы автоматтандыру*, — 154, — 105020.

Әдара, Колумбия Округі, Ванукури, Лос-Анджелес, Систла В. және Колли В.К.К. (2023). Көңіл-күйді талдау және lstm көмегімен онкологиялық аурулардың медициналық жазбаларын мәтіндік санаттарға бөлу. *Қоршаған Органы Барлау және Ізгілендірілген Есептеу журналы*, — 14 (5). — 5309–5325.

REFERENCES

Balaji T.K., Annavarapu C.S.R. & Bablani A. (2021). Machine learning algorithms for social media analysis: — A survey. *Computer Science Review*, — 40. — 100395.

Chung S., Moon S., Kim J., Kim J., Lim S. & Chi S. (2023). Comparing natural language processing (NLP) applications in construction and computer science using preferred reporting items for systematic reviews (PRISMA). *Automation in Construction*, — 154. — 105020.

Edara D.C., Vanukuri L.P., Sistla V. & Kolli V.K.K. (2023). Sentiment analysis and text categorization of cancer medical records with LSTM. *Journal of Ambient Intelligence and Humanized Computing*. — 14(5). — 5309–5325.

Fu G., Li B., Yang Y. & Li C. (2023). Re-ranking and TOPSIS-based ensemble feature selection with multi-stage aggregation for text categorization. *Pattern Recognition Letters*, — 168. — 47–56.

Hupkes D., Giulianelli M., Dankers V., Artetxe M., Elazar Y., Pimentel T. & Jin Z. (2023). A taxonomy and review of generalization research in NLP. *Nature Machine Intelligence*. — 5(10). — 1161–1174.

Manias G., Mavrogiorgou A., Kiourtis A., Symvoulidis C. & Kyriazis D. (2023). Multilingual text categorization and sentiment analysis: a comparative analysis of the utilization of multilingual approaches for classifying twitter data. *Neural Computing and Applications*, — 35(29). — 21415–21431.

McKitrick M.K., Schuurman N. & Crooks V.A. (2023). Collecting, analyzing, and visualizing location-based social media data: review of methods in GIS-social media analysis. *GeoJournal*, — 88(1). — 1035–1057.

Saranya S. & Usha G. (2023). A Machine Learning-Based Technique with IntelligentWordNet Lemmatize for Twitter Sentiment Analysis. *Intelligent Automation & Soft Computing*, — 36(1).

Xin Q., He Y., Pan Y., Wang Y. & Du S. (2023). The implementation of an AI-driven advertising push system based on a NLP algorithm. *International Journal of Computer Science and Information Technology*, — 1(1). — Pp. 30–37.

