

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



«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ

# Х А Б А Р Л А Р Ы

---

---

**ИЗВЕСТИЯ**

РОО «НАЦИОНАЛЬНОЙ  
АКАДЕМИИ НАУК РЕСПУБЛИКИ  
КАЗАХСТАН»

**N E W S**

OF THE ACADEMY OF SCIENCES  
OF THE REPUBLIC OF  
KAZAKHSTAN

**SERIES**  
**PHYSICS AND INFORMATION TECHNOLOGY**

**3 (351)**

**JULY – SEPTEMBER 2024**

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

### **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 3. Number 351 (2024). 25-36

<https://doi.org/10.32014/2024.2518-1726.289>

УДК 28.23.29

MPHTH 28.17.2

**A.A. Abibullayeva, A.S. Baimakhanova.**

Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan,  
Kazakhstan. [aygerim.baymakhanova@ayu.edu.kz](mailto:aygerim.baymakhanova@ayu.edu.kz)

## USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES IN KEYWORD EXTRACTION

**Abibullayeva Aiman**- PhD, senior lecturer, Khoja Akhmet Yassawi International Kazakh-Turkish University, E-mail: [aiman.abibullayeva@ayu.edu.kz](mailto:aiman.abibullayeva@ayu.edu.kz). [https://orcid.org/\(0000-0003-2449-2540\)](https://orcid.org/(0000-0003-2449-2540));

**Baimakhanova Aygerim** – master of science, teacher, Khoja Akhmet Yassawi International Kazakh-Turkish University, E-mail: [aygerim.baymakhanova@ayu.edu.kz](mailto:aygerim.baymakhanova@ayu.edu.kz). [https://orcid.org/\(0000-0002-5364-0146\)](https://orcid.org/(0000-0002-5364-0146));

**Abstract.** With the rapid development of Internet technology and the rapid spread of news content in recent times, the importance of extracting keywords from text is increasing. The use of keywords related to such needs allows the reader to easily and quickly access the necessary information. Keyword extraction is an important research area in text processing, natural language processing, and information retrieval. Due to the importance of the keyword and the difficulty of manually tagging it, automatic keyword generation allows you to get significant results. Keyword extraction, which is one of the sub-topics of natural language processing, includes automatic indexing, automatic aggregation, automatic classification, automatic clustering, and automatic filtering. In the fields of natural language processing, artificial intelligence-based methods began to be used in the early 2000s, thanks to machine learning and deep learning approaches. There are many methods and algorithms proposed for keyword extraction. These methods can be grouped under two main headings: supervised and unsupervised approaches. Currently, machine learning and deep learning methods and natural language processing research are limited to the Kazakh language. Since there is no corpus representing the Kazakh language, there are no studies on the extraction of keywords to date. This area requires more effort due to the significant increase in the amount of online Kazakh content and the rarity of manually entered keywords in existing electronic documents. This study provides an overview of machine learning techniques and deep learning techniques for keyword extraction and reviews important research.

**Keywords:** keyword extraction, natural language processing, machine learning, deep learning, supervised method, unsupervised method.

**А.А. Абибуллаева, А.С. Баймаханова**

Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті,  
Түркістан, Қазақстан.

aygerim.baymakhanova@ayu.edu.kz

## **КІЛТТІК СӨЗДЕРДІ ШЫҒАРУДА МАШИНАЛЫҚ ЖӘНЕ ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ**

**Абибуллаева Айман Абибуллақызы** - PhD, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан, E-mail: aiman.abibullayeva@ayu.edu.kz. [https://orcid.org/\(0000-0003-2449-2540\)](https://orcid.org/(0000-0003-2449-2540))

**Баймаханова Айгерим Саттаровна**-магистр оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан, Қазақстан. E-mail: aygerim.baymakhanova@ayu.edu.kz. [https://orcid.org/\(0000-0002-5364-0146\)](https://orcid.org/(0000-0002-5364-0146)).

**Аннотация.** Интернет технологиясының қарқынды дамуымен және соңғы уақытта жаңалықтар мазмұнының жылдам таралуымен мәтіннен кілттік сөздерді шығару маңыздылығы артып келеді. Осындай қажеттіліктерге байланысты кілттік сөздерді қолдану оқырманға қажетті ақпаратқа оңай және жылдам қол жеткізуге мүмкіндік береді. Кілттік сөзді шығару мәтінді өңдеудегі, табиғи тілді өңдеудегі және ақпаратты іздеудегі маңызды зерттеу бағыты болып табылады. Кілттік сөздің маңыздылығына және оны қолмен белгілеудің қиындығына байланысты, кілттік сөздерді автоматты түрде шығару маңызды нәтижелерді алуға мүмкіндік береді. Табиғи тілді өңдеудің қосалқы тақырыптарының бірі болып табылатын кілттік сөзді шығаруға автоматты индекстеу, автоматты жинақтау, автоматты классификация, автоматты кластерлеу, автоматты сүзгілеу кіреді. Табиғи тілді өңдеу салаларында машиналық оқыту мен терең оқыту тәсілдерінің арқасында жасанды интеллектке негізделген әдістер 2000 жылдардың басында қолданыла бастады. Кілтсөзді шығару үшін ұсынылған көптеген әдістер мен алгоритмдер бар. Бұл әдістерді екі негізгі тақырып бойынша топтастыруға болады: бақыланатын және бақыланбайтын тәсілдер. Қазіргі таңда қазақ тілі үшін машиналық оқыту және терең оқыту әдістері және табиғи тілді өңдеу зерттеулері шектеулі. Қазақ тілін танытатын корпус болмағандықтан, бүгінгі күнге дейін кілттік сөздерді шығару бойынша зерттеулер жоқ. Бұл сала онлайн қазақша контент көлемінің айтарлықтай ұлғаюына және қолданыстағы электрондық құжаттарға қолмен берілген кілт сөздердің сирек болуына байланысты көбірек күш салуды қажет етеді. Бұл зерттеуде кілттік сөздерді шығарудың машиналық оқыту әдістері және терең оқыту әдістері туралы жалпы ақпарат берілген және маңызды зерттеулер қарастырылған.

**Түйін сөздер:** кілттік сөз шығару, табиғи тілді өңдеу, машиналық оқыту, терең оқыту, бақылатын әдіс, бақыланбайтын әдіс.



**А.А. Абибуллаева, А.С. Баймаханова.**

Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан.  
aygerim.baymakhanova@ayu.edu.kz

## **ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО И ГЛУБОКОГО ОБУЧЕНИЯ ПРИ ИЗВЛЕЧЕНИИ КЛЮЧЕВЫХ СЛОВ**

**Абибуллаева Айман Абибуллаевна** – PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: aiman.abibullayeva@ayu.edu.kz. [https://orcid.org/\(0000-0003-2449-2540\)0000-0003-1800-7350](https://orcid.org/(0000-0003-2449-2540)0000-0003-1800-7350);  
**Баймаханова Айгерим Саттаровна** – магистр, преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан, E-mail: aygerim.baymakhanova@ayu.edu.kz. [https://orcid.org/\(0000-0002-5364-0146\)](https://orcid.org/(0000-0002-5364-0146)).

