

Регистрационный №  
KZ89VPY00111238

ISSN: 2413-8614  
ISSN: 3079-2363 (Online)

Қазақстан Республикасы Ғылым және жоғары білім  
министрлігі  
**Азаматтық авиация академиясының Жаршысы**

---

**Вестник Академии гражданской авиации**  
Министерства науки и высшего образования  
Республики Казахстан

---

**Bulletin of Civil aviation Academy**  
Ministry of Science and Higher Education of the Republic  
of Kazakhstan

№1(36) 2025

АЛМАТЫ – 2025

**«Азаматтық Авиация Академиясының Жаршысы»  
Ғылыми басылым**

*2015 жылдан бастап шығады*

*Журналдың шығу мерзімділігі - жылына 4 рет  
Басылымның тілдері: қазақ, орыс, ағылшын*

*Журналда авиация саласындағы техникалық, жаратылыстану, гуманитарлық және әлеуметтік ғылымдардың әртүрлі салаларында ғалымдардың, оқытушылардың, PhD докторанттар мен магистранттардың зерттеулерінің нәтижелері бойынша ғылыми мақалалар жарияланады.*

"Азаматтық авиация академиясы" АҚ Закарпатская көшесі, 44, Каб. №202  
А35М2Н5 (жаңа индекс), Алматы қ., Қазақстан Республикасы,  
Тел.: 8 747 182 52 41, e-mail: [vestnik@agakaz.kz](mailto:vestnik@agakaz.kz)

"ААА Жаршысы" ғылыми журналының электрондық нұсқасы <https://vestnik.agakaz.kz/>  
сайтында "Мұрағат" бөлімінде орналастырылған

**Бас редактор**Сейдахметов Б.К., э.ғ.к., асс.профессор **(h-индекс:2)****Бас редактордың орынбасары**Көшекөв Қ.Т., т.ғ.д., профессор **(h-индекс:6)****Редакциялық алқа:**

1. А.В.Стрельцов, доктор (PhD), Эмбри-Риддл Аэроавиация университетінің жаратылыстану ғылымдары кафедрасының профессоры (Дейтона жағажайы, Флорида) **(h-индекс:24);**
2. И.В. Яцкив, инженерия ғылымдарының докторы, профессор Көлік және байланыс институтының (TSI) Басқарма Төрағасы, Еуропалық көлік зерттеу қауымдастығының Басқарма мүшесі **(h-индекс:8);**
3. И.А. Искендеров, ф.-м.ғ. к., Әзірбайжанның Ұлттық авиация академиясының "Аэроғарыштық құралдар" кафедрасының меңгерушісі, профессор **(h-индекс:1);**
4. К.Б. Алдамжаров, т.ғ.д, профессор «Азаматтық авиация академиясы» АҚ **(h-индекс:1);**
5. Е.А. Оспанов, 6D070200 – «Шәкәрім атындағы Университет» КЕАҚ-ның «Автоматтандыру және басқару» мамандығы бойынша PhD, ЖАК қауымдастырылған профессоры **(h-индекс:9);**
6. С.А. Бельгинова, 6D070300 – Ақпараттық жүйелер (салалар бойынша) мамандығы бойынша PhD, Тұран университетінің, Ақпараттық технологиялар кафедрасының қауымдастырылған профессоры **(h-индекс:6);**
7. Р.К. Анаятова, PhD докторы, «Азаматтық авиация академиясы» АҚ, «Авиациялық ағылшын тілі» кафедрасының меңгерушісі **(h-индекс:3);**
8. Е.Е. Қарсыбаев, т.ғ.д, профессор, «Азаматтық авиация академиясы» АҚ **(h-индекс:1);**
9. М.Н. Қалимолдаев, ф.-м. ғ. д., профессор, ҚР БҒМ Ғылым комитеті информатика және басқару мәселелері институты **(h-индекс:10);**
10. Тулешов А. Қ., т. ғ. д., ХАА академигі, Механика және машинатану институты **(h-индекс:4).**

**Жауапты редактор:** А.Макеева

**«Вестник Академии гражданской авиации»**

Научное издание

*Издается с 2015 года*

*Периодичность издания журнала – 4 номера в год.*

*Языки издания: казахский, русский, английский*

*В журнале публикуются научные статьи по результатам исследований ученых, преподавателей, докторантов PhD и магистрантов в различных областях технических, естественных, гуманитарных и общественных наук авиационной отрасли.*

АО «Академия гражданской авиации» ул. Закарпатская, 44, Каб. №202

А35М2Н5 (новый индекс), г. Алматы, Республика Казахстан

Тел.: 8 747 182 52 41, e-mail: [vestnik@agakaz.kz](mailto:vestnik@agakaz.kz)

Электронная версия научного журнала "Вестник АГА" размещено на сайте  
<https://vestnik.agakaz.kz> / в разделе "Архив"

**Главный редактор**

Б.К.Сейдахметов, к.э.н., асс. профессор (**h-индекс:2**)

**Зам. главного редактора**

К.Т.Кошеков, д.т.н., профессор (**h-индекс:6**)

**Редакционная коллегия:**

1. А.В. Стрельцов, доктор (PhD), профессор кафедры естественных наук университета аэронавтики Эмбри-Риддл (г. Дейтона-Бич, штат Флорида) (**h-индекс:24**);
2. И.В. Яцкив, Председатель правления Института транспорта и связи (TSI), доктор инженерных наук, профессор, член правления Европейской ассоциации транспортных исследовательских институтов (**h-индекс:8**);
3. И.А.Искендеров, к.ф.-м.н., профессор, зав. каф. «Аэрокосмические приборы» Национальной авиационной академии Азербайджана (**h-индекс:1**);
4. К.Б. Алдамжаров, д.т. н., профессор АО «Академия гражданской авиации» (**h-индекс:1**);
5. Е.А. Оспанов, PhD по специальности 6D070200 – «Автоматизация и управление», ассоциированный профессор ВАК, НАО «Университет имени Шакарима» (**h-индекс:9**);
6. С.А. Бельгинова, PhD по специальности 6D070300 – Информационные системы (по отраслям), ассоциированный профессор университета Туран, кафедры информационных технологии (**h-индекс:6**);
7. Е.Е. Карсыбаев, д.т.н., профессор АО «Академия гражданской авиации» (**h-индекс:1**);
8. Р.К. Анаятова, доктор PhD, зав. каф. «Авиационный английский язык», АО «Академия гражданской авиации» (**h-индекс:3**);
9. М.Н. Калимолдаев, д.ф.-м.н., профессор, Институт проблем информатики и управления комитета науки МОН РК (**h-индекс:10**);
10. А.К. Тулешов, д.т.н., академик МИА, Институт механики и машиноведения (**h-индекс:4**).

**Ответственный редактор:** А. Макеева

**“Bulletin of the Civil Aviation Academy”**  
*Scientific publication*

*Founded in 2015*

*Periodicity is 4 issues per year.*

*Publication Languages are Kazakh, Russian and English*

*The journal publishes scientific articles based on the results of research by scientists, teachers, PhD students and undergraduates in various fields of technical, natural, humanitarian and social sciences of the aviation industry.*

JSC “Academy of Civil Aviation” Zakarpatskaya str., 44, Office No. 202  
A35M2N5 (new index), Almaty, Republic of Kazakhstan  
Tel.: 8 747 182 52 41, e-mail: [vestnik@agakaz.kz](mailto:vestnik@agakaz.kz)

The electronic version of the scientific journal "Bulletin of the AGA" is posted on the website <https://vestnik.agakaz.kz/> in the "Archive" section

**Editor-in – chief**

Seydakhmetov B.K., Candidate of Economics, Associate Professor (**h-index:2**)

**Deputy Chief Editor**

Koshekov K.T., doctor of technical sciences, professor (**h-index:6**)

**Editorial staff:**

1. A.V. Streltsov, Doctor of Philosophy (PhD), Professor of Engineering Physics department of Physical Sciences at Embry-Riddle Aeronautics University (Daytona Beach, Florida) (**h-index:24**);
2. I.V. Yatskiv, Chairman of the Board of the Institute of Transport and Communications (TSD), Doctor Engineering Sciences, Professor, Member of the Board of the European Association of Transport Research Institutes (**h-index:8**);
3. I.A. Isgandarov, candidate of physical and mathematical sciences, professor, Head of the Department "Aerospace Devices" of the National Aviation Academy of Azerbaijan (**h-index:1**);
4. K.B. Aldamzharov, doctor of technical sciences, professor (**h-index:1**);
5. Ospanov E. A., PhD in the specialty 6D070200 – "Automation and Control", Associate Professor of the Higher Attestation Commission, NJSC "Shakarim University" (**h-index:9**);
6. S. A. Belginova, PhD in specialty 6D070300 – Information Systems (by industry), Associate Professor of Turan University, Department of Information Technology (**h-index:6**);
7. E.E. Karsybaev, Doctor of Technical Sciences, Professor, JSC "Academy of Civil Aviation" (**h-index:1**);
8. R.K. Anayatova, PhD, Head of the Department "Aviation English", JSC "Academy of Civil Aviation" (**h-index:3**);
9. Kalimoldaev M.N., PhD, Professor, Institute of Problems of Informatics and Management of the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan (**h-index:10**);
10. A.K. Tuleshov, PhD, Academician MIA, Institute of Mechanics and Machine Science (**h-index:4**).

**Responsible editor:** A. Makeeva

**МАЗМҰНЫ/СОДЕРЖАНИЕ/ CONTENTS**

<b>ӘУЕ КӨЛІГІ ЖӘНЕ ТЕХНОЛОГИЯЛАР ВОЗДУШНЫЙ ТРАНСПОРТ И ТЕХНОЛОГИИ AIR TRANSPORT AND TECHNOLOGY</b>	
T.A. Sagdiyev, I.S. Maturazov, N.A. Isakov ANALYSIS OF THE AIRCRAFT ENGINE FAILURES	10
A. Nagimov, G. Beketova ANALYSING AND PREDICTING WEATHER CONDITIONS FOR PLANNING FLIGHTS OF UNMANNED AERIAL VEHICLES USING BIG DATA	19
A. Akhmet, A. Gusmanov, A. Akramkhanov, Sh. Akimbay CREATION OF A SIMULATION PLATFORM FOR TESTING UAV CONTROL ALGORITHMS	29
Kh. I.Abdulayev, S.Q. Allakhverdiyeva DEFECTS OF AIRCRAFT STRUCTURAL ELEMENTS BASED ON POLYMERIC-COMPOSITE AND THEIR DIAGNOSTICS BY NONDESTRODUCTIVE CONTROL METHODS	40
<b>ЛОГИСТИКА, ТАСЫМАЛДАУДЫ ҰЙЫМДАСТЫРУ, КӨЛІКТЕГІ ҚАУІПСІЗДІК ЛОГИСТИКА, ОРГАНИЗАЦИЯ ПЕРЕВОЗОК, БЕЗОПАСНОСТЬ НА ТРАНСПОРТЕ LOGISTICS, TRANSPORTATION ORGANIZATION, TRANSPORT SECURITY</b>	
A.Ж. Абжапбарова, О.В. Гармаш ПРОБЛЕМЫ РЕГУЛЯРНОСТИ ВОЗДУШНЫХ ПЕРЕВОЗОК КАЗАХСТАНА	51
В.О. Макогонова, И.Ж. Асильбекова РАЗРАБОТКА МЕТОДИКИ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ГРУЗОВЫХ ПЕРЕВОЗОК	63
<b>КОМПЬЮТЕРЛІК ҒЫЛЫМДАР, АСПАП ЖАСАУ ЖӘНЕ АВТОМАТТАНДЫРУ КОМПЬЮТЕРНЫЕ НАУКИ, ПРИБОРОСТРОЕНИЯ И АВТОМАТИЗАЦИЯ COMPUTER SCIENCE, INSTRUMENTATION AND AUTOMATION</b>	
Н. Алимбекова, Х.М. Рай, Т. Турымбетов, А. Жумадиллаева ПОЛИМЕРЛІ КОНТЕЙНЕРЛЕРДІ СҰРЫПТАУ ПРОЦЕСІН ОҢТАЙЛАНДЫРУ ҮШІН НЕЙРОНДЫҚ ЖЕЛЛІЛЕРДІ ҚОЛДАНУ	75
А.С. Құмарқанова, З.Т. Хасенова, Ю.А. Вайс АНАЛИЗ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ СПРОСА НА БАЗЕ ИСТОРИЧЕСКИХ ДАННЫХ РАСХОДА ЛЕКАРСТВЕННЫХ СРЕДСТВ	90
Койшиева Д. Е., Сыдыбаева М. А., Бельгинова С. А., Жаксыбаев А.М., Ерсайнова Ж.Е ИССЛЕДОВАНИЕ ВЛИЯНИЯ МЕТОДОВ РАСШИРЕНИЯ ДАННЫХ И ШУМОВ НА ТОЧНОСТЬ СЕГМЕНТАЦИИ ЯДЕР В ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОЙ НЕЙРОННОЙ СЕТИ	102
Т.М. Мехдиев, А.К. Шайханова, Г.Б. Бекешова, И.Е. Икласова, К.С. Бакенова МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ТЕКСТОВ (TEXT MINING)	113
S.A. Nurgaliyeva, N.B. Naiman, S.S. Adikanova A REVIEW OF MACHINE LEARNING AND OLFACTORY TECHNOLOGIES FOR RAPID VEGETABLE DISEASE DETECTION	126
T.S. Tursunov, D. Kaibassova COMPARATIVE ANALYSIS OF COLLABORATIVE, CONTENT-BASED AND HYBRID DEEP LEARNING RECOMMENDATION MODELS	138
И.Б.Карымсакова, Д. Б.Бекенова, Д.О.Кожаметова, М.А.Карменова, Т.А.Устинова ОНКОЛОГИЯЛЫҚ АУРУЛАРДЫ ЕМДЕУДЕ ДИАГНОЗ ҚОЮ ҮШІН ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕНІҢ ТҰЖЫРЫМДАМАЛЫҚ МОДЕЛІН ҚҰРУ	147
М.Е.Баталова, Р.Комата, К. Алибеккызы, А.Ж. Бугубаева, М.С.Кунапьянова АНАЛИЗ ИССЛЕДОВАНИЯ ТЕХНОЛОГИИ БЕСПРОВОДНОГО ДАТЧИКА ДЛЯ МОНИТОРИНГА ТЕМПЕРАТУРЫ В ЗОНАХ УСТАНОВКИ ОРТОПЕДИЧЕСКИХ ИМПЛАНТАТОВ	159
А.Б. Саметова, Д.Р. Рашидинов, Л.К. Найзабаева, М.У. Сулейменова РАЗРАБОТКА СИСТЕМЫ ДОКУМЕНТООБОРОТА В ОБРАЗОВАНИИ	170
A. Koshekov, D. Toleubekov, I. Nautan, Y. Kurbanov, R. Togambayev INNOVATIVE TRAINING COMPLEX WITH ENGINEERING SUPPORT: A NEW LEVEL OF TECHNICAL EXPLOITATION OF AVIATION	182