## МАЗМҰНЫ

<b>Н. Абдразақұлы, Л. Черикбаева, Н. Мұқажанов, Ж. Алибиева</b> АНСАМБЛЬДІК ТӘСІЛ НЕГІЗІНДЕ КЕСКІНДІ ӨНДЕУДІҢ ТИІМДІ АЛГОРИТМІН ҚҰРУ.....	7
<b>Б.Т Абыканова, А.А. Таугенбаева, А.Г. Амангосова, Г.Т. Бекова, А.Ж. Ақматбекова</b> ӨЗДІГІНЕН БІЛІМ АЛУШЫЛАРДЫ ЖЕТІЛДІРУ МЕН ДАМУДАҒЫ ИНТЕРАКТИВТІ БІЛІМ БЕРУ ТЕХНОЛОГИЯЛАРЫ.....	30
<b>Ж.Ж. Ажибекова, Д.И. Усипбекова, Б.Н. Джаханова, К. Жыланбаева, Ә.Н. Тұрсун</b> МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІМЕН ҒАРЫШТЫҚ КЕСКІНДЕРДЕН БҰЛТТАР МЕН ТҰМАНДЫҚТАРДЫ ЖОЮ.....	43
<b>М. Айтимов, Г.Б. Абдикеримова, К.К. Макулов, Б.А. Досжанов, Р.У. Альменаева</b> МАШИНАЛЫҚ ЖӘНЕ ТЕРЕҢ ОҚЫТУ АЛГОРИТМДЕРІ АРҚЫЛЫ МӘТІННІҢ ЭМОЦИОНАЛДЫҚ ЖАҒДАЙЫН ЗЕРТТЕУ.....	57
<b>А.Т. Ақынбекова, А.А. Муханова, Salah Al-Majeed, Г.С. Алтаева</b> АЙМАҚТЫ ДАМУДАҒЫ ӨЛЕУМЕТТІК ПРОЦЕСТЕРІН БАҒАЛАУ ҮШІН ШЕШІМДЕР ҚАБЫЛДАУДЫҢ БҰЛДЫР МОДЕЛЬДЕРІ.....	69
<b>К.М. Алдабергенова, А.Б. Касекеева, М.Ж. Айтимов, К.К. Дауренбеков, Т.Н. Есикова</b> АГРОӨНЕРКӘСІП КЕШЕНІНІҢ ЛОГИСТИКАСЫНЫҢ МАРКЕТИНГТІК БАСҚАРУЫН ЖЕТІЛДІРУ.....	85
<b>А.Е. Әбжанова, А.А. Быков, С.К. Сағнаева, Е.Ә. Әбжанов, Д.И. Суржик</b> ЖЕР АСТЫ ЖЕР АСТЫ СУЛАРЫН ЕСКЕРЕ ОТЫРЫП, ТОПЫРАҚТЫ МОДЕЛЬДЕУДІ ОҢТАЙЛАНДЫРУ.....	96
<b>А.М. Бисенгалиева, А.У. Исембаева, Т.К. Душаева, Н.М. Алмабаева, Г.О. Ильясова</b> СЕМАНТИКАЛЫҚ ДЕРЕКТЕРДІ ТАЛДАУ АРҚЫЛЫ КІЛТ СӨЗДЕРДІ ҚАМТУ.....	108
<b>А.Х. Давлетова, Н.Н. Оразова, Ж.Б. Сайлау, Д.Н. Қурмангалиева, Г.Л. Абдугалимов</b> БАСТАУЫШ СЫНЫП ОҚУШЫЛАРЫН ХАЛЫҚАРАЛЫҚ PIRLS ЗЕРТТЕУІНЕ АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР АРҚЫЛЫ ДАЯРЛАУ ЖОЛДАРЫ.....	120
<b>Г. Есмагамбетова, А. Кубигенова, А. Ақтаева, И. Цэрэн-Онолт, М. Есмагамбет</b> КВАНТТЫҚ ЕСЕПТЕУЛЕРГЕ НЕГІЗДЕЛГЕН БИОМЕТРИЯЛЫҚ ДЕРЕКТЕРДІ ҚОРҒАУ ӘДІСТЕРІ.....	137
<b>Г.Қ. Ешмұрат, Л.С. Қанбаева,</b> МАТЕМАТИКАЛЫҚ ҮРЕЙ ЖӘНЕ ОНЫҢ БОЛАШАҚ МАТЕМАТИКА ПӘНІ МҰҒАЛІМДЕРІНІҢ МАНСАБЫНА ӨСЕРІ.....	149
<b>Т.К. Жукабаева, В.А. Десницкий, Е.М. Марденев</b> СЫМСЫЗ СЕНСОРЛЫҚ ЖЕЛІЛЕРДЕГІ ДЕРЕКТЕРДІ ЖИНАУ, ӨНДЕУ ЖӘНЕ ТАЛДАУ ӘДІСТ ЕМЕСІ.....	163
<b>А.М. Джумагалиева, А.Ә. Шекербек, Ж.Ж. Хамитова, М. Свобода, С.А. Қалдар</b> АДАПТИВТІ АНОМАЛИЯНЫ АНЫҚТАУ ЖҮЙЕЛЕРІНІҢ КИБЕРҚАУІПСІЗДІГІН МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ АРТТЫРУ.....	177