**Аннотация.** В связи с быстрым развитием интернет-технологий и быстрым распространением новостного контента в последнее время важность извлечения ключевых слов из текста возрастает. Использование ключевых слов, связанных с такими потребностями, позволяет читателю легко и быстро получить доступ к необходимой информации. Извлечение ключевых слов — важная область исследований в области обработки текста, естественного языка и поиска информации. Из-за важности ключевого слова и сложности его маркировки вручную автоматическая генерация ключевых слов позволяет получить значительные результаты. Извлечение ключевых слов, которое является одной из подтем обработки естественного языка, включает в себя автоматическое индексирование, автоматическое агрегирование, автоматическую классификацию, автоматическую кластеризацию, автоматическую фильтрацию. В области обработки естественного языка методы на основе искусственного интеллекта начали использоваться в начале 2000-х годов благодаря подходам машинного обучения и глубокого обучения. Для извлечения ключевых слов предложено множество методов и алгоритмов. Эти методы можно сгруппировать под двумя основными заголовками: контролируемые и неконтролируемые подходы. В настоящее время методы машинного и глубокого обучения, а также исследования в области обработки естественного языка ограничены для казахского языка. Поскольку корпуса, представляющего казахский язык, не существует, исследования по выделению ключевых слов на сегодняшний день отсутствуют. Эта область требует больше усилий из-за значительного увеличения количества онлайн-казахского контента и редкости ключевых слов, вводимых вручную в существующих электронных документах. В этом исследовании представлен обзор методов машинного обучения и методов глубокого обучения для извлечения ключевых слов, а также рассмотрены важные исследования.

**Ключевые слова:** извлечение ключевых слов, обработка естественного языка, машинное обучение, глубокое обучение, контролируемый метод, неконтролируемый метод.

## **Кіріспе**

Интернеттегі ақпаратты шолу күнделікті өмірде компьютер пайдаланушылары үшін әдеттегі әрекетке айналды. Күн өткен сайын ақпараттар тобында қажетті ақпаратқа қол жеткізу қиындай түсуде. Интернетте күн сайын мыңдаған интернет жаңалықтары жарияланып отырғандықтан, тиісті құжатты тиімді түрде алу және қорытындылау қиынға соғады. Сондықтан белгілі бір веб-беттің негізгі мазмұнын қамтамасыз ету үшін кілттік сөзді немесе негізгі сөз тіркестерін шығару әдісі қолданылады. Кілттік сөздерді шығару – құжаттың тақырыбын тиісті түрде көрсете алатын құжаттан негізгі терминдерді, негізгі сөз тіркестерін, негізгі бөлімдерді немесе кілттік сөздерді анықтау процесі (Paragiannopoulou, E., Tsoumakas, G., 2020). Жаңалық белгілерін жасау, ғылыми мақалаларға кілт сөздерді тағайындау және құжаттарды қорытындылау сияқты көптеген күнделікті тапсырмалар қымбат және уақытты қажет етеді. Бұл тапсырмалар негізгі ойды сипаттайтын ең кішкентай бөліктер болып табылатын дұрыс кілт сөздерді табуға байланысты (H. Kılıç, A. Çetin, 2023). Кілт сөзді шығару мәтінді өңдеудің маңызды бөлігі болып табылады. Қолмен кілт сөзді шығару көп уақытты алады, сондықтан табиғи тілді өңдеу (Natural Language Processing - NLP) саласындағы зерттеушілер осы процессті автоматтандыру әдістерін дайындауды басты назарда ұстап отыр.

Әдебиеттерге көз жүгіртсек, түйінді сөз дегеніміз;

- ең маңызды ақпаратты қамтитын терминдер (Kaur, et all, 2010),
- оқырманға мазмұн туралы жиынтық ақпарат беретін терминдер жинағы (Liu, et.all, 2010),
- құжаттың негізгі тақырыптарын қамтитын сөздер (Kim, et all, 2013),
- мәтіннің негізгі идеясын немесе тақырыбын қамтитын терминдер (Awajan, 2014),
- мәтінді барынша анық көрсететін сөздер немесе сөз топтары (Birdevrim, 2018),
- құжатта талқыланатын негізгі тақырыпты қамтитын өрнектер (Basaldella, 2018),
- құжат мазмұнының барлық маңызды аспектілерін білдіретін сөздер (Paragiannopoulou, et all, 2020),
- мәтін мазмұны туралы мәлімет беретін сөздер (Kılıç Ünlü, et all, 2023).

Мәтінді өңдеудің ішкі тақырыптары мәтінді қорытындылау және мәтінді сипаттайтын кілттік сөздерді анықтау сияқты қосымшаларды қамтиды. Мәтіндердің мағынасын ашу және бұл мағынаны әрекетке айналдыру табиғи тілді өңдеу саласындағы басқа бір зерттеу тақырыбы болып табылады. Кілт сөзді шығару ақпарат іздеу жүйелері, сандық кітапхананы іздеу, веб-мазмұнды басқару, құжаттарды кластерлеу және мәтінді қорытындылау сияқты табиғи тілді өңдеудің әртүрлі қолданбаларында қолданылған. Қолмен белгіленген немесе автоматты түрде таңдалған кілттік сөз топтары пайдаланушыға мәтін мазмұнының жалпы идеясын жасау үшін қолданылады. Негізгі сөздер мен



сөз тіркестері мәтіннің мазмұны туралы ең маңызды ақпаратты қамтитын мәтіннің құрылымдық бірліктері ретінде қарастырылады. Кілттік сөздер мен сөз тіркестерін алуды автоматтандыру мәселесін шешу үшін екі негізгі тәсілді ажырату қажет: кілттік сөз тіркесін тағайындау (keyphrase assignment) және кілттік сөз тіркесін алу (keyphrase extraction). Кілт сөзді шығаруда құжаттағы сөздер жиілік пен ұзындық сияқты ерекшеліктерге негізделген мағыналы болып көрінетін сөздерді анықтау үшін талданады. Кілттік сөздерді тағайындау кезінде кілт сөздер терминдер сөздігінен таңдалады және құжаттар мазмұны негізінде лексикалық элементтерге сәйкес келетін сыныптарға жіктеледі. Кілттік сөзді шығару мәтіннен тікелей негізгі ақпаратты таңдауды қамтиды. Екі тәсіл де машиналық оқыту әдістерін пайдаланады және негізгі сөз тіркестері бар оқыту құжаттары қажет етеді (Nguyen, et all, 2007).

Бұл зерттеу жұмысымыздың барысында кілт сөзді шығарудың әртүрлі әдістері мен алгоритмдері зерттелген. Машиналық оқыту және терең оқыту әдістері туралы ақпарат беріледі, басқа тілдердегі зерттеулер туралы айтылады, негізгі тақырыптар және осы саладағы бұрын зерттелген әдістерді қарастырамыз.