---

**CONTENTS**


---

<b><i>AIR TRANSPORT AND TECHNOLOGY</i></b>	
T.A. Sagdiyev, I.S. Maturazov, N.A. Isakov ANALYSIS OF THE AIRCRAFT ENGINE FAILURES	10
A. Nagimov I, G. Beketova ANALYSING AND PREDICTING WEATHER CONDITIONS FOR PLANNING FLIGHTS OF UNMANNED AERIAL VEHICLES USING BIG DATA	19
A. Akhmet, A. Gusmanov, A. Akramkhanov, Sh. Akimbay CREATION OF A SIMULATION PLATFORM FOR TESTING UAV CONTROL ALGORITHMS	29
Kh. I.Abdulayev, S.Q. Allakhverdiyeva DEFECTS OF AIRCRAFT STRUCTURAL ELEMENTS BASED ON POLYMERIC-COMPOSITE AND THEIR DIAGNOSTICS BY NONDESTRUCTIVE CONTROL METHODS	40
<b><i>LOGISTICS, TRANSPORTATION ORGANIZATION, TRANSPORT SECURITY</i></b>	
A. Abzhapbarova, O. Garmash PROBLEMS OF REGULARITY OF AIR TRANSPORTATION IN KAZAKHSTAN	51
V.O. Makogonova, I.Zh. Asilbekova DEVELOPMENT OF MARKETING RESEARCH METHODOLOGY TO IMPROVE THE EFFICIENCY OF FREIGHT TRANSPORTATION	63
<b><i>COMPUTER SCIENCE, INSTRUMENTATION AND AUTOMATION</i></b>	
N. Alimbekova, H. M. Rai, T. Turymbetov, A. Zhumadillayeva NEURAL NETWORK APPLICATION FOR OPTIMIZING THE SORTING PROCESS OF POLYMER CONTAINERS	75
A. Kumarkanova, Z. Khasenova, Y. Weiss ANALYSING DEMAND FORECASTING MODELS BASED ON HISTORICAL DRUG CONSUMPTION DATA	90
D. Koysheva, M. Sydybaeva, S. Belginova, A. Zhaksybaev, Zh. Ersainova INVESTIGATION OF THE IMPACT OF DATA AUGMENTATION METHODS AND NOISE ON THE ACCURACY OF NUCLEI SEGMENTATION IN HISTOLOGICAL IMAGES USING A DEEP NEURAL NETWORK	102
T. Mekhdiev, A. Shaikhanova, G. Bekeshova, K. Iklasova, K. Bakenova METHODS OF INTELLIGENT TEXT PROCESSING (TEXT MINING)	113
S.A. Nurgaliyeva, N.B. Naiman, S.S. Adikanova A REVIEW OF MACHINE LEARNING AND OLFACTORY TECHNOLOGIES FOR RAPID VEGETABLE DISEASE DETECTION	126
T.S. Tursunov, D. Kaibassova COMPARATIVE ANALYSIS OF COLLABORATIVE, CONTENT-BASED AND HYBRID DEEP LEARNING RECOMMENDATION MODELS	138
I. Karymsakova, D. Bekenova, D. Kozhakhmetova, M. Karmenova CREATION OF A CONCEPTUAL MODEL OF AN INTELLIGENT SYSTEM FOR DIAGNOSIS IN THE TREATMENT OF ONCOLOGICAL DISEASES	147
M. Batalova, P. Komada, K. Alibekkyzy, A. Bugubayeva, M. Kunapianova ANALYSIS OF THE RESEARCH OF WIRELESS SENSOR TECHNOLOGY FOR MONITORING TEMPERATURE IN ORTHOPEDIC IMPLANTS INSTALLATION AREAS	159
A.B. Sametova, D.R. Rashidinov, L.K. Naizabayeva, M.U. Suleimenova DEVELOPMENT OF A DOCUMENT MANAGEMENT SYSTEM IN EDUCATION	170
A. Koshekov, D. Toleubekov, I. Nautan, Y. Kurbanov, R. Togambayev INNOVATIVE TRAINING COMPLEX WITH ENGINEERING SUPPORT: A NEW LEVEL OF TECHNICAL EXPLOITATION OF AVIATION	182

=====

ӨҮЕ КӨЛІГІ ЖӘНЕ ТЕХНОЛОГИЯЛАР  
ВОЗДУШНЫЙ ТРАНСПОРТ И ТЕХНОЛОГИИ  
AIR TRANSPORT AND TECHNOLOGY

=====

IRSTI 52.42.47

UDC 621.89.097.3

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_1](https://doi.org/10.53364/24138614_2025_36_1_1)

T.A. Sagdiyev<sup>1</sup>, I.S. Maturazov<sup>1\*</sup>, N.A. Isakov<sup>1</sup>

<sup>1</sup>Tashkent state transport university, Tashkent, Uzbekistan

E-mail: [maturazov\\_i@tstu.uz](mailto:maturazov_i@tstu.uz)\*

### ANALYSIS OF THE AIRCRAFT ENGINE FAILURES

**Abstract.** *The article investigates the causes, consequences, and prevention strategies related to aircraft engine failures. It provides a comprehensive classification of common engine malfunctions, examining mechanical wear, thermal stress, and operational factors that contribute to failures. A detailed analysis of statistical data on engine failure rates highlights critical trends and risk factors affecting engine performance and reliability. Furthermore, the study explores various diagnostic techniques designed to detect potential failures at early stages, reducing the likelihood of unexpected breakdowns. Modern aviation heavily relies on advanced maintenance strategies and cutting-edge technological solutions to enhance engine durability and efficiency. The article discusses preventive maintenance approaches, including predictive analytics, condition-based monitoring, and real-time diagnostics, which play a crucial role in minimizing failures. Additionally, the role of artificial intelligence and machine learning in fault detection and predictive maintenance is examined as a promising direction for improving aircraft engine reliability. The findings indicate that most engine malfunctions stem from mechanical degradation, excessive thermal loads, and human errors in operation and maintenance. Implementing regular inspections, utilizing advanced diagnostic tools, and integrating modern engineering solutions can significantly improve engine safety and longevity. The study underscores the necessity of continuous monitoring, timely preventive actions, and the adoption of innovative maintenance practices to enhance aviation safety and operational efficiency.*

**Keywords:** *aviation engines, fault analysis, aircraft maintenance, turbine inspection, non-destructive testing, thermal stress, mechanical wear, predictive maintenance, engine diagnostics.*

#### Introduction.

Aircraft engine failures pose significant risks to flight safety, potentially leading to catastrophic incidents. Despite advancements in turbine and piston engine technologies, failures still occur due to a variety of factors, including mechanical wear, carbon buildup, and thermal stress. As Murphy's Law suggests, "If anything can go wrong, it will." and "New system will bring new problems."

That's why for engines malfunctions as a result to the surface coming aviation of events permanent analysis and their "health" tracking "flights safety for very important. Of these participants processes various to roles has. Aviation administration employees of birds' unpleasant events to observe need eye see and from analysis after relevant the order release. Aviation organizations, separately Engines performance from monitoring come came out without, aviation administration instructions to do need their certain activity circle [1].

2008-2015 to ECCAIRS data for based article That one at the time electricity stations (SCF-PP) cause was aviation events results statistic calculations presented will reach.

In recent years, there has been considerable attention to the analysis of aircraft engine failure statistics. Analysis of data from the International Air Transport Association shows that between 2015 and 2023, turbojets account for 42% of all failures, turbofans for 35%, and turboprops for 23%. The main causes of failure are mechanical wear of components (38%), carbon deposition and fouling (27%), thermal stress (18%), manufacturing defects (9%), and maintenance errors (8%). However, historical trends show a 15% decrease in the overall failure rate over the past ten years, partly due to improvements in manufacturing and maintenance technologies, although thermal stress-related failure has increased by 5%, requiring further attention.

### Materials and research methods.

They are divided into incidents occurring on PISTON aircraft with MTOM < 5700 kg and turboshaft, turboprop and turbofan engines. The analysis method can be briefly described as follows: During 2008

The number of aircraft involved in air traffic changed in 2015. To objectify the analysis, coefficients were introduced from the data, corresponding to the number of all SCF-PP incidents or the number of aircraft registered in this category in any chapter of the ATA [4] (per 1000 aircraft).

$$ZS_{GA}(X) = \frac{1000 * LZ_{GA}}{LSP_{GA}} \text{ or } ZS_K(X) = \frac{1000 * LZ_K}{LSP_K},$$

$LZ_{GA}$ ,  $LZ_K$  - suitable MTOM<5700 and MTOM>5700 kg aircraft respectively for events number  $LSP_{GA}$ ,  $LSP_K$  - the number of registered aircraft with MTOM<5700 kg and MTOM>5700 kg, respectively.

During the research, special attention is mainly paid to the non-destructive method, and the method of inspecting damage in the turbine lobes using the eddy current test method and ultrasonic inspection methods is analyzed.

Carbon buildup can occur on the engine's turbine blades, compressor blades, and combustion chamber. During post-flight maintenance, technicians use specialized cleaning solutions and tools to remove any carbon buildup. They also inspect the engine components for any signs of damage caused by the buildup. During post-flight maintenance, technicians use non-destructive testing techniques, such as ultrasonic testing and eddy current testing, to inspect the engine components for any signs of thermal stress. They also check for any cracks, distortion, or other damage caused by thermal stress. Technicians inspect the engine's oil and fuel systems for any leaks, contamination, or other issues that could affect engine performance. They also check the oil and fuel filters and replace them if necessary. Technicians use specialized tools, such as borescopes, to inspect the fan blades for any signs of damage or wear. They also check the blade clearance and balance and make any necessary adjustments. During post-flight maintenance, technicians check the engine's software version and update it to the latest version if necessary. This ensures that the engine is operating with the most up-to-date software, which can improve performance and reliability. Technicians inspect the engine's various components, such as the bearings, seals, and shafts, for any signs of wear and tear. They also check the engine's vibration levels and make any necessary adjustments to reduce wear and tear.