<b>А.А. Исмаилова, Г.Е. Мырзабекова, М.Ж. Базарова, Г.Ж. Нурова, Г.Т. Азиева</b> ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІН ПАЙДАЛАНУ АРҚЫЛЫ ҚАРЖЫ НАРЫҒЫНДАҒЫ БАҒАЛАРДЫ БОЛЖАУ.....	190
<b>К. Кошанова, Сапарбайқызы, К.Е. Жангазакова, А.С. Сағынбай, Э. Куриэль-Марин</b> STEM-ДЕ БІЛІМ БЕРУ ӘЛЕУЕТІН БАРЫНША ПАЙДАЛАНУ: ОҚУ НӘТИЖЕЛЕРІН ЖАҚСARTУҒА ҮЛЕС, ҚИЫНДЫҚТАР ЖӘНЕ СТРАТЕГИЯЛАР.....	205
<b>А.А. Мұханова, С.К. Кожукаева, Л.Г. Рзаева, Ж.Е. Доумчариева, У.Т. Махажанова</b> МЕДИЦИНАЛЫҚ БЕЙНЕЛЕР НЕГІЗІНДЕ КӨЗ ТОРЫНЫҢ АУРУЛАРЫН ДИАГНОСТИКАЛАУ ҮШІН ТЕРЕҢ ОҚЫТУ МОДЕЛЬДЕРІН ҚОЛДАНУ ЖӘНЕ ТАЛДАУ..	218
<b>Ә.Ж. Омуртаева, У.Т. Махажанова, М.А. Кантуреева, Г. Ускенбаева, Т.Н. Есикова</b> БІЛІМ БЕРУ НЕГІЗІНДЕ АУЫЛ ШАРУАШЫЛЫҒЫ КӘСІПОРЫНДАРЫНЫҢ ИНВЕСТИЦИЯЛЫҚ ТАРТЫМДЫЛЫҒЫН БАҒАЛАУ ӘДІСТЕМЕСІ.....	235
<b>А.Р. Оразаева, Д.А. Тусупов, В. Войчик, А.К. Шайханова, Г.Б. Бекешова</b> МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІМЕН СҮТ БЕЗІ ПАТОЛОГИЯСЫН ТИІМДІ АНЫҚТАУ...	246
<b>Б.Б. Оразбаев, Б.У. Асанова, Ж.Ж. Молдашева, Ж.Е. Шангитова</b> АЙҚЫНСЫЗДЫҚТА КОКСТЕУ РЕАКТОРЛАРЫНЫҢ ЖҰМЫС РЕЖИМДЕРІН КӨПКРИТЕРИЙЛІК ОПТИМИЗАЦИЯЛАУ ЕСЕБІНІҢ ҚОЙЫЛЫМЫ МЕН ОНЫ ШЕШУ ЭВРИСТИКАЛЫҚ ТӘСІЛІ.....	258