### Әдістер мен материалдар

Кілттік сөздерді шығарудың машиналық әдістері. Зерттеу барысында кілттік сөзді шығаруда машиналық және терең оқыту әдістеріне әдебиеттерге талдау жүргізілді. Кілттік сөзді шығару мәтінді өңдеудегі, табиғи тілді өңдеудегі және ақпаратты іздеудегі маңызды зерттеу бағыты болып табылады. Кәсіби индекстеушілермен кілт сөздерді нақты анықталған стандарттарға және жиі қолданылатын сөздіктерге немесе басқарылатын сөздіктерге сәйкес қолмен шығару көп уақытты қажет етеді, жалықтырады, қымбат, субъективті және практикалық емес. Сондықтан құжаттардан кілт сөздерді шығару үшін автоматты әдістер қажет. Кілттік сөзді шығару – құжаттың негізгі мазмұнын сипаттай алатын маңызды сөздердің немесе терминдердің шағын тобын анықтайтын жүйелі процесс. Siddiqi S., Sharan A., (2015) кілттік сөздерді шығару тәсілдерін 4 негізгі тақырып бойынша топтаған:



Сурет 1. Кілт сөздерді автоматты түрде алудың классификациясы.

- *Лингвистикалық әдістер:* Бұл әдістер әдетте ережеге негізделген және сөйлемдер мен құжаттарды тілдік ерекшеліктерінің негізі ретінде пайдаланады. Бұл есептеу қарқынды және лингвистикалық домен сарапшысынан басқа домен білімін қажет етеді. Бұл әдістер негізінен сөйлемдер мен құжаттардағы сөздердің тілдік ерекшеліктерін пайдаланады. Лингвистикалық әдіс-лексикалық талдау, синтаксистік талдау – дискурстық талдау, т.б. кіреді.

- *Статистикалық әдістер:* Статистикалық әдістер әдетте лингвистикалық корпусқа және осы корпусдан алынған статистикалық белгілерге негізделеді. Олардың ең айқын артықшылығы - олар тілге тәуелсіз және бірдей әдісті бірнеше тілде қолдануға болады. Кілттік сөзді шығару үшін ең жиі қолданылатын статистикалық әдістер сөз жиілігі, термин жиілігі (Term Frequency-TF), термин жиілігі – кері құжат жиілігі (Term Frequency-Inverse Document Frequency-TF-IDF), сөздің бірігуі және PAT-Tree алгоритмдері болып табылады.

- *Машиналық оқыту әдістер:* Машиналық оқыту әдістері негізінен бақыланатын оқыту әдістерін пайдаланады. Бұл әдістерде үлгіні үйрену үшін оқу құжаттары алынады, содан кейін модель тестілеу модулі арқылы тексеріледі. Жақсы нәтижелерге қол жеткізгеннен кейін ол жаңа құжаттардан кілт сөздерді табу үшін қолданылады. Бұл әдіс Найве Байес алгоритмдерін (Naive Bayes), Қолдау векторлық машиналарын (SVM), Кластерлеу алгоритмдерін (K-Means, K-Medians) және жіктеу алгоритмдерін қамтиды. Дегенмен, бақыланатын оқыту әдістерінде бізде таңбаланған құжаттар болуы керек. Мұндай корпус болмаған жағдайда бақылаусыз және жартылай бақыланатын оқыту әдістері қолданылады.

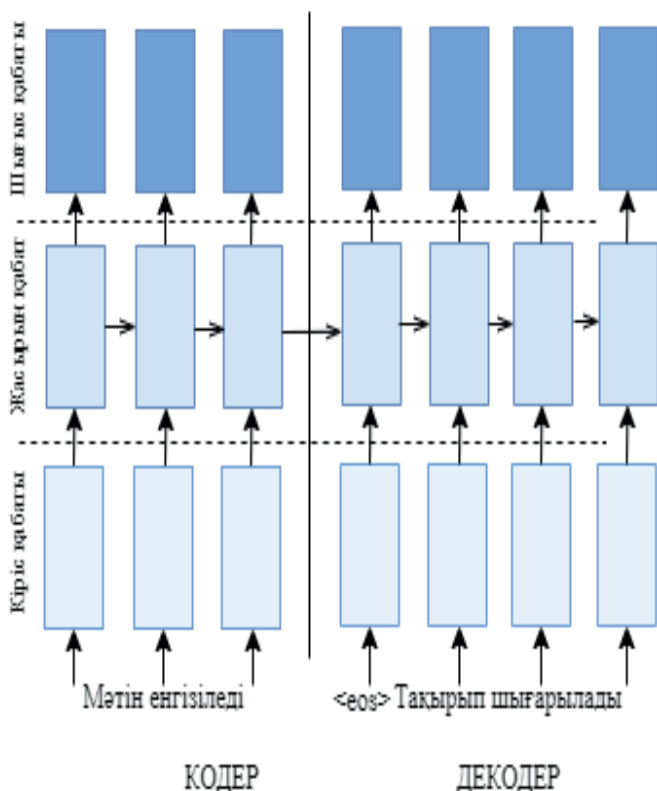
- *Гибридті әдістер:* Кілтсөзді шығарудың басқа әдістері негізінен жоғарыда аталған әдістерді біріктіреді немесе сөздердің орналасуын, ұзындығын, орналасу мүмкіндігін, сөздердің айналасындағы Гипермәтіндік белгілеу тілі (HyperText Markup Language - HTML) тегтерін және т.б. сияқты кейбір интуитивті ақпаратты пайдаланады (Abibullayeva A. A., 2023).

#### *Кілттік сөздерді шығаруда қолданылатын терең оқыту алгоритмдері*

Табиғи тілді өңдеуде мәтінді қорытындылау және жіктеу сияқты мәселелерді шешуде терең оқыту алгоритмдері сәтті қолданылып келеді. Осы мақсатта мәтіндерден тақырып алу, кілттік сөздерді анықтау, сұрақ-жауап жүйесін жасау сияқты мәселелерді шешу үшін басқа тілдерде зерттеулер жүргізілген. Терең оқыту - бұл машиналық оқытудың бір тармағы. Терең оқыту саласында әзірленген архитектуралардың саны көбейіп, қолданыстағы статистикалық және графикалық әдістерден табысты нәтижелерге қол жеткізілгендіктен, терең оқыту әдістері әртүрлі мәселелерді шешу үшін қолданыла бастады. Ағылшын тілінде және басқа тілдерде кілт сөздерді алу бойынша көптеген зерттеулер жүргізілгенімен, қазақ тілінде кілттік сөздерді шығару мәселесі әлі қолға алынғанбаған. Қазақ тілі үшін машиналық оқыту және терең оқыту әдістері және табиғи тілді өңдеу зерттеулері шектеулі.

Қазақ тіліндегі мәтіндерден кілттік сөздерді алу үшін тереңдетіп оқытылатын үлгі әлі жоқ. Дегенмен қазақ тілінің құрылымдық жүйесін басқа тілдермен салыстырып, оның жеке ұқсастықтары мен айырмашылықтарын көрсететін салыстырмалы-типологиялық бағытта көптеген ғылыми зерттеулер бар.