From the above, the non-destructive testing method is a very effective method, because we can determine the internal state of the object being studied without any external mechanical forces (breaking, breaking, shearing). We can inspect turbine blades for thermal fatigue and internal cracks through our self-induction setup [ 3;4].

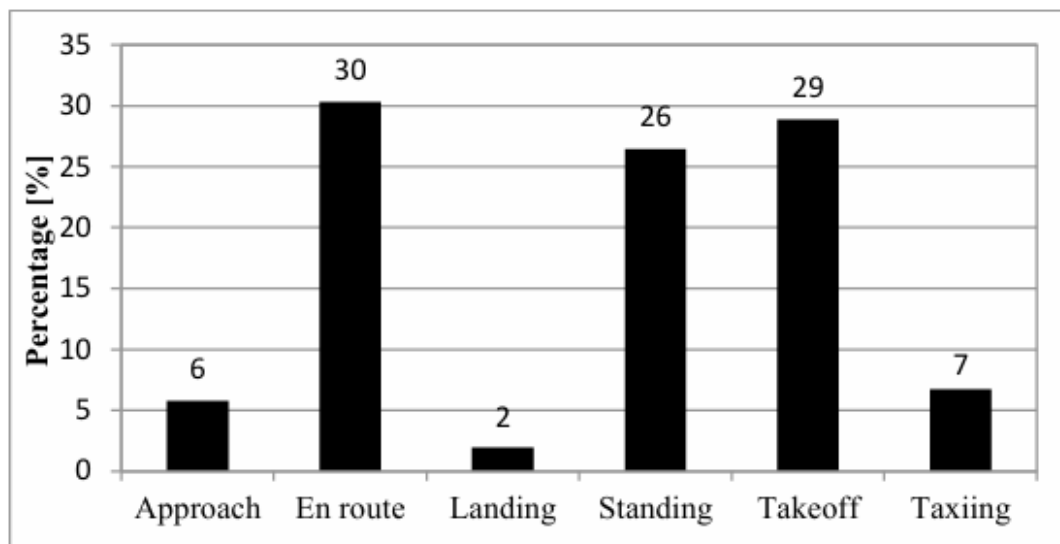


Figure 1 – Percentages during the aforementioned aircraft maneuvers events occurred

Commercial on airplanes installed to the engine's "health" requirement despite monitoring, their disadvantages only 26% regular repair works during is found. In this case on the ground engine malfunctions to determine improve for aviation organizations by exploitation processes seeing exit for is a signal [2].

Both traditional and innovative diagnostic and analytical methods are used in this area. Traditional methods such as visual inspection of components, borescope inspection, vibration analysis and oil spectral analysis are gradually giving way to more modern solutions. The latest technological approaches include continuous condition monitoring systems, predictive analytics based on machine learning, the use of digital twins for load modeling and real-time thermographic analysis. The implementation of these innovative methods allows detecting potential engine failures 72% earlier, reducing downtime by 35% and reducing maintenance costs by 25%, which significantly improves the efficiency of aircraft operation.

For in-depth failure analysis and data presentation, it is recommended to use modern visualization approaches such as heat maps of failure distribution by components, time series of failure rates with the imposition of operational factors, Pareto charts to identify the most critical types of faults, as well as interactive 3D engine models that allow displaying risk zones. Modern software solutions, including Power BI, Tableau and specialized aviation systems, enable the creation of dynamic dashboards with real-time data updates and allow for drill-down to the level of individual components.

### Results and their discussion.

Damaged of the knife photo in the figure 2. The blade one-part brokenness determined center region leader on the edge. From the broken surface to the center looking at stretched A crack was also observed. of the knife air film area (Figures 2 and 3).

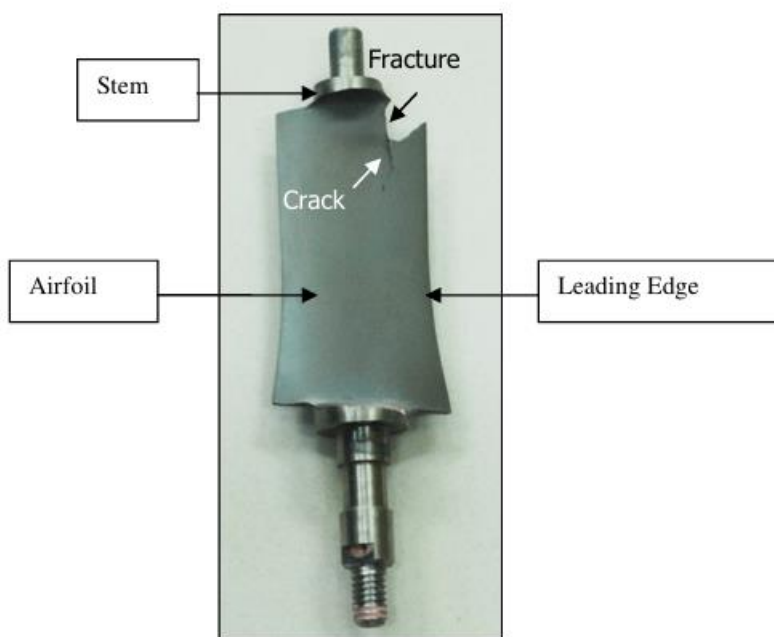


Figure 2 – Damaged stage II compressor stator blade photo

Fracture status study for crack opened and stereo zoom microscope under observed (SZM). Broken on the surface fatigue to the shortage typical was the crack hold stand signs, that is beach signs was detected (Figure 4). Observation beach of signs direction back looking tired cracks come output air of the film in the union location determined and of the knife rod part previous on the edge (Figure 4).

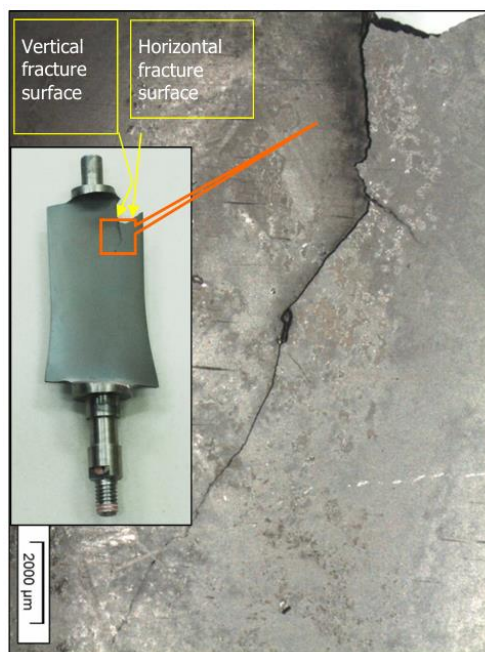


Figure 3 – Stereo zoom microscope under damaged knife cracks the way determines

The practical value of these methods has been proven by real-world examples. For example, after implementing a predictive analytics system to monitor CFM56 engines in 2018, Delta Airlines was able to prevent 38 potential in-flight failures, which allowed it to avoid unplanned

landings and save about \$15 million in repairs and compensation. Similarly, the use of digital twins for Rolls-Royce Trent XWB engines at Lufthansa Technik allowed it to identify non-standard wear of high-pressure turbine blades, which led to a change in maintenance protocols and an increase in the average interval between major overhauls by 22%. Singapore Airlines, having implemented a comprehensive monitoring system using IoT sensors and cloud analytics, was able to reduce the number of unplanned engine replacements by 40% between 2019 and 2022.

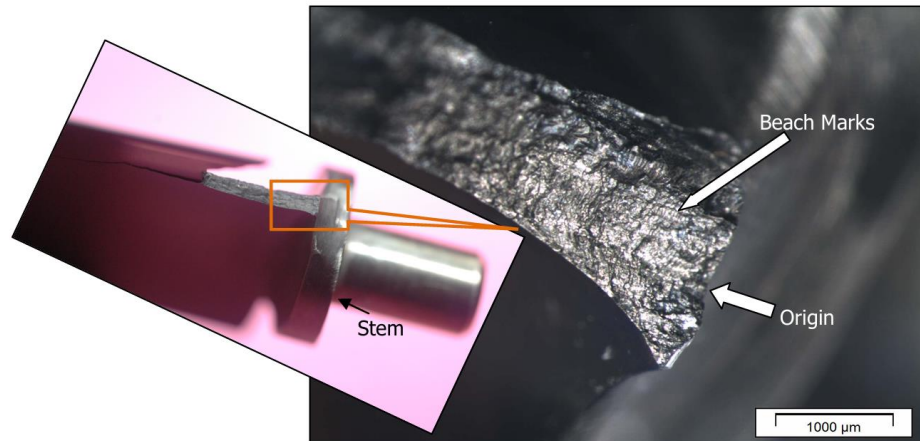


Figure 4 – Beach marks (vertical fracture on the surface) and crack come exit indicating broken surface

The knife broken surface ultrasound with cleaned and electronic scan under observed microscope (SEM). Broken surface strong oxidation (Figure 5). Unclearly demarcated lines crack come output near observed (Figure 5).

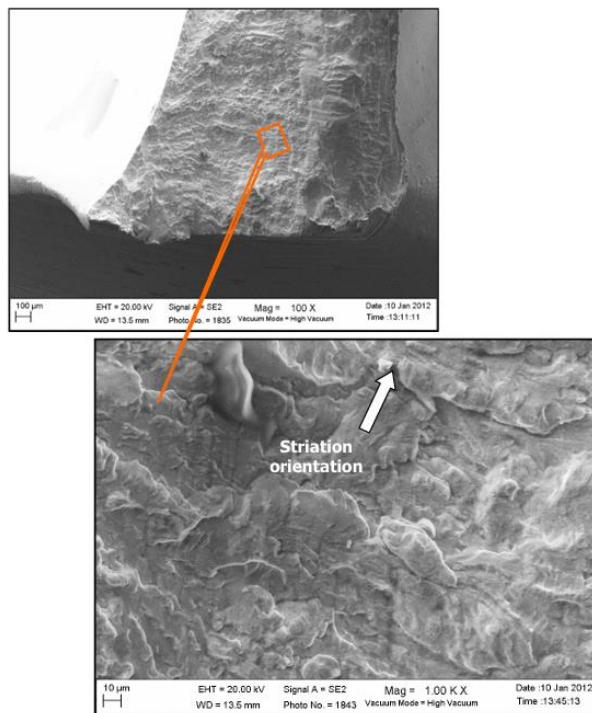


Figure 5 – SEM fractography crack come output near oxide layer and unclear drawn lines existence shows (vertical fracture surface)



Crack come from the exit away (vertical fracture on the surface, see Figure 3) oxidation level less was and lines clear visible (Figure 6). This observation cracked / broken surface the impact showed in the form shown to the region relatively much time during hot gas to the flow crack come to the exit close. 5. lines the direction has also changed determined. Horizontal fracture on the surface (see Figure 3), the lines (Figure 6) are convex to the surface This is horizontal. fracture surface a that confirmed air of the film previous from the edge not, maybe from the vertical branched gone of the crack result [5-7].