<b>Г.А. Салтанова, К.Б. Багитова, Г.А. Дашева, М.Е. Шангитова, Э.Г. Гайсина</b> УНИВЕРСИТЕТ КІТАПХАНАСЫНЫҢ АВТОМАТТАНДЫРЫЛҒАН АҚПАРАТТЫҚ ЖҮЙЕСІН ӨЗІРЛЕУ ЖӘНЕ ЕНГІЗУ: АҚПАРАТТЫҚ РЕСУРСТАРДЫ БАСҚАРУДЫ ОҢТАЙЛАНДЫРУ ЖӘНЕ ПАЙДАЛАНУШЫЛАРҒА ТИІМДІ ҚЫЗМЕТ КӨРСЕТУ.....	269
<b>Л.Т. Салыбек, К.Н. Оразбаева, В.Е. Махатова, Л.Т. Қурмангазиева, Б.Е. Утенова</b> МҰНАЙДЫ АЛҒАШҚЫ ӨНДЕУ ҚОНДЫРҒЫСЫ АТМОСФЕРАЛЫҚ БЛОГЫНЫҢ МОДЕЛЬДЕРІН ТҮРЛІ СИПАТТАҒЫ ҚОЛЖЕТІМДІ АҚПАРАТ НЕГІЗІНДЕ ҚҰРУ.....	285
<b>А. Сейтенов, Т. Жукабаева, С. Ал-Маджид</b> ЭЛЕКТРОНДЫҚ МЕДИЦИНАЛЫҚ ТӨЛҚҰЖАТЫ МЕН ТЕЛЕМЕДИЦИНА АҚПАРАТТЫҚ ЖҮЙЕСІНІҢ МОДЕЛІН ЖОБАЛАУ.....	297
<b>Г.Б. Турмуханова, А.А. Таутенбаева, Г.Т. Бекова, С.Б. Нугуманов, Я. Култан</b> ӘЛЕУМЕТТІК МЕДИА ҚАУЫМДАСТЫҚТАРЫНДАҒЫ ӨЗАРА ІС-ҚИМЫЛ АРҚЫЛЫ УНИВЕРСИТЕТ СТУДЕНТТЕРІНІҢ ЖҰМСАҚ ДАҒДЫЛАРЫН ҚАЛЫПТАСТЫРУ.....	310
<b>А.С. Тынықұлова, А.В. Фаддеев, А.А. Мұханова, А.У. Искалиева, Д.Б. Абулкасова</b> БЕЛГІСІЗДІК ЖАҒДАЙЫНДА ТӘУЕКЕЛДЕРДІ БАСҚАРУДЫ ТАЛДАУ ЖӘНЕ ОҢТАЙЛАНДЫРУ: ЗАМАНАУИ ӘДІСТЕР МЕН ТЕХНОЛОГИЯЛАР.....	325
<b>Ж.Р. Умарова, Г.Ж. Ельбергенава, Н.С. Жуматаев, А.Х. Махатова, С.Б. Ботаева</b> МЕЗОСКОПИЯ ДЕҢГЕЙІНДЕГІ МОЛЕКУЛАЛЫҚ ЕЛЕКТЕРДЕГІ ЗАТ ТАСЫМАЛУЫН ЕСЕПТЕУ АЛГОРИТМІНІҢ ЗИЯЛДЫ ТАЛДАУЫ.....	336