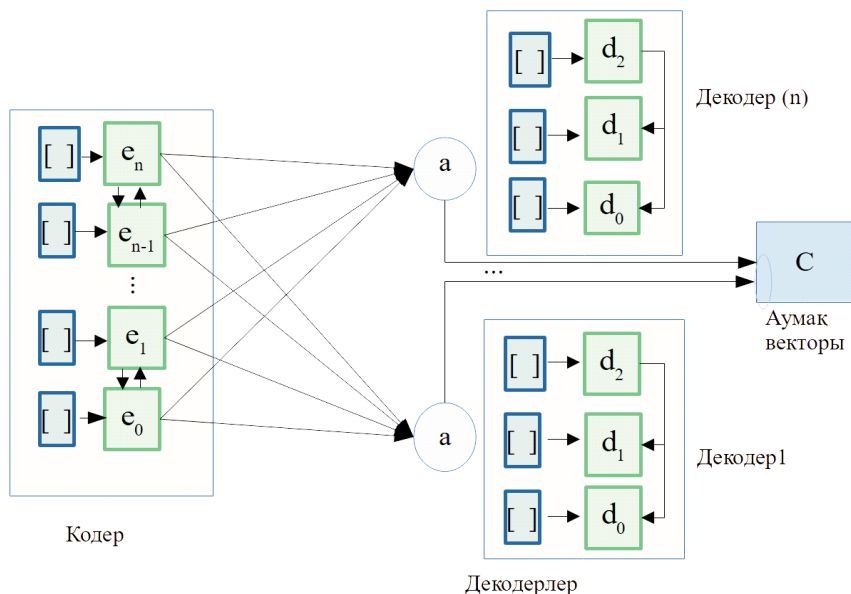
Жақында терең оқыту алгоритмдері Ұзақ қысқа мерзімді жад (Long Short-Term Memory-LSTM, екі бағытты ұзақ қысқа мерзімді жады (Bi-LSTM), конволюциялық нейрондық желі (Convolutional neural network-CNN), қайталанатын нейрондық желі Recurrent neural network-RNN) сөйлеу/мәтін, кескін және бейне өңдеуде өте жақсы нәтижелер беруде. Кілт сөзді шығару үшін ұсынылған әдістер негізінен Sequence-to-sequence моделі (Seq2Seq) кодер/декодер (Encoder/Decoder) құрылымын пайдалану арқылы мәселені шешуге тырысты. Seq2Seq архитектурасы машиналық аударма, кескінді тегтеу, сезімді талдау, жаңалықтар тақырыбын шығару, негізгі фразаны шығару сияқты көптеген мәселелерде қолданылады. Seq-to-seq моделі – бірнеше кірістен бірнеше шығыс шығарылатын кезде қолданылатын RNN моделі. 2-суретте seq-to-seq RNN моделі көрсетілген (Lopyrev, et all, 2015). Seq-to-seq моделі екі RNN желісінен тұрады. Олар кодер және декодер деп аталады.



Сурет 2. Кодер/декодер архитектурасының реттілігі

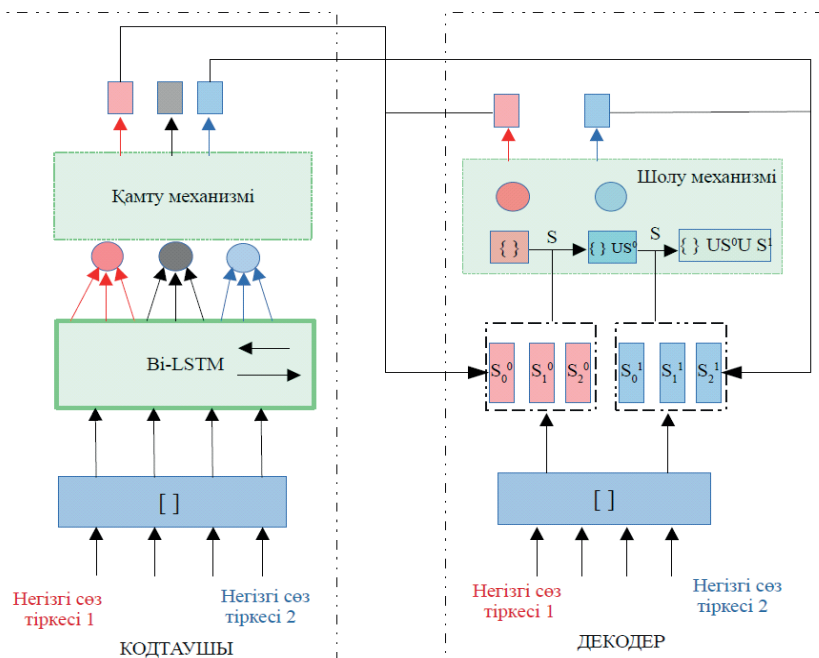
Қайталанатын нейрондық желіні қамту (Coverage RNN) архитектурасында үлгі маңызды механизмі және транскрипция механизмі бар кілт сөзді шығару үшін seq2seq ұсынылды (Zhang, et all, 2018). CovRNN үлгісінде бастапқы мәтіндік ақпарат  $c^t$  векторымен сақталды және  $c^t$  алдыңғы шешу қадамдарындағы барлық назар бөлулердің қосындысы арқылы есептелді. Қамту механизмі алдыңғы қадамнан шешу фазасындағы маңыздылық үлестірімінің жиынтық сомасы болып табылады. Ол төменде көрсетілген формула бойынша есептеледі:

$$c^t = \sum_{t_s=0}^{t-1} a^{t_s} \quad (1)$$



Сурет 3. CovRNN моделі

Өз зерттеулерінде Chen et al., (2018) Корреляциялық қайталанатын нейрондық желі (Correlation RNN) деп аталатын кілт сөздерді генерациялауға арналған жаңа seq2seq архитектурасын ұсынды, ол репликация және қамту сұрақтарын шешудің екі әдісімен бірнеше кілт сөздер арасындағы корреляцияны анықтайды. Бұл мәселелерді шешу үшін Бұл әдіс бірнеше кілт сөздер арасындағы корреляцияны екі жолмен түсіреді; қамту векторы алдыңғы кілт сөздің және бастапқы мәтіннің жинақталғанын түсіну үшін пайдаланылады, ал шолу механизмі бірдей мағынаны білдіретін қайталанатын сөз тіркестерін жою және кілт сөздерден қосымша ақпаратты болдырмау үшін қолданылады. Шолу механизмі назар аудару механизміне өте ұқсас деп есептелген. 4-суретте көрсетілген бұл архитектура seq2seq үлгісін кеңейтеді және кілт сөзді генерациялауда бір-көп қатынастарын түсіреді.



Сурет 4. CorrRNN моделі

Соңғы уақытта мәтіндердегі кілттік сөздерді анықтау үшін ұзақ қысқа мерзімді жад (LSTM) және Жабық қайталау бірлігі (Gated Repeating Unit-GRU) архитектуралары қолданыла бастады. GRU және LSTM-де ұяшық жады жасырын қабатпен бірге келесі ұяшыққа тасымалданады. GRU кодер және декодер ретінде пайдаланылады. LSTM-CNN (Song et al., 2019) негізіндегі басқа қорытындылау алгоритмі хэш өнімділігін жақсарту үшін әзірленді. Кодер өрнектегі сөздерді CNN қабаты арқылы өткізу арқылы LSTM кірісіне жібереді. Басқа зерттеуде LSTM-CNN негізіндегі архитектура (Meem et al., 2019) қолданылды, сөздерді енгізу векторларының бастапқы мәндері fastText көмегімен инициализацияланды. Басқа бір мақалада кілттік сөзді шығару үшін қазақ тіліндегі газет жаңалықтарынан құралған деректер жинағынан кілт сөздерді алу үшін BERT және BERT base-multilingual-uncased тіл үлгісін пайдалану арқылы алынды (Abibullayeva, et all, 2022).