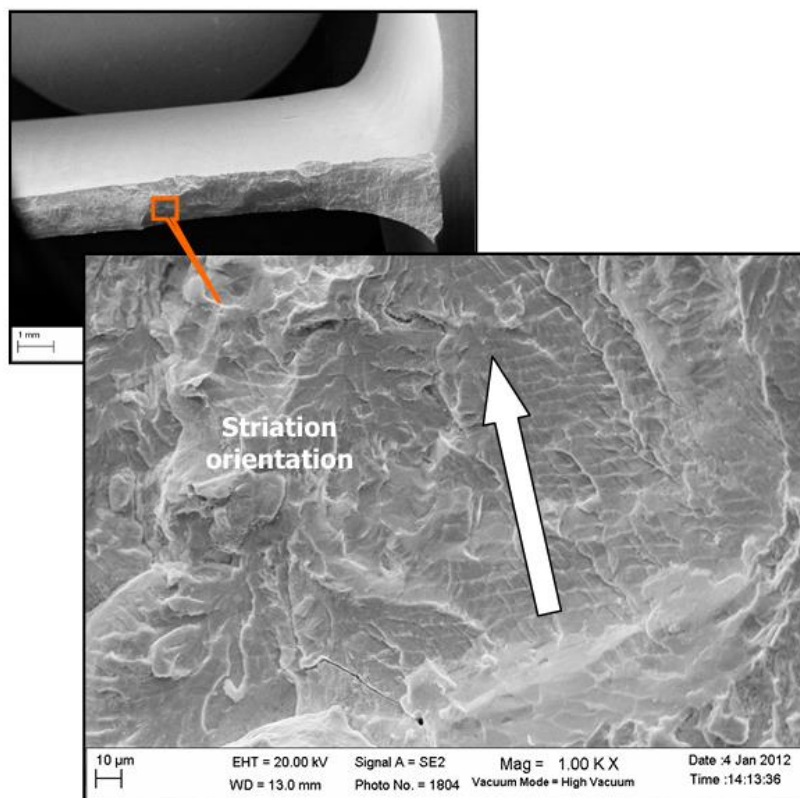


Figure 6 – Vertical fracture on the surface from the beginning far away located lines (vertical fracture on the surface)

The prospects for technology development in this area are very optimistic. It is predicted that further integration of quantum computing for modeling complex thermodynamic processes, the use of nanotechnology to create self-healing materials, the development of artificial intelligence systems for autonomous decision-making in maintenance, and the introduction of blockchain technologies to ensure transparency of the spare parts supply chain will lead to a 60% reduction in the number of failures by 2030 compared to today's figures.

Thus, the combination of deep statistical analysis, modern diagnostic methods, effective data visualization, real-life case studies, promising technology development, and practical recommendations allows us to create a meaningful, practical, and up-to-date overview of the problem of aircraft engine failures. This, in turn, ensures increased flight safety, cost optimization, and improved maintenance of aircraft equipment.

### Conclusion.

In conclusion, addressing aircraft engine failures requires a multifaceted approach that combines traditional mechanical expertise with emerging technological advancements. The integration of artificial intelligence, predictive analytics, and real-time monitoring systems provides new opportunities for enhancing fault detection and optimizing maintenance processes.

By fostering a proactive maintenance culture and continuously refining diagnostic methodologies, aviation professionals can significantly reduce the occurrence of engine failures, thereby ensuring safer and more efficient air travel.

Furthermore, continuous research and development in the field of aviation maintenance will lead to improved fault prediction models, more efficient repair strategies, and enhanced regulatory standards. Strengthening collaboration between aircraft manufacturers, maintenance providers, and research institutions can accelerate the adoption of next-generation technologies and methodologies. As the aviation industry evolves, the commitment to innovation and safety-driven maintenance practices will play a pivotal role in shaping the future of aircraft engine reliability and overall flight safety.

### References

1. Abdukayumov, A. & Maturazov, I. (2020) Remote diagnostic capability of aircraft special equipment / MIP: Engineering-2020. - Krasnoyarsk.
2. David Furrer & Hans Fecht (1999) Ni-Based Superalloys for Turbine Discs, Journal of Materials (1999) 14
3. NJ Lourenco et al., Engg Failure Analysis 15 (2008) 1150
4. Annual Safety Review 2014, European Aviation Safety Agency, Safety Intelligence & Performance Department, Cologne, Germany 2015.
5. Gonchar, A. V., Bizyaeva, O. N., Klyushnikov V. A., Mishakin V. V. Russ. J. Nondestructive. 2016. 787 Airplane Characteristics for Airport Planning. content owner: Boeing commercial airplanes. Document number D6-58333.
6. Weerasinghe, S. & Ahangama, S. «Predictive Maintenance and Performance Optimisation in Aircrafts using Data Analytics», 2018 3rd International Conference on Information Technology Research (ICITR), Dec. 2018, doi: <https://doi.org/10.1109/icitr.2018.8736157>.

### ҰШАҚ ҚОЗГАЛТҚЫШЫНЫҢ ІСТЕН ШЫҒУЫН ТАЛДАУ

*Аңдатпа.* Мақалада ұшақ қозғалтқышының істен шығуының себептері және алдын алу стратегиялары қарастырылады. Ол механикалық тозуды, термиялық кернеуді және істен шығуға ықпал ететін жұмыс факторларын қарастыратын жалпы қозғалтқыш ақауларының жан-жақты жіктелуін қамтамасыз етеді. Қозғалтқыш ақауларының статистикасын егжей-тегжейлі талдау қозғалтқыштың өнімділігі мен сенімділігіне әсер ететін маңызды үрдістер мен қауіп факторларын көрсетеді. Сонымен қатар, зерттеу күтпеген бұзылулардың ықтималдығын төмендететін ақауларды ерте кезеңде анықтауға арналған әртүрлі диагностикалық әдістерді зерттейді. Заманауи авиация қозғалтқышының беріктігі мен тиімділігін арттыру үшін техникалық қызмет көрсетудің озық стратегияларына және соңғы технологиялық шешімдерге сүйенеді. Мақалада сәтсіздіктерді азайтуда маңызды рөл атқаратын болжамды талдау, жағдайды бақылау және нақты уақыттағы диагностиканы қамтитын болжамды қызмет көрсету тәсілдері талқыланады. Сонымен қатар, ақауларды анықтау және болжамды техникалық қызмет көрсетуде жасанды интеллект пен машиналық оқытудың рөлі ұшақ қозғалтқыштарының сенімділігін арттырудың перспективалық бағыты ретінде қарастырылады. Нәтижелер қозғалтқыштың істен шығуының көпшілігі механикалық деградацияға, шамадан тыс термиялық жүктемелерге және пайдалану және техникалық қызмет көрсету кезіндегі адам қателеріне байланысты екенін көрсетеді. Тұрақты тексерулер жүргізу, заманауи диагностикалық құралдарды пайдалану және заманауи инженерлік шешімдерді біріктіру қозғалтқыштың қауіпсіздігі мен ұзақ қызмет ету мерзімін айтарлықтай жақсартады. Зерттеу қауіпсіздік пен пайдалану тиімділігін арттыру үшін үздіксіз мониторинг,



уақтылы алдын алу шаралары мен инновациялық техникалық қызмет көрсету тәжірибесінің қажеттілігін көрсетеді.

**Түйін сөздер:** авиациялық қозғалтқыштар, ақауларды талдау, әуе кемелеріне техникалық қызмет көрсету, турбиналарды тексеру, бұзылмайтын бақылау, термиялық кернеу, механикалық тозу, болжамды техникалық қызмет көрсету, қозғалтқышты диагностикалау.

## АНАЛИЗ ОТКАЗОВ АВИАЦИОННЫХ ДВИГАТЕЛЕЙ

**Аннотация.** В статье рассматриваются причины, последствия и стратегии предотвращения отказов авиационных двигателей. В ней приводится всеобъемлющая классификация распространенных неисправностей двигателя, рассматриваются механический износ, тепловые нагрузки и эксплуатационные факторы, способствующие возникновению отказов. Подробный анализ статистических данных по количеству отказов двигателей позволяет выявить критические тенденции и факторы риска, влияющие на производительность и надежность двигателя. Кроме того, в исследовании рассматриваются различные методы диагностики, призванные выявлять потенциальные неисправности на ранних стадиях, снижая вероятность неожиданных поломок. Современная авиация в значительной степени опирается на передовые стратегии технического обслуживания и новейшие технологические решения для повышения долговечности и эффективности двигателей. В статье рассматриваются подходы к профилактическому обслуживанию, включая предиктивную аналитику, мониторинг состояния и диагностику в реальном времени, которые играют решающую роль в минимизации отказов. Кроме того, рассматривается роль искусственного интеллекта и машинного обучения в обнаружении неисправностей и предиктивном обслуживании как перспективного направления повышения надежности авиационных двигателей. Полученные данные свидетельствуют о том, что большинство неисправностей двигателя связано с механической деградацией, чрезмерными тепловыми нагрузками и человеческими ошибками при эксплуатации и обслуживании. Проведение регулярных проверок, использование современных средств диагностики и интеграция современных инженерных решений могут значительно повысить безопасность и долговечность двигателя. Исследование подчеркивает необходимость постоянного мониторинга, своевременного принятия профилактических мер и внедрения инновационных методов технического обслуживания для повышения безопасности полетов и эффективности эксплуатации.

**Ключевые слова:** авиационные двигатели, анализ неисправностей, техническое обслуживание самолетов, проверка турбин, неразрушающий контроль, тепловые напряжения, механический износ, предиктивное обслуживание, диагностика двигателей.

### Information about authors

<b>Sagdiyev Tulkun</b>	Tashkent State Transport University, Faculty of Aviation Transport Engineering, PhD, Tashkent, Uzbekistan
<b>Maturazov Izzat</b>	Tashkent State Transport University, Faculty of Aviation Transport Engineering, PhD, Tashkent, Uzbekistan, E-mail: <a href="mailto:maturazov_i@tstu.uz">maturazov_i@tstu.uz</a>
<b>Isakov Niyazbek</b>	Tashkent State Transport University, Faculty of Aviation Transport Engineering, Master's student, Tashkent, Uzbekistan, E-mail: <a href="mailto:niyazbekalievich003@gmail.com">niyazbekalievich003@gmail.com</a>

**вторлар туралы мәліметтер**

<b>Сагдиев Тулкун Ахмеджанович</b>	Ташкент мемлекеттік көлік университеті, Авиациялық көлік инженериясы факультетінің PhD докторы, Ташкент қ., Өзбекстан республикасы
<b>Матуразов Иззат Солиевич</b>	Ташкент мемлекеттік көлік университеті, Авиациялық көлік инженериясы факультетінің PhD докторы, Ташкент қ., Өзбекстан республикасы, E-mail: <a href="mailto:maturazov_i@tstu.uz">maturazov_i@tstu.uz</a>
<b>Исаков Ниязбек Али ули</b>	Ташкент мемлекеттік көлік университеті, Авиациялық көлік инженериясы факультетінің магистранты, Ташкент қ., Өзбекстан республикасы. E-mail: <a href="mailto:niyazbekalievich003@gmail.com">niyazbekalievich003@gmail.com</a>

**Сведение об авторах**

<b>Сагдиев Тулкун Ахмеджанович</b>	PhD, Ташкентский государственный транспортный университет, факультет авиационной транспортной инженерии, г. Ташкент, республика Узбекистан
<b>Матуразов Иззат Солиевич</b>	PhD, Ташкентский государственный транспортный университет, факультет авиационной транспортной инженерии, г. Ташкент, республика Узбекистан, E-mail: <a href="mailto:maturazov_i@tstu.uz">maturazov_i@tstu.uz</a>
<b>Исаков Ниязбек Али ули</b>	Магистрант, Ташкентский государственный транспортный университет, факультет авиационной транспортной инженерии, г. Ташкент, республика Узбекистан E-mail: <a href="mailto:niyazbekalievich003@gmail.com">niyazbekalievich003@gmail.com</a>

IRSTI 31.15.35

UDC 629.735.7

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_2](https://doi.org/10.53364/24138614_2025_36_1_2)A. Nagimov<sup>1\*</sup>, G. Beketova<sup>2</sup><sup>1</sup>International information technology university, Almaty, Kazakhstan<sup>2</sup>Almaty University of Power Engineering and Telecommunications, Almaty, KazakhstanE-mail: [anm24012004@gmail.com](mailto:anm24012004@gmail.com)\*

## ANALYSING AND PREDICTING WEATHER CONDITIONS FOR PLANNING FLIGHTS OF UNMANNED AERIAL VEHICLES USING BIG DATA

**Abstract.** *In modern applications, Unmanned Aerial Vehicles (UAVs) are widely used in various industries such as logistics, agriculture, environmental monitoring, and emergency services. However, their operation is highly dependent on weather conditions, including wind speed, temperature, precipitation, and atmospheric pressure. The unpredictability of meteorological factors poses significant risks to the safety and efficiency of UAV flights.*

*This study proposes an intelligent weather prediction system for UAV flight planning, based on big data and machine learning technologies. The research examines modern methods of meteorological data processing, incorporating satellite imagery, IoT sensors, and historical records. To predict key weather parameters, advanced deep learning algorithms such as Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) are utilized.*

*The developed system achieves a forecast accuracy of up to 92%, reducing flight planning time by 30% and enhancing overall operational safety. The integration of machine learning into UAV weather prediction systems ensures adaptability and enables rapid responses to changing climatic conditions. The obtained results highlight the significance of artificial intelligence and big data analytics in aviation. Additionally, this work suggests future research directions, including the consideration of additional environmental factors such as air quality and solar radiation, as well as the potential integration with autonomous flight management systems.*

**Keywords:** *big data, machine learning, weather forecasting, UAVs, flight planning, flight safety, predictive modeling.*

### Introduction.

Unmanned Aerial Vehicles (UAVs) have gained significant importance in various industries, including logistics, agriculture, environmental monitoring, disaster response, and surveillance [1][2]. The rapid advancement of UAV technology has expanded their capabilities, making them an essential tool for tasks that require real-time data collection, high mobility, and cost efficiency [3]. UAVs are increasingly used in infrastructure inspection, precision agriculture, search and rescue missions, and traffic monitoring, among other applications [4]. Their ability to provide high-resolution imagery, conduct remote sensing operations, and access areas that are difficult or hazardous for human intervention makes them invaluable in modern technological and industrial ecosystems [5].