## СОДЕРЖАНИЕ

<b>Н. Абдразакулы, Л. Черикбаева, Н. Мукажанов, Ж. Алибиева</b> СОЗДАНИЕ ЭФФЕКТИВНОГО АЛГОРИТМА ОБРАБОТКИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ АНСАМБЛЕВОГО ПОДХОДА.....	7
<b>Б.Т. Абыканова, А.А. Таугенбаева, А.Г. Амангосова, Г.Т. Бекова, А.Ж. Акматбекова</b> ИНТЕРАКТИВНЫЕ ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ В СОВЕРШЕНСТВОВАНИИ И РАЗВИТИИ САМОСТОЯТЕЛЬНОСТИ ОБУЧАЮЩИХСЯ.....	30
<b>Ж.Ж. Ажибекова, Д.И. Усипбекова, Б.Н. Джаханова, К. Жыланбаева, Ә.Н. Түрсун</b> УДАЛЕНИЯ ОБЛАКОВ И ТУМАННОСТЕЙ С КОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ.....	43
<b>М. Айтимов, Г.Б. Абдикеримова, К.К. Макулов, Б.А. Досжанов, Р.У. Альменаева</b> ИССЛЕДОВАНИЕ ЭМОЦИОНАЛЬНОЙ ТОНАЛЬНОСТИ ТЕКСТА С ПРИМЕНЕНИЕМ АЛГОРИТМОВ МАШИННОГО И ГЛУБОКОГО ОБУЧЕНИЯ.....	57
<b>А.Т. Акынбекова, А.А. Муханова, Salah Al-Majeed, Г.С. Алтаева</b> НЕЧЕТКИЕ МОДЕЛИ ПРИНЯТИЯ РЕШЕНИЙ ОЦЕНКИ СОЦИАЛЬНЫХ ПРОЦЕССОВ РАЗВИТИЯ РЕГИОНА.....	69
<b>К.М. Алдабергенова, А.Б. Касекеева, М.Ж. Айтимов, К.К. Дауренбеков, Т.Н. Есикова</b> СОВЕРШЕНСТВОВАНИЕ МАРКЕТИНГОВОГО УПРАВЛЕНИЯ ЛОГИСТИКОЙ АГРОПРОМЫШЛЕННОГО КОМПЛЕКСА.....	85
<b>А.Е. Абжанова, А.А. Быков, С.К. Сагнаева, Е.А. Абжанов, Д.И. Суржик</b> ОПТИМИЗАЦИЯ МОДЕЛИРОВАНИЯ ГРУНТА С УЧЕТОМ ПОДЗЕМНЫХ ГРУНТОВЫХ ВОД.....	96
<b>А.М. Бисенгалиева, А.У. Исембаева, Т.К. Душаева, Н.М. Алмабаева, Г.О. Ильясова</b> ОХВАТ КЛЮЧЕВЫХ СЛОВ С ПРИМЕНЕНИЕМ СЕМАНТИЧЕСКОГО АНАЛИЗА ДАННЫХ.....	108
<b>А.Х. Давлетова, Н.Н. Оразова, Ж.Б. Сайлау, Д.Н. Курмангалиева, Г.Л. Абдугалимов</b> ПУТИ ПОДГОТОВКИ УЧАЩИХСЯ НАЧАЛЬНЫХ КЛАССОВ К МЕЖДУНАРОДНОМУ ИССЛЕДОВАНИЮ PIRLS С ПОМОЩЬЮ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.....	120
<b>Г. Есмагамбетова, А. Кубигенова, А. Актаева, И. Цэрэн-Онолт, М. Есмагамбет</b> МЕТОДЫ ЗАЩИТЫ БИОМЕТРИЧЕСКИХ ДАННЫХ НА ОСНОВЕ КВАНТОВЫХ ВЫЧИСЛЕНИЙ.....	137
<b>Г.К. Ешмурат, Л.С. Каинбаева</b> МАТЕМАТИЧЕСКАЯ ТРЕВОЖНОСТЬ И ЕЁ ВЛИЯНИЕ НА КАРЬЕРУ БУДУЩИХ УЧИТЕЛЕЙ МАТЕМАТИКИ.....	149
<b>Т.К. Жукабаева, В.А. Десницкий, Е.М. Марденов</b> МЕТОДИКА СБОРА, ПРЕДОБРАБОТКИ И АНАЛИЗА ДАННЫХ В БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЯХ.....	163
<b>А.М. Джумагалиева, А.А. Шекербек, Ж.Ж. Хамитова, М. Свобода, С.А. Калдар</b> ПОВЫШЕНИЕ КИБЕРБЕЗОПАСНОСТИ С ПОМОЩЬЮ АДАПТИВНЫХ СИСТЕМ ОБНАРУЖЕНИЯ АНОМАЛИЙ ПОСРЕДСТВОМ МАШИННОГО ОБУЧЕНИЯ.....	177
<b>А.А. Исмаилова, Г.Е. Мырзабекова, М.Ж. Базарова, Г.Ж. Нурова, Г.Т. Азиева</b> ПРОГНОЗИРОВАНИЕ ЦЕН НА ФОНДОВОМ РЫНКЕ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ	