**Нәтижелер және оларды талқылау**

Модельдер үшін кілт сөздерді шығару архитектурасын бағалау дәлдікті өлшеу үшін өте маңызды. Жалпы мәліметтер жиынтығын қолдана отырып бағалау кілт сөздерді шығару үшін әдебиетке қосқан оң үлесін өлшейді. Құрылған модельдердің дәлдігін өлшеудің еске түсіру (recall), дәлдік (accuracy), нақтылық (precision) F-ұпай (F-score) бағалау әдістері бар. Бұл көрсеткіштердің барлығы шатасу матрицасы (confusion matrix) арқылы есептелді. Нағыз оң (True Positive-TP) - кілттік сөз ретінде дұрыс белгіленген

сөздердің саны, Жалған оң (False Positive-FP) - кілт сөзсіз үлгі бойынша кілт сөз ретінде белгіленген сөздердің саны, Жалған теріс (False Negative-FN) - кілт сөз болмаса да, модель кілт сөз ретінде белгіленген сөздердің саны, және Нағыз теріс (True Negative -TN) - кілттік емес сөздер ретінде дұрыс белгіленген сөздердің саны. 2-тендеуде модельдің дәлдік - дұрыс оң және теріс болжамдардың арақатынасын білдіреді.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

Нақтылық - шынайы позитивтердің жалпы оң болжамдарға нақты қатынасын есептейді.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Еске түсіру - оң мәндердің дұрыс болжалатын арақатынасын есептейді.

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

Модельдің өнімділігін бағалау үшін дәлдік және нақтылық арасындағы гармоникалық орташа мән болып табылатын  $F_1$ - ұпайы өлшенеді.

$$F_1 score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Кесте 1 – Шатасу матрицасы

	<i>Болжамды оң</i>	<i>Болжамды теріс</i>
<i>Нақты оң</i>	Нағыз оң (True Positive)	Жалған теріс (False Negative)
<i>Нақты теріс</i>	Жалған оң (False Positive)	Нағыз теріс (True Negative)

Модельдің өнімділігін дәл өлшеу және бағалау машиналық оқытудың миссиясы үшін өте маңызды. Дұрыс таңдалған көрсеткіштерді бағалау арқылы үлгіні жақсарту және шындыққа барынша жақын болжамдар жасау үшін жұмыс істеу оңайырақ болады.

### **Қорытынды**

Мәтін ішінен ең қысқаша ақпаратты жылдам алуға болатын кілтті сөздер бүгінгі күні деректер көлемі экспоненциалды түрде ұлғайған кезде шешуді қажет ететін маңызды мәселе болып табылады. Кілтсөзді/фразаны шығару мәтінді мағынаға қатысты көптеген мәселелерді шешу үшін қолданылады. Бұл ақпаратты алу, тақырыпты қадағалау және қорытындылау мәселелерінің қадамы болғанымен, жаңалықтар мәтіндері мен ғылыми мақалаларға түйінді сөздерді тағайындау мәселелерінде шешуді қажет ететін негізгі жағдай болып табылады. Түйінді сөздер мен түйінді фразаларды алу үшін соңғы жылдары



эртүрлі зерттеушілер ұсынған эртүрлі тәсілдерге шолу жасалды және талданатын мәтіндегі маңыздылығына қарай үміткер кілт сөздері мен түйінді фразаларды бағалау үшін әдетте қолданылатын эртүрлі мүмкіндіктерді таңдау метрикасының қысқаша сипаттамасы жасалды. Сауалнама жұмысы негізгі санаттарға бөлінді, мысалы, бақылаудағы, бақыланбайтын, жартылай бақыланатын және түйінді сөздерді алудың статистикалық тәсілдері және сол салаларда жасалған негізгі жұмыстар хронологиялық түрде тізімделді.

#### Әдебиеттер

Абибуллаева, А., & ЧЕТИН, А. (2022). Извлечение ключевых слов из набора данных новостей Казахстана с помощью BERT. Эль-Сезери, 9 (4), 1193-1200.

Сиддики, С., & Шаран, А. (2015). Методы извлечения ключевых слов и фраз: обзор литературы. Международный журнал компьютерных приложений, 109 (2).

Килич, Х., & Четин, А. (2023). Новая модель классификации лексем ансамбля на основе графа для извлечения ключевых слов. Арабский журнал науки и техники, 48 (8), 10673-10680.

Папагианнопулу, Э., & Цумакас, Г. (2020). Обзор извлечения ключевых фраз. Междисциплинарные обзоры Wiley: Интеллектуальный анализ данных и раскрытие знаний, 10 (2), e1339.

Нгуен, Т. Д., и Кан, М. Ю. (2007, декабрь). Извлечение ключевых фраз в научных публикациях. На Международной конференции по азиатским цифровым библиотекам (стр. 317-326). Берлин, Гейдельберг: Springer Berlin Heidelberg.

Чжан, Ю., и Сяо, В. (2018). Генерация ключевых фраз на основе модели deep seq2seq. Доступ к IEEE, 6, 46047-46057.

Chen, J., Zhang, X., Wu, Y., Yan, Z., & Li, Z. (2018). Генерация ключевых фраз с ограничениями корреляции. Препринт arXiv arXiv: 1808.07185.

Мим, Н. Т. А., Чоудхури, М. М. Х., и Рахман, М. М. (2019, декабрь). Подход машинного обучения для извлечения ключевых фраз из бенгальского документа с использованием CNN-двунаправленного LSTM. В 2019 году состоится 22-я Международная конференция по компьютерным и информационным технологиям (ICCSIT) (стр. 1-5). IEEE.

Лопырев К. (2015). Генерация заголовков новостей с помощью рекуррентных нейронных сетей. Препринт arXiv arXiv: 1512.01712.

Абибуллаева, А. Новая модель извлечения ключевых слов из ансамбля в казахском языке с помощью машинного обучения // Кандидатская диссертация, - 2023, университет Гази, Турция.

Каур, Дж., & Гупта, В. (2010). Эффективные подходы к извлечению ключевых слов. Международный журнал проблем компьютерных наук (IJCSI), 7 (6), 144.

Лю, З., Хуан, В., Чжэн, Ю., и Сун, М. (2010, октябрь). Автоматическое извлечение ключевых фраз с помощью тематической декомпозиции. В материалах конференции 2010 года по эмпирическим методам обработки естественного языка (стр. 366-376).

Ким, С. Н., Меделян, О., Кан, М. Ю., & Болдуин, Т. (2013). Автоматическое извлечение ключевых фраз из научных статей. Языковые ресурсы и оценка, 47, 723-742.

Аваджан, А. А. (2014, октябрь). Неконтролируемый подход к автоматическому извлечению ключевых слов из арабских документов. В материалах 26-й конференции по компьютерной лингвистике и обработке речи (ROCLING 2014) (стр. 175-184).