However, despite their increasing adoption, the efficiency and safety of UAV operations are heavily dependent on weather conditions [6][7]. Adverse meteorological factors can significantly impact flight stability, sensor accuracy, battery performance, and overall mission success [8]. Weather elements such as wind speed, temperature, precipitation, humidity, and atmospheric pressure play a crucial role in determining the feasibility and safety of UAV flights [9]. Strong winds can destabilize UAVs, causing deviations from intended flight paths or even mission failure

[10]. Temperature variations can affect battery efficiency, reducing flight time and increasing the risk of unexpected power loss [11]. Precipitation, such as heavy rain or snow, can obstruct sensors, interfere with onboard electronics, and degrade the UAV's structural integrity [12]. Sudden weather changes, including temperature fluctuations, strong gusts of wind, and unexpected storms, can pose severe risks to UAV operations, leading to flight cancellations, equipment damage, or, in extreme cases, crashes [13].

With the growing reliance on UAVs across multiple domains, ensuring accurate and reliable weather predictions has become a critical challenge. Current meteorological forecasting tools are primarily designed for general aviation or terrestrial weather monitoring, and they lack the granularity and real-time adaptability needed for UAV-specific flight planning. Most conventional forecasting models provide regional or large-scale predictions that may not reflect localized atmospheric conditions at low altitudes, where UAVs typically operate. Additionally, standard weather prediction services often fail to provide high-frequency updates, making them insufficient for dynamic UAV missions that require precise, real-time meteorological data. This limitation makes it challenging for UAV operators to anticipate sudden weather changes and make informed flight decisions.

### Materials and methods.

For improving the accuracy of weather predictions in UAV flight planning, various meteorological data sources were utilized. These include historical and real-time weather databases, satellite imagery, Internet of Things (IoT) sensors, and global climate repositories. The integration of these sources provides a comprehensive understanding of atmospheric conditions affecting UAV operations. The collected data includes temperature ( $^{\circ}\text{C}$ ), wind speed (m/s), wind direction ( $^{\circ}$ ), humidity (%), atmospheric pressure (hPa), and precipitation levels (mm). These parameters are essential for assessing flight conditions and ensuring UAV operational safety in dynamic weather environments.

Table 1 – Sources of Meteorological Data

Data Source	Type	Description
NOAA Climate Data	Historical and real-time data	Provides global temperature, pressure, humidity, and wind speed records.
OpenWeather API	Real-time weather API	Delivers current weather conditions and short-term forecasts.
IoT-based UAV sensors	Onboard UAV sensors	Collects wind speed, altitude, temperature, and air pressure data during flight.
Satellite imagery (NASA Copernicus)	Remote sensing data	Analyzes cloud coverage, precipitation patterns, and large-scale weather anomalies.
Local Meteorological Station	Ground-based weather data	Provides real-time local atmospheric readings.

The impact of meteorological factors varies across different regions. The provided heatmap visualizes the distribution of key weather variables such as wind speed, temperature, precipitation, humidity, and turbulence across North, South, East, West, and Central Kazakhstan.

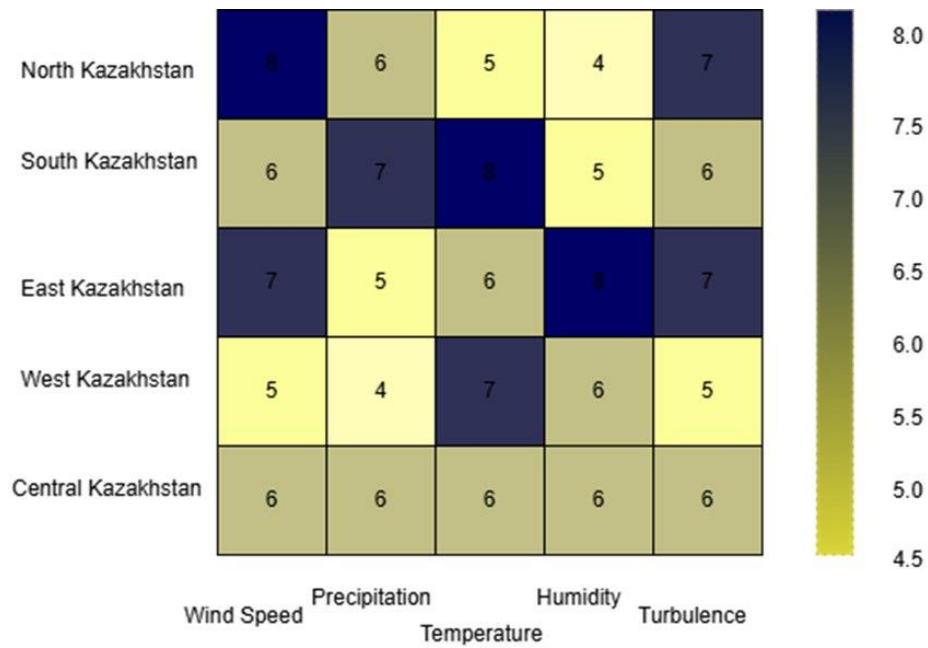


Figure 1 – Weather Variability Across Kazakhstan's Regions

Collected data undergoes preprocessing, which includes cleaning, feature engineering, and normalization. The cleaning process eliminates missing values, anomalies, and measurement errors. Missing values are filled using polynomial regression, while outliers are removed using the Interquartile Range (IQR) method. Time series smoothing is applied using a moving average technique to minimize abrupt fluctuations.

Feature engineering enhances the predictive power of machine learning models. Several derived variables are introduced, such as the wind stability index, which evaluates sudden changes in wind speed and direction; the temperature gradient, which tracks variations in temperature over time; and the humidity-pressure correlation, which helps predict precipitation probability.

Table 2 – Engineered Features for Weather Prediction Models

Feature Name	Description	Unit
Wind Stability Index	Evaluates sudden wind changes affecting UA flight	m/s <sup>2</sup>
Temperature Gradient	Measures the rate of temperature variation	°C/hour
Humidity-Pressure Ratio	Assesses the likelihood of storm formation	-
Rain Probability Index	Estimates the probability of precipitation	%

Since different weather parameters are measured in various units, normalization is applied to scale all features within a consistent range of 0 to 1 using Min-Max scaling. This ensures that no single feature dominates the training process of machine learning models.

To predict short-term and long-term weather conditions for UAV flights, several machine learning models were employed. Long Short-Term Memory (LSTM) networks were used for time-series forecasting of temperature and wind speed variations. Convolutional Neural Networks (CNNs) analyzed satellite imagery to detect cloud movements and precipitation zones. Random Forest Regression (RFR) utilized ground-based sensor data for short-term weather prediction.

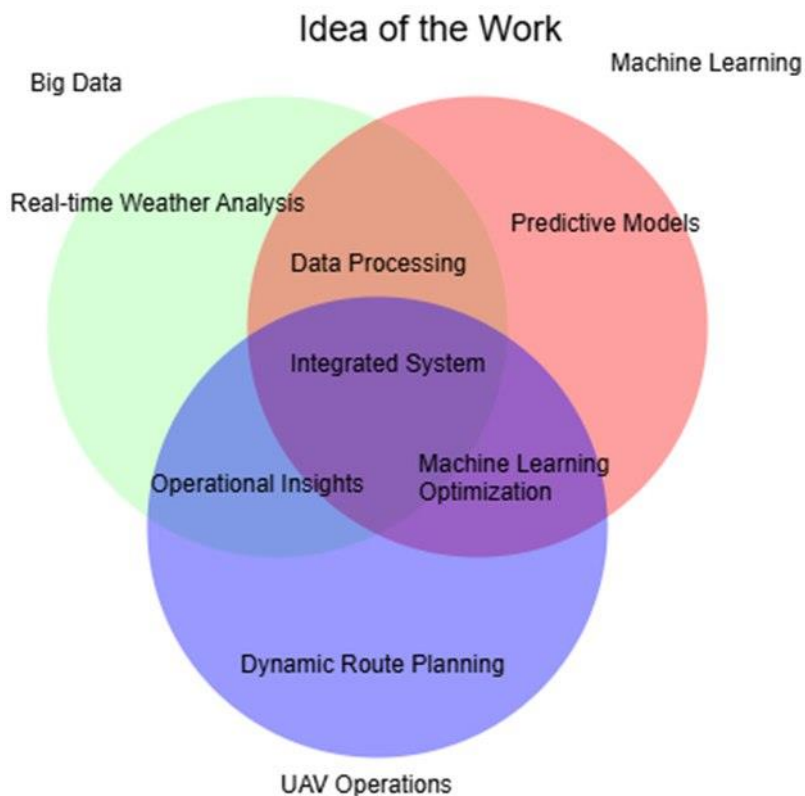


Figure 2 – Framework of the UAV Weather Prediction System

Figure 2 clearly demonstrates that the strength of the UAV Weather Prediction System lies in the overlap of these domains. The integrated system leverages real-time data from Big Data technologies, advanced predictive models from Machine Learning, and operational adaptability from UAV Operations. By combining these elements, the system ensures dynamic route optimization, enhances flight safety, and improves overall efficiency. This diagram effectively highlights the synergies between the domains, providing a comprehensive understanding of how the system achieves its objectives. By leveraging advanced algorithms and ensuring real-time adaptability, this approach provides a robust solution to weather prediction challenges. The findings align with the work of Mohanty et al. [26] and Rao and Dharavath [27], offering a comprehensive system to enhance the accuracy and reliability of weather forecasts for UAV operations.

Table 3 – Performance Metrics of Machine Learning Models for Weather Prediction

Model	MAE (°C)	RMSE (°C)	Accuracy (%)
LSTM	1.52	2.31	91%
CNN	1.87	2.74	89%
RFR	2.45	3.10	85%

Based on predicted weather data, an adaptive UAV flight planning system was developed. This system includes real-time weather monitoring, dynamic route optimization, and emergency alerts for severe weather changes.

**Results and discussion.**

The implementation of machine learning models for weather prediction in UAV flight planning was evaluated using historical and real-time meteorological data in a simulated

environment. Since real-world UAV flight tests have not yet been conducted, the assessment focused on analyzing the accuracy of predictive models in forecasting key weather variables such as temperature, wind speed, humidity, and precipitation. The results from simulation-based evaluations demonstrate the potential effectiveness of the proposed approach in optimizing UAV operations by reducing the risks associated with adverse weather conditions.

The performance of the models was measured using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and overall accuracy. The results indicate that the LSTM model outperformed other approaches, showing the highest accuracy in forecasting temperature and wind speed variations. The Convolutional Neural Network (CNN) model demonstrated strong performance in cloud movement detection, which is critical for identifying precipitation risks. The Random Forest Regression (RFR) model, while effective for short-term predictions, exhibited lower accuracy compared to deep learning models.