ГЛУБОКОГО ОБУЧЕНИЯ.....	190
<b>К. Кошанова, Ш. Сапарбайқызы, К.Е. Жангазакова, А.С. Сагынбай, Э. Куриэль-Марин</b>	
МАКСИМАЛЬНОЕ ИСПОЛЬЗОВАНИЕ ПОТЕНЦИАЛА ОБРАЗОВАНИЯ В STEM: ВКЛАД, ПРОБЛЕМЫ И СТРАТЕГИИ ДЛЯ УЛУЧШЕНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ.....	205
<b>А.А. Муханова, С.К. Кожукаева, Л.Г. Рзаева, Ж.Е. Доумчариева, У.Т. Махажанова</b>	
ПРИМЕНЕНИЕ И АНАЛИЗ МОДЕЛЕЙ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ДИАГНОСТИКИ ЗАБОЛЕВАНИЙ СЕТЧАТКИ ГЛАЗА НА ОСНОВЕ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ.....	218
<b>Ә.Ж. Омуртаева, У.Т. Махажанова, М.А. Кантуреева, Г. Ускенбаева, Т.Н. Есикова</b>	
МЕТОДИКА ОЦЕНКИ ИНВЕСТИЦИОННОЙ ПРИВЛЕКАТЕЛЬНОСТИ СЕЛЬСКОХОЗЯЙСТВЕННЫХ ПРЕДПРИЯТИЙ НА ОСНОВЕ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ...235	
<b>А.Р. Оразаева, Д.А. Тусупов, В. Войчик, А.К. Шайханова, Г.Б. Бекешова</b>	
ЭФФЕКТИВНОЕ ВЫЯВЛЕНИЕ ПАТОЛОГИИ МОЛОЧНОЙ ЖЕЛЕЗЫ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ.....	246
<b>Б.Б. Оразбаев, Б.У. Асанова, Ж.Ж. Молдашева, Ж.Е. Шангитова</b>	
ПОСТАНОВКА ЗАДАЧИ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ РЕЖИМОВ РАБОТЫ КОКСОВЫХ РЕАКТОРОВ В УСЛОВИЯХ НЕЧЕТКОСТИ И ЭВРИСТИЧЕСКИЙ МЕТОД ЕЕ РЕШЕНИЯ.....	258
<b>Г.А. Салтанова, К.Б. Багитова, Г.А. Дашева, М.Е. Шангитова, Э.Г. Гайсина</b>	
РАЗРАБОТКА И ВНЕДРЕНИЕ АВТОМАТИЗИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ УНИВЕРСИТЕТСКОЙ БИБЛИОТЕКИ: ОПТИМИЗАЦИЯ УПРАВЛЕНИЯ ИНФОРМАЦИОННЫМИ РЕСУРСАМИ И ОБЕСПЕЧЕНИЕ ЭФФЕКТИВНОГО ОБСЛУЖИВАНИЯ ПОЛЬЗОВАТЕЛЕЙ.....	269
<b>Л.Т. Салыбек, К.Н. Оразбаева, В.Е. Махатова, Л.Т. Курмангазиева, Б.Е. Утенова</b>	
РАЗРАБОТКА МОДЕЛЕЙ АТМОСФЕРНОГО БЛОКА УСТАНОВКИ ПЕРВИЧНОЙ ПЕРЕРАБОТКИ НЕФТИ НА ОСНОВЕ ДОСТУПНОЙ ИНФОРМАЦИИ РАЗЛИЧНОГО ХАРАКТЕРА .....	285
<b>А. Сейтенов, Т. Жукабаева, С. Ал-Маджид</b>	
ПРОЕКТИРОВАНИЕ МОДЕЛИ ИНФОРМАЦИОННОЙ СИСТЕМЫ ТЕЛЕМЕДИЦИНЫ С ЭЛЕКТРОННОЙ МЕДИЦИНСКОЙ КАРТОЙ.....	297
<b>Г.Б. Турмуханова, А.А. Таутенбаева, Г.Т. Бекова, С.Б. Нугуманов, Я. Култан</b>	
ФОРМИРОВАНИЕ МЯГКИХ НАВЫКОВ СТУДЕНТОВ УНИВЕРСИТЕТА ПОСРЕДСТВОМ ВЗАИМОДЕЙСТВИЯ В СООБЩЕСТВАХ СОЦИАЛЬНЫХ СЕТЕЙ.....	310
<b>А.С. Тыныкулова, А.В. Фаддеенков, А.А. Муханова, А.У. Искалиева, А.Б. Абулкасова</b>	
АНАЛИЗ И ОПТИМИЗАЦИЯ УПРАВЛЕНИЯ РИСКАМИ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ: СОВРЕМЕННЫЕ МЕТОДЫ И ТЕХНОЛОГИИ.....	325
<b>Ж.Р. Умарова, Г.Ж. Ельбергенава, Н.С. Жуматаев, А.Х. Махатова, С.Б. Ботаева</b>	
ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ АЛГОРИТМА РАСЧЕТА ПЕРЕНОСА ВЕЩЕСТВА В МОЛЕКУЛЯРНЫХ СИТАХ НА МЕЗОСКОПИЧЕСКОМ УРОВНЕ.....	336