Бердеврим, С. А., Боячи, А. И Аль Тани, Д. А. С. (2018). Ийилештирилмиш отоматик анахтар келиме чикарими тормоз. Стамбульский университет Тикарет Текнолоджи ве Уйгуламалы Билимлер Дергиси, 1(1), 11-19.

Базалделла, М., Антолли, Э., Серра, Г. и Тассо, С. (2018). Двунаправленная рекуррентная нейронная сеть lstm для извлечения ключевой фразы. В цифровых библиотеках и мультимедийных архивах: 14-я итальянская исследовательская конференция по цифровым библиотекам, IRCDL 2018, Удине, Италия, 25-26 января 2018 г., Материалы 14 (стр. 180-187). Springer International Publishing.

Килич Юнлю, Х., & Çetin, A. (2023). Извлечение ключевых слов как маркировка последовательности с помощью алгоритмов классификации. Нейронные вычисления и приложения, 35 (4), 3413-3422.

### References

Abibullayeva, A., & Çetin, A. (2022). Keyword Extraction from Kazakh News Dataset with BERT. *El-Cezeri*, 9(4), 1193-1200.

Siddiqi, S., & Sharan, A. (2015). Keyword and keyphrase extraction techniques: a literature review. *International Journal of Computer Applications*, 109(2).

Kılıç, H., & Çetin, A. (2023). A Novel Graph-Based Ensemble Token Classification Model for Keyword Extraction. *Arabian Journal for Science and Engineering*, 48(8), 10673-10680.

Papagiannopoulou, E., & Tsoumakas, G. (2020). A review of keyphrase extraction. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(2), e1339.

Nguyen, T. D., & Kan, M. Y. (2007, December). Keyphrase extraction in scientific publications. In *International conference on Asian digital libraries* (pp. 317-326). Berlin, Heidelberg: Springer Berlin Heidelberg.

Zhang, Y., & Xiao, W. (2018). Keyphrase generation based on deep seq2seq model. *IEEE access*, 6, 46047-46057.

Chen, J., Zhang, X., Wu, Y., Yan, Z., & Li, Z. (2018). Keyphrase generation with correlation constraints. *arXiv preprint arXiv:1808.07185*.

Meem, N. T. A., Chowdhury, M. M. H., & Rahman, M. M. (2019, December). A Machine Learning Approach to Extract Keyphrases from Bengali Document using CNN-Bidirectional LSTM. In *2019 22nd International Conference on Computer and Information Technology (ICCIIT)* (pp. 1-5). IEEE.

Lopyrev, K. (2015). Generating news headlines with recurrent neural networks. *arXiv preprint arXiv:1512.01712*.

Abibullayeva, A. A novel ensemble keyword extraction model in the kazakh language with machine learning // Ph. D. Thesis,- 2023, Gazi university, Turkey.

Kaur, J., & Gupta, V. (2010). Effective approaches for extraction of keywords. *International Journal of Computer Science Issues (IJCSI)*, 7(6), 144.

Liu, Z., Huang, W., Zheng, Y., & Sun, M. (2010, October). Automatic keyphrase extraction via topic decomposition. In *Proceedings of the 2010 conference on empirical methods in natural language processing* (pp. 366-376).

Kim, S. N., Medelyan, O., Kan, M. Y., & Baldwin, T. (2013). Automatic keyphrase extraction from scientific articles. *Language resources and evaluation*, 47, 723-742.

Awajan, A. A. (2014, October). Unsupervised approach for automatic keyword extraction from Arabic documents. In *Proceedings of the 26th Conference on Computational Linguistics and Speech Processing (ROCLING 2014)* (pp. 175-184).

Birdevrim, S. A., Boyacı, A., & Al Thani, D. A. S. (2018). İyileştirilmiş otomatik anahtar kelime çıkarımı brake. *İstanbul Ticaret Üniversitesi Teknoloji ve Uygulamalı Bilimler Dergisi*, 1(1), 11-19.

Basaldella, M., Antolli, E., Serra, G., & Tasso, C. (2018). Bidirectional lstm recurrent neural network for keyphrase extraction. In *Digital Libraries and Multimedia Archives: 14th Italian Research Conference on Digital Libraries, IRCDL 2018, Udine, Italy, January 25-26, 2018, Proceedings 14* (pp. 180-187). Springer International Publishing.

Kılıç Ünlü, H., & Çetin, A. (2023). Keyword extraction as sequence labeling with classification algorithms. *Neural Computing and Applications*, 35(4), 3413-3422.

## CONTENTS

## INFORMATICS

<b>Zh.K. Abdugulova, M. Tlegen, A.T. Kishubaeva, N.M. Kisikova, A.K. Shukirova</b> AUTOMATION OF MINING EQUIPMENT USING DIGITAL CONTROL MACHINES.....	5
<b>A.A. Abibullayeva, A.S. Baimakhanova</b> USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES IN KEYWORD EXTRACTION.....	25
<b>M. Ashimgaliyev, K. Dyussekeyev, T. Turymbetov, A. Zhumadillayeva</b> ADVANCING SKIN CANCER DETECTION USING MULTIMODAL DATA FUSION AND AI TECHNIQUES.....	37
<b>D.S. Amirkhanova, O.Zh. Mamyrbayev</b> EL-GAMAL'S CRYPTOGRAPHIC ALGORITHM: MATHEMATICAL FOUNDATIONS, APPLICATIONS AND ANALYSIS.....	52
<b>A.Sh. Barakova, O.A. Ussatova, Sh.E. Zhussipbekova, Sh.M. Urazgalieva, K.S. Shadinova</b> USE OF BLOCKCHAIN FOR DATA PROTECTION AND TECHNOLOGY DRAWBACKS.....	67
<b>M. Kantureyev<sup>1</sup>, G. Bekmanova, A. Omarbekova, B. Yergesh, V. Franzoni</b> ARTIFICIAL INTELLIGENCE TECHNOLOGIES AND SOLVING SOCIAL PROBLEMS.....	78
<b>A.B. Kassekeyeva, A.B. Togissova*, A.M. Bakiyeva, Z.B. Lamasheva, Y.N. Baibakty</b> ANALYSIS OF COMPARATIVE OPINIONS USING INFORMATION TECHNOLOGY.....	88
<b>M. Mussaif, A. Kintonova, A. Nazyrova, G. Muratova, I.F. Povkhan</b> IMPROVED PUPIL LOCALIZATION METHOD BASED ON HOUGH TRANSFORM USING ELLIPTICAL AND CIRCULAR COMPENSATION.....	103
<b>Zh. S. Mutalova, A.G. Shaushenova, G.O. Issakova, A.A. Nurpeisova, M.B. Ongarbayeva, G.A. Abdygalikova</b> THE METHOD FOR RECOGNIZING A PERSON FROM A FACE IMAGE BASED ON MOVING A POINT ALONG GUIDES.....	118