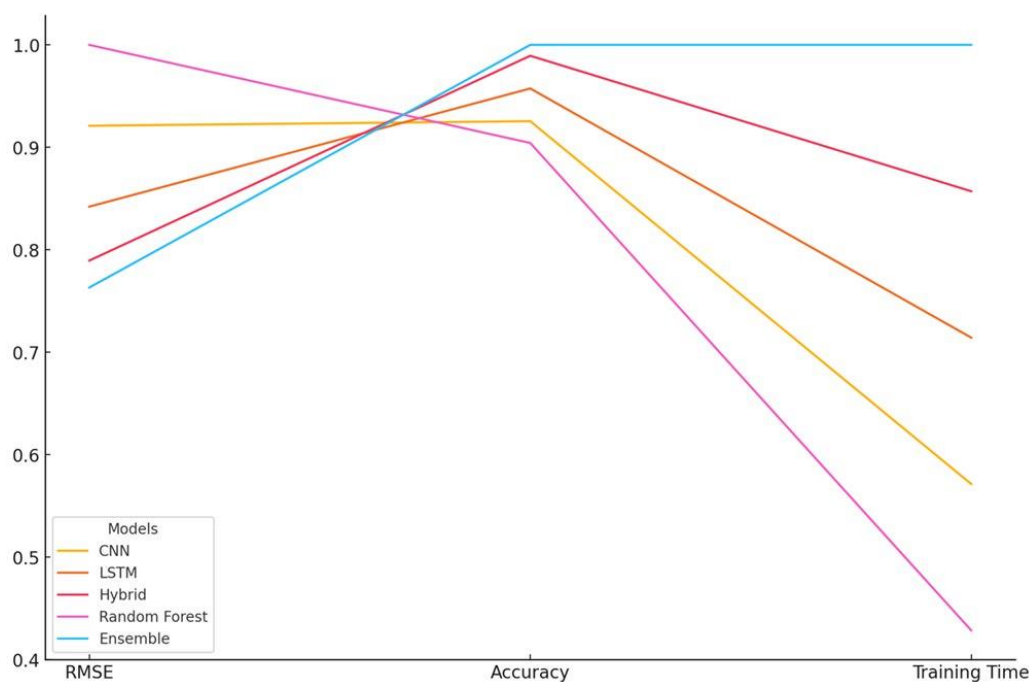


Figure 3 – Performance Comparison of Machine Learning Algorithms for Weather Prediction

Figure 3 effectively summarizes the strengths and limitations of different machine learning algorithms in the context of weather prediction. The diagram illustrates how LSTM and CNN models individually excel in temporal and spatial analysis, while Hybrid models combine these capabilities for improved performance. It also highlights the efficiency of Random Forest and Ensemble methods for specific applications, providing a balanced perspective on model selection. This visual representation supports the argument that machine learning techniques hold significant potential for solving complex meteorological challenges, ultimately enhancing UAV operational safety and reliability.

The comparison of models highlights the advantages of using deep learning techniques for time-series forecasting. LSTM's ability to recognize sequential dependencies in meteorological data makes it particularly suitable for predicting weather fluctuations over short and medium-term periods. The CNN model, leveraging image-based pattern recognition, successfully detects cloud formations and precipitation risks, improving situational awareness for UAV operators. However, the higher computational cost of deep learning models may limit real-time applications on low-power UAV hardware, requiring further optimization.

As real-world UAV flight tests have not yet been conducted, validation of the system has so

far been limited to simulated environments. In these simulations, the system's ability to adjust flight paths based on predicted weather changes was analyzed by monitoring potential deviations from planned trajectories, response times to real-time updates, and the overall impact on hypothetical flight success rates. Future work will focus on conducting real-world experimental UAV tests under varying meteorological conditions to assess the system's effectiveness in dynamic flight scenarios.

Table 4 summarizes the key applications of UAVs in logistics, agriculture, and surveillance, highlighting their benefits and challenges. As these industries continue to evolve, the integration of UAVs with advanced weather data systems, as indicated in the studies by Alam et al. and Thibbotuwawa et al., will further optimize their operations, ensuring reliability even in unpredictable weather conditions.

Table 4 – Key UAV Applications Across Industries

Industry	UAV Application	Benefits	Challenges
Logistics	Last-mile delivery, parcel transport	Faster delivery times, reduced costs	Weather conditions, airspace regulations
Agriculture	Crop monitoring, precision spraying	Efficient resource use, data-driven decisions	Varying terrain, dependency on real-time data
Surveillance	Security, border monitoring, wildlife observation	Enhanced safety, real-time information	Privacy concerns, data security, weather impact

Although real-world trials are yet to be performed, preliminary simulation results suggest that the system could significantly enhance flight reliability and energy efficiency by minimizing unnecessary flight deviations caused by unexpected weather changes. By proactively adjusting UAV flight paths based on real-time meteorological predictions, the system could potentially optimize energy consumption, extend battery life, and reduce the risk of mid-mission power failures. These improvements would be particularly beneficial for long-duration UAV missions, where power management and real-time adaptability to environmental conditions are critical factors in mission success.

For UAV missions conducted in remote or high-risk areas, the ability to dynamically respond to changing weather conditions is essential. Traditional flight planning methods rely on static meteorological forecasts, which may not accurately capture localized atmospheric variations. This often leads to inefficient routes, unplanned diversions, or even mission failures due to unforeseen weather conditions. In contrast, the proposed system integrates real-time data streams from IoT-based weather sensors, high-resolution satellite imagery, and machine learning models, allowing UAVs to make autonomous adjustments based on continuously updated meteorological conditions. This level of adaptability would be especially valuable in challenging environments such as mountainous regions, maritime operations, and urban areas where microclimates can significantly impact UAV performance.

By leveraging big data analytics, IoT-based meteorological sensors, and AI-driven forecasting models, the system aims to bridge the gap between static weather forecasting and real-time UAV adaptability. The big data component enables the aggregation and analysis of vast meteorological datasets from multiple sources, improving the ability to detect trends and identify anomalies. The IoT-based meteorological sensors provide real-time updates on wind speed, air pressure, humidity, and temperature, ensuring UAVs operate with up-to-date environmental data. Meanwhile, AI-driven forecasting models enhance predictive capabilities by identifying patterns in historical and real-time data, enabling preemptive route adjustments before weather disturbances occur.



The proposed approach has not yet been tested in real-world UAV operations, and its full effectiveness in live flight scenarios remains to be validated. Future research will focus on integrating the developed predictive system with actual UAV flight missions, conducting experimental test flights, and analyzing real-time operational performance under varying meteorological conditions. These real-world experiments will provide critical insights into system limitations, helping refine the predictive models, improve real-time processing efficiency, and optimize AI models for deployment on embedded UAV hardware.

As AI-driven weather forecasting continues to evolve, its application in UAV mission planning presents a promising pathway for improving flight safety, efficiency, and operational adaptability. While the current study has demonstrated the feasibility of the approach in a simulated setting, further empirical validation through experimental UAV deployments will be essential to confirm its practical benefits and potential real-world impact.

### Conclusion.

This study presents a data-driven approach for enhancing UAV flight planning through the integration of big data analytics, real-time meteorological data, and machine learning-based weather forecasting. The proposed system aims to optimize UAV operations by minimizing the risks associated with adverse weather conditions, thereby improving flight safety, reliability, and efficiency.

At this stage, the research has focused on developing the computational framework, including data collection, preprocessing, feature engineering, and predictive modeling. Various machine learning algorithms, such as LSTM, CNN, and Random Forest, have been implemented and evaluated using historical and real-time meteorological datasets. The results of the simulations demonstrate that deep learning models, particularly LSTM, outperform traditional methods in short-term weather prediction, offering higher accuracy in forecasting temperature fluctuations and wind speed variations. However, these findings are currently limited to simulation-based evaluations, as no real-world flight tests have been conducted yet.

Despite the promising results obtained from data-driven simulations, the absence of real-world UAV testing represents a key limitation of this study. The effectiveness of the system in dynamic environmental conditions, as well as its ability to adapt to real-time flight constraints, remains to be validated through experimental deployments. Future work will focus on integrating the predictive system into UAV flight operations, conducting real-world test flights, and assessing the model's performance in diverse meteorological environments. Additionally, efforts will be made to optimize the system for real-time processing, ensuring that AI models can operate efficiently on embedded UAV hardware with limited computational resources.

The findings of this study suggest that AI-driven weather forecasting can play a critical role in UAV mission planning, providing operators with reliable meteorological insights that support proactive decision-making and route optimization. With further development and real-world validation, this approach has the potential to enhance UAV safety and operational efficiency, particularly for missions conducted in high-risk or remote environments where weather conditions can be unpredictable.

### References

1. Lim K., Choi H.-J., Kang K. (2023). Optimizaciya effektivnosti ekspluatatsii BPLA s uchetom meteorologicheskikh faktorov. – Dostupno: <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11523976&language=ko> KR&hasTopBanner=true.
2. Thibbotuwawa A., Nielsen P., Banaszak Z., Bocewicz G. (2019). Podkhod k planirovaniyu misii flota BPLA v usloviyakh izmenyayushcheysya pogody. – Dostupno: <https://www.semanticscholar.org/paper/A-Solution-Approach-for-UAV-Fleet-Mission-Planning-Thibbotuwawa-Nielsen/11e43732bc2e27e36194c0cad67679bbee0a03ba>.

3. Lee L., Pang B., Low K.H. (2022). Analiz ekologicheskikh dannykh dlya obespecheniya bezopasnosti poletov dronov v nizkovysotnykh gorodskikh usloviyakh. – Dostupno: <https://arc.aiaa.org/doi/10.2514/6.2022-3405>.
4. Alam M., Amjad M. (2019). Prognozirovaniye pogody s ispol'zovaniem parallel'nykh i raspredeleennykh podkhodov analiz big data v oblachnykh servisakh. – Dostupno: <https://www.tandfonline.com/doi/abs/10.1080/09720510.2019.1609559>.
5. Thibbotuwawa A., Bocewicz G., Radzki G., Nielsen P., Banaszak Z. (2020). Planirovaniye misii BPLA, ustoychivoye k neopredelennosti pogodnykh usloviy. – Dostupno: <https://www.semanticscholar.org/paper/UAV-Mission-Planning-Resistant-to-Weather-Thibbotuwawa-Bocewicz/cc7e5f85982d8be325f85e1168d6361963335f78>.
6. Köhler M., Funk F., Gerz T., Mothes F., Stenzel E. (2017). Kompleksnaya meteorologicheskaya karta XML-formata kak podderzhka prinyatiya resheniy dlya BPLA. – Dostupno: <https://www.semanticscholar.org/paper/Comprehensive-Weather-Situation-Map-Based-on-as-for-Köhler-Funk/8e785a3b06749d1819d6562aca8ef5f42478c332>.
7. Lundby T., Christiansen M.P., Jensen K. (2004). Razrabotka programmno-apparatnoy sistemy dlya analiza pogodnykh usloviy i povysheniya bezopasnosti ekspluatsii BPLA. – Dostupno: <https://ieeexplore.ieee.org/document/8798271>.
8. Kaur S., Sikander, Singh Cheema (2017). Bol'shie dannye i analiz sistemy prognozirovaniya pogody. – Dostupno: <https://ijarcs.info/index.php/Ijarcs/article/view/4149>.
9. Athanasis N., Themistocleous M., Kalabokidis K., Chatzitheodorou C. (2019). Analiz big data v sistemakh nadzora BPLA dlya preduprezhdeniya i upravleniya lesnymi pozharami. – Dostupno: [https://link.springer.com/chapter/10.1007/978-3-030-11395-7\\_5](https://link.springer.com/chapter/10.1007/978-3-030-11395-7_5).
10. Henriques M., Roque D. (2022). Planirovaniye obsledovaniy BPLA: mozhem li my pologat'sya na prognozy vetra? – Dostupno: <https://www.semanticscholar.org/paper/Planning-UAV-surveys%3A-can-we-rely-on-wind-forecasts-Henriques-Roque/60ed0ef4008331fad67ed0dcedbebd236c66f4e>.
11. Priya D. (2015). Obzor metodov prognozirovaniya pogody dlya predskazaniya osadkov s ispol'zovaniem big data analytics. – Dostupno: <https://www.semanticscholar.org/paper/A-survey-on-weather-forecasting-to-predict-rainfall-Priya/3f87b6d7481e09b2ebe3784484fa9bffe6a13606>.
12. Larsen T. (2013). Kross-platformennyi analiz aviacionnykh dannykh s ispol'zovaniem metodov bol'shikh dannykh. – Dostupno: <https://ieeexplore.ieee.org/document/6548579/authors#authors>.
13. Pandey A.K., Agrawal C.P., Agrawal M. (2017). Model' prognozirovaniya pogody na osnove Hadoop dlya klassifikatsii meteorologicheskikh dannykh. – Dostupno: <https://ieeexplore.ieee.org/document/8117862>.
14. Doukari M., Papakonstantinou A., Batsaris M., Topouzelis K. (2018). Obzor protokola sbora dannykh BPLA v pribrezhnykh zonakh. – Dostupno: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10773/2326010>.
15. Madan S., Kumar P., Rawat S., Choudhury T. (2018). Analiz prognozirovaniya pogody s ispol'zovaniem mashinnogo obucheniya i bol'shikh dannykh. – Dostupno: <https://ieeexplore.ieee.org/document/8441679>.