## CONTENTS

<b>N. Abdrazakuly, L. Cherikbayeva, N. Mukazhanov, Zh. Alibiyeva</b> CREATING AN EFFECTIVE IMAGE PROCESSING ALGORITHM BASED ON AN ENSEMBLE APPROACH.....	7
<b>B.T. Abykanova, A.A. Tautenbayeva, A.Γ. Amangosova, G.T. Bekova, A.Zh. Akmatbekova</b> INTERACTIVE EDUCATIONAL TECHNOLOGIES IN IMPROVING AND DEVELOPING STUDENTS' AGENCY.....	30
<b>Zh.Zh. Azhibekova, D.I. Ussipbekova, B. Djakhanova, B.K. Zhylanbaeva, A.N. Tursun</b> REMOVING CLOUDS AND NEBULAE FROM SPACE IMAGES USING MACHINE LEARNING METHOD.....	43
<b>M. Aitimov, G.B. Abdikerimova, K.K. Makulov, B.A. Doszhanov, R.U. Almenayeva</b> STUDY OF THE EMOTIONAL TONE OF A TEXT USING MACHINE AND DEEP LEARNING ALGORITHMS.....	57
<b>A. Akynbekova, A. Mukhanova, Salah Al-Majeed, G. Altayeva</b> FUZZY DECISION MAKING MODELS FOR ASSESSING SOCIAL PROCESSES OF REGIONAL DEVELOPMENT.....	69
<b>K.M. Aldabergenova, A.B. Kassekeyeva, M. Aitimov, K. Daurenbekov, T.N. Esikova</b> IMPROVEMENT OF MARKETING MANAGEMENT OF LOGISTICS OF THE AGRICULTURAL COMPLEX.....	85
<b>A.E. Abzhanova, A.A. Bykov, S.K. Sagnaeva, E.A. Abzhanov, D.I. Surzhik</b> OPTIMIZATION OF SOIL MODELING WITH CONSIDERATION OF UNDERGROUND GROUNDWATER.....	96
<b>A.M. Bissengaliyeva, A.U. Issembayeva, T.K. Dushayeva, N.M. Almabayeva, G.O. Ilyassova</b> KEYWORD COVERAGE USING SEMANTIC DATA ANALYSIS.....	108
<b>A.Kh. Davletova, N.N. Orazova, Zh.B. Sailau, D.N. Kurmangalieva, G.L. Abdugalimov</b> WAYS TO PREPARE PRIMARY SCHOOL STUDENTS FOR INTERNATIONAL PIRLS RESEARCH USING INFORMATION TECHNOLOGY.....	120
<b>G. Yesmagambetova, A. Kubigenova, A. Aktayeva, I. Tseren-Onolt, M. Esmaganbet</b> METHODS OF BIOMETRIC DATA PROTECTION BASED ON QUANTUM COMPUTING.....	137
<b>G.K. Yeshmurat, L.S. Kainbayeva</b> UNDERSTANDING MATH ANXIETY AND ITS IMPACT ON MATH EDUCATION STUDENTS' CAREERS.....	149
<b>T.K. Zhukabayeva, V.A. Desnitsky, E.M. Mardenov</b> A TECHNIQUE FOR COLLECTION, PREPROCESSING AND ANALYSIS OF DATA IN WIRELESS SENSOR NETWORKS.....	163
<b>A.M. Jumagaliyeva, A.A. Shekerbek, Zh.Zh. Khamitova, M. Svoboda, S. Kaldar</b> ENHANCING CYBERSECURITY WITH ADAPTIVE ANOMALY DETECTION SYSTEMS THROUGH MACHINE LEARNING.....	177
<b>A.A. Ismailova, G. Murzabekova, M.Zh. Bazarova, G.Zh. Nurova, G.T. Azieva</b> FORECASTING PRICES IN THE STOCK MARKET USING DEEP LEARNING METHODS.....	190