<b>G. Nurzhaubayeva, K. Chezhimbayeva, H. Norshakila</b> THE DEVELOPMENT AND ANALYSIS OF A WEARABLE TEXTILE YAGI-UDA ANTENNA DESIGN FOR SECURITY AND RESCUE PURPOSES.....	138
<b>A.A. Oxenenko, A.S.Yerimbetova, A. Kuanayev, R.I. Mukhamediev, Ya.I. Kuchin</b> TECHNICAL TOOLS FOR REMOTE MONITORING USING UNMANNED AERIAL PLATFORMS.....	152
<b>B.S. Omarov, A.B. Toktarova, B.S. Kaldarova, A.Z. Tursynbayev, R.B. Abdrakhmanov</b> DETECTING OFFENSIVE LANGUAGE IN LOW-RESOURCE LANGUAGES WITH BILSTM.....	174
<b>G.Taganova, D.A. Tussupov, A. Nazyrova, A.A. Abdildaeva, T.Zh. Yermek</b> SHORT-TERM FORECAST OF POWER GENERATION OF PHOTOVOLTAIC POWER PLANTS BY COMPARING LSTM AND MLP MODELS.....	190
<b>Zh. Tashenova, E. Nurlybaeva, Zh.Abdugulova, Sh. Amanzholova</b> CREATION OF SOFTWARE BASED ON SPECTRAL ANALYSIS FOR STEGOANALYSIS OF DIGITAL AUDIO FILES.....	203
<b>Zh.U. Shermantayeva, O.Zh. Mamyrbayev</b> DEVELOPMENT AND CREATION OF HYBRID EWT-LSTM-RELM- IEWT MODELING IN HIGH-VOLTAGE ELECTRIC NETWORKS.....	223

## МАЗМҰНЫ

### ИНФОРМАТИКА

<b>Ж.К. Абдугулова, М. Тлеген, А.Т. Кишубаева, Н.М. Кисикова, А.К. Шукирова</b> САНДЫҚ БАСҚАРУ СТАНОКТАРЫНЫҢ КӨМЕГІМЕН ТАУ-КЕН-ШАХТА ЖАБДЫҚТАРЫН АВТОМАТТАНДЫРУ.....	5
<b>А.А. Абибуллаева, А.С. Баймаханова</b> КІЛТТІК СӨЗДЕРДІ ШЫҒАРУДА МАШИНАЛЫҚ ЖӘНЕ ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ.....	25
<b>М. Ашимғалиев, К. Дюсекеев, Т. Турымбетов, А. Жумадиллаева</b> МУЛЬТИМОДАЛЬДЫ ДЕРЕКТЕРДІ БІРІКТІРУ ЖӘНЕ ЖАСАНДЫ ИНТЕЛЛЕКТ ӘДІСТЕРІН ҚОЛДАНА ОТЫРЫП, ТЕРІ ҚАТЕРЛІ ІСІГІН АНЫҚТАУДЫ ЖЕТІЛДІРУ.....	37
<b>Д.С. Әмірханова, Ө.Ж. Мамырбаев</b> ЭЛЬ-ГАМАЛЬДЫҢ КРИПТОГРАФИЯЛЫҚ АЛГОРИТМІ: МАТЕМАТИКАЛЫҚ НЕГІЗДЕРІ, ҚОЛДАНУ ЖӘНЕ ТАЛДАУ.....	52
<b>А.Ш. Баракова, О.А.Усатова, Ш.Е.Жусипбекова, Ш.М. Уразғалиева, К.С. Шадинова</b> ДЕРЕКТЕРДІ ҚОРҒАУДА БЛОКЧЕЙНДІ ПАЙДАЛАНУ ЖӘНЕ ТЕХНОЛОГИЯНЫҢ КЕМШІЛІКТЕРІ.....	67
<b>М.А. Кантуреева, Г.Т. Бекманова, А.С. Омарбекова, Б.Ж. Ергеш, V. Franzoni</b> ЖАСАНДЫ ИНТЕЛЛЕКТТІК ТЕХНОЛОГИЯЛАР ЖӘНЕ ӘЛЕУМЕТТІК МӘСЕЛЕЛЕРДІ ШЕШУ.....	78
<b>А.Б. Касекеева, А.Б. Тогисова, А.М. Бакиева, Ж.Б. Ламашева, Е.Н. Байбақты</b> АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАРДЫ ҚОЛДАНУ АРҚЫЛЫ САЛЫСТЫРМАЛЫ ПІКІРЛЕРДІ ТАЛДАУ.....	88
<b>М. Мұсайф, А.Ж. Кинтонова, А.Е. Назырова, Г. Муратова, И.Ф. Повхан</b> ЭЛЛИПТИКАЛЫҚ ЖӘНЕ ДӨҢГЕЛЕК КОМПЕНСАЦИЯНЫ ҚОЛДАНА ОТЫРЫП, ХАФ ТҮРЛЕНДІРУІНЕ НЕГІЗДЕЛГЕН КӨЗДІҢ ҚАРАШЫҒЫҢ ЛОКАЛИЗАЦИЯЛАУДЫҢ ЖЕТІЛДІРІЛГЕН ӘДІСІ.....	103

<b>Ж.С. Муталова, А.Г. Шаушенова, Г.О. Исакова, А. Нұрпейісова, М.Б. Оңғарбаева, Г.А. Әбдіғалықова</b> НҮКТЕНІ БАҒЫТТАУШЫЛАР БОЙЫМЕН ЖЫЛЖЫТУ НЕГІЗІНДЕ АДАМДЫ БЕТ БЕЙНЕСІ АРҚЫЛЫ ТАНУ ӘДІСІ.....	118
<b>Г. Нуржаубаева, К. Чежимбаева, Х. Норшакила</b> ҚҰТҚАРУ ҚЫЗМЕТІ МАҚСАТЫНДА КИІМГЕ ОРНАЛАСТЫРЫЛАТЫН ТЕКСТИЛЬДІ ЯГИ-УДА АНТЕННАСЫНЫҢ ДИЗАЙНЫН ҚҰРУ ЖӘНЕ ТАЛДАУ.....	138
<b>А.А. Оксененко, А.С. Еримбетова, А. Куанаев, Р.И. Мухамедиев, Я.И. Кучин</b> ҰШҚЫШСЫЗ ӘУЕ ПЛАТФОРМАЛАРЫН ПАЙДАЛАНАТЫН ҚАШЫҚТАН МОНИТОРИНГ ЖҮРГІЗУ ҮШІН ТЕХНИКАЛЫҚ ҚҰРАЛДАР.....	152
<b>Б.С. Омаров, А.Б. Тоқтарова, Б.С. Қалдарова, А.З. Турсынбаев, Р.Б. Абдрахманов</b> БЕЙӘДЕП СӨЗДЕРДІ АЗ РЕСУРСТЫ ТІЛДЕРДЕН АНЫҚТАУДА BILSTM- ДІ ҚОЛДАНУ.....	174
<b>Г.Ж. Таганова, Д.А. Тусупов, А. Назырова, А.А. Абдильдаева, Т.Ж. Ермек</b> LSTM ЖӘНЕ MLP МОДЕЛЬДЕРІН САЛЫСТЫРУ АРҚЫЛЫ ФОТОЭЛЕКТРЛІК ЭЛЕКТР СТАНЦИЯЛАРЫНЫҢ ЭЛЕКТР ЭНЕРГИЯСЫН ӨНДІРУДІҢ ҚЫСҚА МЕРЗІМДІ БОЛЖАМЫ.....	190
<b>Ж.М. Ташенова, Э. Нурлыбаева, Ж.К. Абдугулова, Ш.А. Аманжолова</b> САНДЫҚ АУДИОФАЙЛДАРДЫ СТЕГО ТАЛДАУ ҮШІН СПЕКТРАЛДЫ ТАЛДАУ НЕГІЗІНДЕ БАҒДАРЛАМАЛЫҚ ҚҰРАМДЫ ҚҰРУ.....	203
<b>Ж.У. Шермантаева, О.Ж. Мамырбаев</b> ЖОҒАРЫ КЕРНЕУЛІ ЭЛЕКТР ЖЕЛІЛЕРІНДЕ ГИБРИДТІ EWT-LSTM- RELM-IEWT МОДЕЛЬДЕУДІ ДАМЫТУ ЖӘНЕ ҚҰРУ.....	223