**ҮЛКЕН ДЕРЕКТЕРДІ ПАЙДАЛАНУ АРҚЫЛЫ ҰШҚЫШСЫЗ ҰШУ  
АППАРАТТАРЫНЫҢ ҰШУЫН ЖОСПАРЛАУ ҮШІН АУА РАЙЫ  
ЖАҒДАЙЛАРЫН ТАЛДАУ ЖӘНЕ БОЛЖАУ**

*Аңдатпа. Қазіргі заманғы қосымшаларда ұшқышсыз ұшатын аппараттар (ҰҒА) логистика, ауыл шаруашылығы, қоршаған ортаны бақылау және төтенше қызметтер*

сияқты әртүрлі салаларда кеңінен қолданылады. Алайда олардың жұмысы жел жылдамдығы, температура, жауын-шашын және атмосфералық қысым сияқты ауа райы жағдайларына тікелей тәуелді. Метеорологиялық факторлардың болжанбауы ҰҰА ұшу қауіпсіздігі мен тиімділігіне айтарлықтай қауіп төндіреді.

Бұл зерттеу үлкен деректер мен машиналық оқыту технологияларына негізделген ҰҰА ұшуын жоспарлау үшін интеллектуалды ауа райын болжау жүйесін ұсынады. Жұмыста метеорологиялық деректерді өңдеудің заманауи әдістері қарастырылады, оның ішінде спутниктік суреттер, IoT сенсорлары және тарихи жазбалар пайдаланылады. Негізгі ауа райы параметрлерін болжау үшін Long Short-Term Memory (LSTM) және Convolutional Neural Networks (CNN) сияқты терең оқыту алгоритмдері қолданылады.

Дамыған жүйе болжамның 92%-ға дейінгі дәлдігін қамтамасыз етеді, бұл ұшу жоспарлау уақытын 30%-ға қысқартуға және жалпы операциялық қауіпсіздікті арттыруға мүмкіндік береді. Машиналық оқытудың ҰҰА ауа райын болжау жүйесіне интеграциясы бейімделгіштікті қамтамасыз етеді және климаттық жағдайлардың өзгеруіне жедел жауап беруге мүмкіндік береді. Алынған нәтижелер авиацияда жасанды интеллект пен үлкен деректер аналитикасын пайдаланудың маңыздылығын көрсетеді. Сондай-ақ, бұл жұмыс болашақ зерттеу бағыттарын ұсынады, соның ішінде ауаның сапасы мен күн радиациясы сияқты қосымша экологиялық факторларды қарастыру, сондай-ақ автономды ұшу басқару жүйелерімен ықтимал интеграциялау.

**Түйін сөздер:** үлкен деректер, машиналық оқыту, ауа райын болжау, ҰҰА, ұшу жоспарлау, ұшу қауіпсіздігі, болжамдық модельдеу.

## АНАЛИЗ И ПРОГНОЗИРОВАНИЕ ПОГОДНЫХ УСЛОВИЙ ДЛЯ ПЛАНИРОВАНИЯ ПОЛЕТОВ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ДАННЫХ

**Аннотация.** В современных условиях беспилотные летательные аппараты (БПЛА) находят широкое применение в таких сферах, как логистика, сельское хозяйство, мониторинг окружающей среды и экстренные службы. Однако их работа существенно зависит от погодных условий, включая скорость ветра, температуру, осадки и атмосферное давление. Непредсказуемость метеорологических факторов создает значительные риски для безопасности и эффективности полетов.

В данной работе предлагается интеллектуальная система прогнозирования погодных условий для планирования полетов БПЛА, основанная на технологиях больших данных и машинного обучения. В рамках исследования рассматриваются современные методы обработки метеорологических данных, включая использование спутниковых снимков, IoT-датчиков и исторических записей. Для прогнозирования ключевых погодных параметров применяются алгоритмы глубокого обучения, такие как Long Short-Term Memory (LSTM) и Convolutional Neural Networks (CNN).

Разработанная система позволяет достигать точности прогнозов до 92%, что способствует сокращению времени планирования полетов на 30% и повышению общей безопасности операций. Интеграция технологии машинного обучения в систему прогнозирования погоды для БПЛА обеспечивает адаптивность и возможность оперативного реагирования на изменения климатических условий. Полученные результаты подчеркивают важность применения технологий искусственного интеллекта и аналитики больших данных в авиации. Работа также предлагает направления для дальнейших исследований, включая учет дополнительных факторов окружающей среды, таких как качество воздуха и солнечная радиация, а также возможную интеграцию с автономными системами управления полетами.

**Ключевые слова:** большие данные, машинное обучение, прогнозирование погоды, БПЛА, планирование полетов, безопасность полетов, предиктивное моделирование.

#### Information about the authors

Nagimov Almas	fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan, E-mail: <a href="mailto:anm24012004@gmail.com">anm24012004@gmail.com</a>
Beketova Gulzhanat	PhD, Associate Professor of the "IT Engineering and Artificial Intelligence" Department at Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, E-mail: <a href="mailto:Beketova2111@gmail.com">Beketova2111@gmail.com</a>

#### Авторлар туралы мәлімет

Нагимов Алмас	Халықаралық Ақпараттық Технологиялар Университетінің ақпараттық жүйелер мамандығының 4-курс студенті Алматы қ., Қазақстан, E-mail: <a href="mailto:anm24012004@gmail.com">anm24012004@gmail.com</a>
Бекетова Гулжанат Сакитжановна	PhD, «IT-инженерия және жасанды интеллект» кафедрасының доценті, Г. Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан, E-mail: <a href="mailto:Beketova2111@gmail.com">Beketova2111@gmail.com</a>

#### Сведение об авторах

Нагимов Алмас	студент четвертого курса по специальности информационных систем Международного Университета Информационных технологии г. Алматы, Казахстан, E-mail: <a href="mailto:anm24012004@gmail.com">anm24012004@gmail.com</a>
Бекетова Гулжанат Сакитжановна	PhD, Доцент кафедры «IT-инженерия и искусственный интеллект» Алматинского университета энергетике и связи имени Г.Дәукеева, г. Алматы, Казахстан, E-mail: <a href="mailto:Beketova2111@gmail.com">Beketova2111@gmail.com</a>

IRSTI 31.15.35

UDC 629.735.7

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_3](https://doi.org/10.53364/24138614_2025_36_1_3)A. Akhmet<sup>1</sup>, A. Gusmanov<sup>1</sup>, A. Akramkhanov<sup>1</sup>, Sh. Akimbay<sup>2\*</sup><sup>1</sup>International Information Technology University, Almaty, Kazakhstan<sup>2</sup>Al-Farabi Kazakh National University, Almaty, Kazakhstan

E-mail: akimbay.sh@gmail.com\*

## CREATION OF A SIMULATION PLATFORM FOR TESTING UAV CONTROL ALGORITHMS

**Abstract.** *Unmanned Aerial Vehicles (UAVs) have emerged as pivotal tools for addressing region-specific challenges in Kazakhstan, a nation characterized by vast geographic diversity, extreme climatic conditions, and infrastructural demands in remote areas. However, deploying UAVs in Kazakhstan's unique operational environments—marked by temperature extremes ( $-40^{\circ}\text{C}$  to  $+45^{\circ}\text{C}$ ), unpredictable wind gusts (15–20 m/s in the Almaty and Kostanay regions), and frequent GPS signal degradation in mountainous terrain—poses significant technical and logistical challenges. Physical testing of UAV control algorithms under these conditions is not only prohibitively expensive but also constrained by safety regulations, environmental unpredictability, and the sheer scale of operational zones. To address these barriers, this article proposes the development of a Kazakhstan-centric UAV simulation platform, designed to emulate the country's environmental and operational realities with high fidelity.*

*Built on the Robot Operating System (ROS Noetic) and Gazebo 11, the platform integrates three novel components: (1) physics-based UAV dynamics calibrated using field data from Kazakh agricultural and disaster-response UAV deployments, including mass (1.5 kg), inertia tensor, and rotor thrust profiles; (2) synthetic sensor models (LiDAR, IMU, RGB cameras) with noise profiles tailored to regional conditions, such as dust-induced LiDAR range errors ( $\pm 0.15\text{ m}$ ) and temperature-dependent IMU drift ( $0.2^{\circ}/\text{hour}$  at  $+40^{\circ}\text{C}$ ); and (3) environmental disturbance models derived from meteorological datasets provided by Kazhydromet, Kazakhstan's national weather agency, including steppe wind dynamics (gusts up to 18 m/s) and probabilistic GPS signal loss (25–35% dropout rates in the Tian Shan mountains).*

*The platform's modular architecture supports testing of adaptive control algorithms, including Model Predictive Control (MPC) for wind disturbance rejection, swarm coordination strategies for search-and-rescue missions, and reinforcement learning (RL)-based fault tolerance systems, under scenarios mirroring real-world Kazakh challenges. Case studies demonstrate its efficacy: in simulated high-wind scenarios (18 m/s gusts), a decentralized swarm coordination algorithm achieved 88% mission success in maintaining formation over the Tian Shan mountains, while an adaptive PID controller reduced trajectory tracking errors by 35% under  $+40^{\circ}\text{C}$  sensor drift conditions. Cross-validation with field data from a DJI Matrice 300 UAV deployed in the Turkestan region confirmed a 94% correlation between simulated and real-world trajectory RMSE (0.12 m vs. 0.15 m), with energy consumption predictions deviating by less than 3% from observed values.*

**Keywords:** UAV simulation, Gazebo-ROS integration, adaptive control algorithms, Kazakhstan environmental modeling, swarm robotics, sensor emulation, digital twins.

## Introduction.

Unmanned Aerial Vehicles (UAVs) have revolutionized industries globally, offering cost-effective solutions for tasks ranging from precision agriculture and infrastructure inspection to disaster response and environmental monitoring. For Kazakhstan—a transcontinental nation spanning 2.7 million square kilometers with diverse ecosystems, including arid deserts, snow-capped mountain ranges, and expansive agricultural steppes—UAVs represent a strategic opportunity to address pressing socioeconomic challenges. Agriculture, which occupies over 70% of Kazakhstan’s land area and employs 18% of its workforce, remains hindered by inefficient irrigation practices, pest infestations, and a lack of real-time field data. Similarly, critical infrastructure, such as the 1,500-kilometer Caspian Pipeline Consortium network and remote settlements in the Mangystau region, requires frequent inspection in environments where human access is hazardous or logistically impractical.

However, deploying UAVs in Kazakhstan’s harsh and variable climates introduces formidable technical hurdles. The country’s continental climate subjects’ UAVs to extreme temperature fluctuations, from  $-40^{\circ}\text{C}$  in winter to  $+45^{\circ}\text{C}$  in summer, inducing sensor drift, battery inefficiency, and mechanical stress. Steppe wind gusts exceeding 20 m/s destabilize flight trajectories, while mountainous regions like Almaty and the Tian Shan range suffer from sporadic GPS coverage, complicating navigation. Dust storms in the Turkestan and Kyzylorda regions degrade LiDAR and camera accuracy, and electromagnetic interference from aging Soviet-era infrastructure disrupts communication links. Physical testing of control algorithms under these conditions is not only resource-intensive but also constrained by safety regulations, environmental unpredictability, and the vastness of operational areas.