<b>G. Kochshanova, Sh. Saparbaykyzy, K.Y. Zhangazakova, A.S. Sagynbay, E. Curiel-Marin</b> MAXIMIZING THE POTENTIAL OF STEM EDUCATION: CONTRIBUTIONS, CHALLENGES, AND STRATEGIES TO IMPROVE LEARNING OUTCOMES.....	205
<b>A.A. Mukhanova, S.K. Kozhukaeva, L.G. Rzayeva, Zh.E. Doumcharieva, U.T. Makhazhanova</b> APPLICATION AND ANALYSIS OF DEEP LEARNING MODELS FOR DIAGNOSIS OF RETINAL DISEASES FROM MEDICAL IMAGES.....	218
<b>A. Omurtayeva, U. Makhazhanova, M. Kantureyeva, G. Uskenbayeva, T.N. Esikova</b> METHODOLOGY FOR ASSESSING THE INVESTMENT ATTRACTIVENESS OF AGRICULTURAL ENTERPRISES BASED ON THE PRESENTATION OF KNOWLEDGE.....	235
<b>A.R. Orazayeva, J.A. Tussupov, W. Wójcik, A.K. Shaikhanova, G.B. Bekeshova</b> EFFECTIVE DETECTION OF BREAST PATHOLOGY USING MACHINE LEARNING METHODS.....	246
<b>B.B. Orazbayev, B.U. Asanova, Zh.Zh. Moldasheva, Zh.E. Shangitova</b> FORMULATION OF THE PROBLEM OF MULTICRITERIAL OPTIMIZATION OF OPERATING MODES OF COKE REACTORS UNDER FUZZY CONDITIONS AND A HEURISTIC METHOD FOR ITS SOLUTION.....	258
<b>G.A. Saltanova, K.B. Bagitova, G.A. Dasheva, M.E. Shangitova, E.G. Gaisina</b> DEVELOPMENT AND IMPLEMENTATION OF AN AUTOMATED UNIVERSITY LIBRARY INFORMATION SYSTEM: INFORMATION RESOURCE MANAGEMENT OPTIMIZATION AND EFFECTIVE USER SERVICE PROVISION.....	269
<b>L. Salybek, K. Orazbayeva, V. Makhatova, L. Kurmangazieva, B. Utenova</b> DEVELOPMENT OF MODELS OF THE ATMOSPHERIC BLOCK OF A PRIMARY OIL PROCESSING PLANT BASED ON AVAILABLE INFORMATION OF VARIOUS NATURE.....	285
<b>A. Seitenov, T. Zhukabayeva, S. Al-Majeed</b> DESIGNING A MODEL OF A TELEMEDICINE INFORMATION SYSTEM WITH ELECTRONIC MEDICAL RECORD.....	297
<b>G.B. Turmukhanova, A.A. Tautenbayeva, G.T. Bekova, S.B. Nugumanov, K. Yaroslav</b> FORMATION OF UNIVERSITY STUDENTS' SOFT SKILLS THROUGH INTERACTION I N SOCIAL NETWORKING COMMUNITIES.....	310
<b>A.S. Tynykulova, A.V. Faddeenkov, A.A. Mukhanova, A. Iskaliyeva, D.B. Abulkassova</b> ANALYSIS AND OPTIMIZATION OF RISK MANAGEMENT IN CONDITIONS OF UNCERTAINTY: MODERN METHODS AND TECHNOLOGIES.....	325
<b>Zh. Umarova, G. Yelbergenova, N. Zhumatayev, A. Makhatova, S. Botayeva</b> INTELLIGENT ANALYSIS OF SUBSTANCE TRANSPORT ALGORITHM IN MOLECULAR SIEVES AT THE MESOSCOPIC LEVEL.....	336

**Publication Ethics and Publication Malpractice  
the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct ([http://publicationethics.org/files/u2/New\\_Code.pdf](http://publicationethics.org/files/u2/New_Code.pdf)). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

**[www.nauka-nanrk.kz](http://www.nauka-nanrk.kz)**

**<http://physics-mathematics.kz/index.php/en/archive>**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Подписано в печать 15.06.2024.

Формат 60x881/8. Бумага офсетная. Печать-ризограф.

21,0 п.л. Тираж 300. Заказ 2.