## СОДЕРЖАНИЕ

## ИНФОРМАТИКА

- Ж.К. Абдугулова, А.Т. Кишубаева, Н.М. Кисикова, А.К. Шукирова**  
АВТОМАТИЗАЦИЯ ГОРНО-ШАХТНОГО ОБОРУДОВАНИЯ С  
ПОМОЩЬЮ СТАНКОВ ЦИФРОВОГО УПРАВЛЕНИЯ.....5
- А.А. Абибуллаева, А.С. Баймаханова**  
ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО И ГЛУБОКОГО  
ОБУЧЕНИЯ ПРИ ИЗВЛЕЧЕНИИ КЛЮЧЕВЫХ СЛОВ.....25
- М. Ашимгалиев, К. Дюсекеев, Т. Турымбетов, А. Жумадиллаева**  
СОВЕРШЕНСТВОВАНИЕ МЕТОДОВ ВЫЯВЛЕНИЯ РАКА КОЖИ  
С ИСПОЛЬЗОВАНИЕМ МУЛЬТИМОДАЛЬНОГО ОБЪЕДИНЕНИЯ  
ДАННЫХ И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА.....37
- Д. С. Эмірханова, О. Ж. Мамырбаев**  
КРИПТОГРАФИЧЕСКИЙ АЛГОРИТМ ЭЛЬ-ГАМАЛЯ:  
МАТЕМАТИЧЕСКИЕ ОСНОВЫ, ПРИМЕНЕНИЕ И АНАЛИЗ.....52
- А.Ш. Баракова, О.А. Усатова, Ш.Е. Жусипбекова, Ш.М. Уразгалиева,  
К.С. Шадинова**  
ИСПОЛЬЗОВАНИЕ БЛОКЧЕЙНА ДЛЯ ЗАЩИТЫ ДАННЫХ И  
НЕДОСТАТКИ ТЕХНОЛОГИИ.....67
- М.А. Кантуреева, Г.Т. Бекманова, А.С. Омарбекова, Б.Ж. Ергеш,  
V. Franzon**  
ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И  
РЕШЕНИЕ СОЦИАЛЬНЫХ ПРОБЛЕМ.....78
- А.Б. Касекеева, А.Б. Тогисова, А.М. Бакиева, Ж.Б. Ламашева,  
Е.Н. Байбакты**  
АНАЛИЗ СРАВНИТЕЛЬНЫХ МНЕНИЙ С ИСПОЛЬЗОВАНИЕМ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.....88
- М. Мусайф, А.Ж. Кинтонова, А.Е. Назырова, Г. Муратова, И.Ф. Повхан**  
УЛУЧШЕННЫЙ МЕТОД ЛОКАЛИЗАЦИИ ЗРАЧКА НА ОСНОВЕ  
ПРЕОБРАЗОВАНИЯ ХАФА С ИСПОЛЬЗОВАНИЕМ  
ЭЛЛИПТИЧЕСКОЙ И КРУГОВОЙ КОМПЕНСАЦИИ.....103

<b>Ж.С. Муталова, А.Г. Шаушенова, Г.О. Исакова, А.А. Нурпейсова, М.Б. Онгарбаева, Г.А. Абдыгаликова</b> МЕТОД РАСПОЗНАВАНИЯ ЧЕЛОВЕКА ПО ИЗОБРАЖЕНИЮ ЛИЦА НА ОСНОВЕ ПЕРЕМЕЩЕНИЯ ТОЧКИ ПО НАПРАВЛЯЮЩИМ.....	118
<b>Г. Нуржаубаева, К. Чежимбаева, Х. Норшакила</b> РАЗРАБОТКА И АНАЛИЗ ДИЗАЙНА ВСТРАИВАЕМОЙ ТЕКСТИЛЬНОЙ ЯГИ-УДА АНТЕННЫ ДЛЯ ПРИМЕНЕНИЯ В СФЕРЕ СПАСАТЕЛЬНЫХ СЛУЖБ.....	138
<b>А.А. Оксененко, А.С. Еримбетова, А. Куанаев, Р.И. Мухамедиев, Я.И. Кучин</b> ТЕХНИЧЕСКИЕ СРЕДСТВА ДИСТАНЦИОННОГО МОНИТОРИНГА С ПОМОЩЬЮ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ ПЛАТФОРМ.....	152
<b>Б.С. Омаров, А.Б. Токтарова, Б.С. Калдарова, А.З. Турсынбаев, Р.Б. Абдрахманов</b> ИСПОЛЬЗОВАНИЕ VILSTM ДЛЯ ОПРЕДЕЛЕНИЯ ОСКОРБИТЕЛЬНОГО ЯЗЫКА В ЯЗЫКАХ С НИЗКИМ УРОВНЕМ РЕСУРСОВ.....	174
<b>Г.Ж. Таганова, Д.А. Тусупов, А. Назырова, А.А. Абдильдаева, Т.Ж. Ермек</b> КРАТКОСРОЧНЫЙ ПРОГНОЗ ВЫРАБОТКИ ЭЛЕКТРОЭНЕРГИИ ФОТОЭЛЕКТРИЧЕСКИМИ ЭЛЕКТРОСТАНЦИЯМИ ПУТЕМ СРАВНЕНИЯ МОДЕЛЕЙ LSTM И MLP.....	190
<b>Ж.М. Ташенова, Э. Нурлыбаева, Ж.К. Абдугулова, Ш.А. Аманжолова</b> СОЗДАНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ НА БАЗЕ СПЕКТРАЛЬНОГО АНАЛИЗА ДЛЯ СТЕГОАНАЛИЗА ЦИФРОВЫХ АУДИОФАЙЛОВ.....	203
<b>Ж.У. Шермантаева, О.Ж. Мамырбаев</b> РАЗРАБОТКА И СОЗДАНИЕ ГИБРИДНОГО МОДЕЛИРОВАНИЯ EWT-LSTM-RELM-IEWT В ВЫСОКОВОЛЬТНЫХ ЭЛЕКТРИЧЕСКИХ СЕТЯХ.....	223

**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)**

Директор отдела издания научных журналов НАН РК *А. Ботанқызы*

Редакторы: *Д.С. Аленов, Ж.Ш. Әден*

Верстка на компьютере *Г.Д. Жадыранова*

Подписано в печать 30.09.2024.

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

15,5 п.л. Тираж 300. Заказ 3.