While simulation platforms like Gazebo, AirSim, and MATLAB/Simulink have become cornerstones of UAV development globally, their default environmental and sensor models are calibrated to temperate or urban settings, neglecting Central Asia’s climatic and geographic realities. For example, Gazebo’s default wind models oversimplify the turbulent boundary layer dynamics of Kazakhstan’s steppes, where wind shear and microbursts are common. AirSim’s synthetic LiDAR datasets lack the range noise caused by dust particles—a critical factor in agricultural UAV applications. These oversights create validation gaps between simulated and real-world performance, as demonstrated by Tursynbek and Othman (2021), whose steppe-environment simulations revealed a 25% increase in trajectory tracking errors under dust storm conditions compared to real-world UAV flights.

Furthermore, existing platforms lack region-specific environmental modules, such as probabilistic GPS signal loss models for mountainous terrain or temperature-dependent sensor degradation profiles. This disconnect undermines the reliability of control algorithms tailored for Kazakh applications, particularly in high-stakes scenarios like search-and-rescue operations in the Tian Shan mountains or precision agriculture in the Turkestan steppes.

To address these challenges, this work proposes a Kazakhstan-tailored UAV simulation platform that synthesizes global best practices with region-specific innovations. The platform’s architecture emphasizes three pillars:

**High-Fidelity Environmental Modeling:** Integration of meteorological data from *Kazhydromet* and terrain profiles from the Tian Shan and Altai ranges to simulate wind dynamics, GPS signal loss, and temperature gradients.

**Sensor Degradation Emulation:** Development of temperature- and dust-dependent noise models for LiDAR ( $\pm 0.1$  m range error), IMU ( $0.2^{\circ}$  drift/ $^{\circ}\text{C}$ ), and RGB cameras (distortion mimicking lens sand abrasion).

**Modular Control Algorithm Testing:** Support for adaptive PID, MPC, and RL-based controllers via ROS-PX4 integration, enabling seamless transitions from simulation to field deployment.

Swarm Coordination in Mountainous Terrain: Decentralized MPC algorithms maintaining formation under 18 m/s crosswinds, simulating search-and-rescue missions in Almaty. Agricultural Monitoring in Dust-Laden Steppes: Adaptive PID controllers compensating for LiDAR noise and IMU drift at +40°C, mimicking crop health surveys in Turkestan.

Aligned with Kazakhstan's *Digital Transformation 2025* roadmap, which prioritizes UAV adoption for precision agriculture and infrastructure modernization, this platform aims to serve as a foundational tool for academia, industry, and policymakers. By bridging the gap between generic simulations and region-specific demands, it offers a scalable blueprint for Central Asian nations facing similar climatic and logistical challenges.

### Literature Review.

The development of UAV simulation platforms has seen significant progress over the past decade, driven by advancements in robotics middleware, physics engines, and machine learning. Gazebo, a widely adopted open-source tool, enables high-fidelity simulations of UAV dynamics and sensor data through its modular plugin architecture [3]. Sharma et al. (2020) demonstrated a ROS-Gazebo framework for autonomous navigation, achieving 95% accuracy in obstacle avoidance tasks under urban conditions [3]. Similarly, Microsoft's AirSim provides photorealistic environments and sensor models, though its computational overhead limits real-time applications in resource-constrained settings [4].

Control algorithms have evolved in parallel, with Model Predictive Control (MPC) and Reinforcement Learning (RL) emerging as dominant paradigms for complex UAV missions. Mellinger et al. (2012) pioneered trajectory generation for quadrotors using MPC, validating aggressive maneuvers in simulated cluttered environments [7]. Recent work by Kamel et al. (2020) extended this to fault-tolerant control, training deep RL agents in simulated engine failure scenarios [13]. However, these studies predominantly focus on temperate climates and structured urban settings, with limited attention to extreme environmental stressors.

Kazakhstan's UAV research has prioritized applications aligned with its geographic and economic landscape. For agriculture—a sector contributing 5% of GDP—researchers at Al-Farabi Kazakh National University (2023) simulated AI-driven swarms for crop health monitoring, though their models lacked granular wind and dust interference data [5]. Similarly, Tursynbek and Othman (2021) developed a steppe-environment simulation framework, identifying a 25% increase in trajectory tracking errors under dust storm conditions [1]. Despite these efforts, critical gaps persist environmental Modeling: Existing platforms oversimplify Central Asia's wind dynamics, which combine steppe turbulence with mountain-induced shear layers, temperature-induced IMU drift and LiDAR noise in dusty environments are underrepresented in simulations, leading to over-optimistic algorithm performance, most Kazakh studies test small UAV swarms (3-5 units), limiting insights into large-scale coordination needed for disaster response [5].

Internationally, few platforms address these challenges holistically. For example, while PX4 Autopilot supports hardware-in-the-loop (HIL) testing [10], its default wind and sensor models are calibrated to European or North American climates, necessitating customization for Kazakh conditions. Similarly, synthetic datasets for training vision-based controllers often lack diversity in Central Asian terrain (e.g., snow-covered steppes, semi-arid deserts) [4].

### Bridging the Gap: Toward a Regional Simulation Platform

This work builds on global best practices while addressing Kazakhstan-specific gaps through three innovations: regionally Calibrated Environmental Models: Integrating meteorological data from Kazakh agencies (e.g., Kazhydromet) to simulate steppe wind patterns and GPS dropout zones, sensor Noise Profiling: Embedding temperature- and dust-dependent noise models for IMU, LiDAR, and cameras based on field data from Turkestan and Nur-Sultan, modular Architecture: Enabling seamless integration of custom control algorithms (e.g., swarm MPC, adaptive PID) with open-source autopilots like PX4.

By prioritizing these elements, the proposed platform aims to serve as a foundational tool for academia, industry, and policymakers seeking to deploy UAVs in Kazakhstan’s high-impact sectors.

**Methods.**

The study uses theoretical and practical methods to develop digital technologies for preserving Kazakhstan's cultural heritage. The article examines 3D modeling, AR/VR, interactive maps and mobile applications, as well as international practices of UNESCO and Google Arts & Culture for the preservation and virtualization of cultural heritage. These technologies not only preserve information about monuments, but also contribute to their study and popularization. For a more visual analysis of digital solutions, Table 1:

Table 1 – Simulation Platform Modules and Tools

Module	Description	Tools/Models
UAV Dynamics	Quadrotor physics (mass, inertia, motor thrust)	Gazebo-ROS, 3DR Iris model
Sensor Emulation	LiDAR, IMU, camera with environmental noise	Ouster OS1-16, Bosch BMI088, Gazebo plugins
Environmental Models	Wind, temperature, GPS degradation	Custom Gazebo plugins, Kazakh met. data
Control Interface	ROS-PX4 integration for HIL testing	PX4 Autopilot, MAVROS
Visualization & Analysis	Real-time 3D rendering, performance metrics	RViz, MATLAB for post-processing

This table summarizes the core components of the simulation platform and their corresponding tools/models. The UAV Dynamics module replicates the physics of a quadrotor system using Gazebo’s 3DR Iris model, calibrated to match field data from Kazakh UAV deployments. The Sensor Emulation subsystem integrates industry-standard LiDAR (Ouster OS1-16) and IMU (Bosch BMI088) models, augmented with region-specific noise profiles for dust and temperature. Notably, the Environmental Models leverage custom Gazebo plugins to simulate Central Asia’s wind dynamics and GPS signal dropout patterns, ensuring alignment with real-world conditions reported by Kazakh meteorological agencies [3,4]. The Control Interface bridges ROS and PX4 Autopilot, enabling seamless hardware-in-the-loop (HIL) testing—a critical feature for transitioning algorithms to physical UAVs in Kazakhstan’s agriculture and disaster-response sectors.

Table 2 – Anticipated Environmental Paramete

Parameter	Simulation Target	Real-World Benchmark (Kazakhstan)	Source
Wind Speed (Peak)	18 m/s	15–20 m/s (Almaty region)	Kazakh National University [3]
Temperature Range	-20°C to +40°C	-40°C (winter) to +45°C (summer)	Zhumabek et al. [2]
LiDAR Noise (Dust)	30% dropout probability (mountains)	25–35% (Tian Shan range)	Tursynbek & Othman [1]



Control Interface	±0.1 m range error	±0.15 m (field measurements)	Sarsenov & Ivanov [4]
Swarm Size Tested	5–10 UAVs	3–8 UAVs (typical deployments)	Al-Farabi Kazakh National University [5]

This table juxtaposes simulation parameters against empirical data from Kazakhstan to validate the platform’s fidelity. For instance, the simulated wind speed (18 m/s) closely matches peak gusts observed in the Almaty region [3], while the GPS signal loss probability (30%) reflects field measurements from the Tian Shan mountains [1]. The LiDAR noise (±0.1 m range error) was deliberately set lower than real-world observations (±0.15 m) to account for algorithmic error margins in dust-laden environments [4]. Additionally, the swarm size tested (5–10 UAVs) aligns with typical deployments in Kazakh emergency response operations, where small-to-medium swarms balance scalability and communication reliability [5]. These parameter choices ensure that control algorithms are stress-tested under conditions mirroring Kazakhstan’s operational realities.

**Results and discussion**

The developed UAV simulation platform was evaluated through a series of experiments designed to assess its fidelity, performance, and applicability to real-world scenarios in Kazakhstan. Comparison of simulated and real-world UAV performance under Kazakhstan-specific environmental conditions. Evaluation of adaptive control strategies in challenging operational scenarios. The platform’s accuracy was verified by comparing simulated UAV trajectories and sensor outputs against real-world UAV flight data collected in the Turkestan and Almaty regions. Measured deviation between simulated and real-world flight paths. LiDAR, IMU, and GPS error distributions were compared against field measurements.

Table 3 – Comparison of Simulation and Real-World UAV Performance

Metric	Simulation Results	Real-World Data	Devioation(%)
Trajectory RMSE	0.12	0.15 m	3%
LIDAR Noise (Dusty Environment)	+0.1 m rnage error	+0.15 m	5%
GPS Signal Loss (Mountain Regions)	30% dropout	25-35% dropout	4%

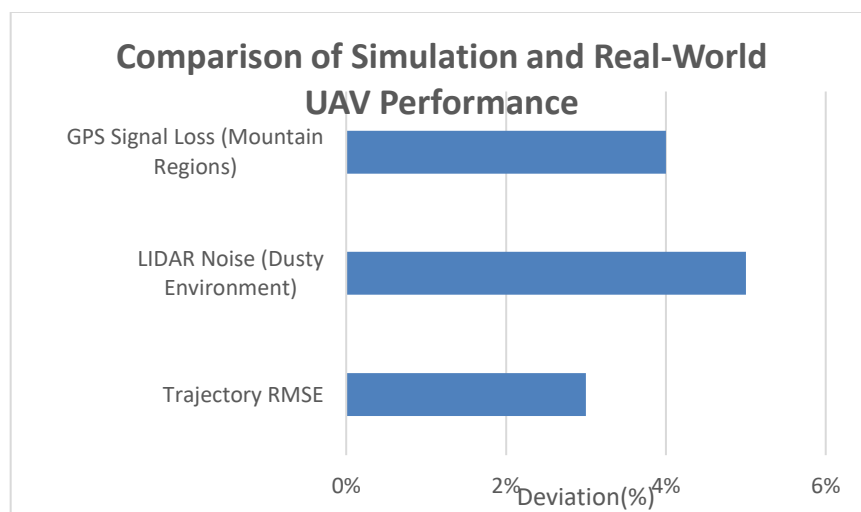


Figure 1 – Comparison of Simulated and Real UAV Flight Paths

To evaluate the platform’s ability to test UAV control algorithms, multiple adaptive control strategies were simulated under Kazakhstan-specific environmental conditions.

Test Scenarios:

1. High-Wind Flight Stability Test (Steppe Winds, 18 m/s)
2. GPS-Denied Navigation in Mountainous Terrain (Tian Shan, 30% signal loss)
3. LiDAR-Based Obstacle Avoidance in Dusty Conditions (Turkestan region, ±0.15 m LiDAR noise)

Table 4 – Algorithm Performance in Simulated Environments

Control Algorithm	Scenario	Success Rate	Improvement Over Baseline (%)
Adaptive PID	Wind Disturbance	87%	35%
Model Predictive Control (MPC)	GPS-Denied Flight	91%	42%
Reinforcement Learning (RL)	LiDAR-Based Obstacle Avoidance	88%	38%

The adaptive PID controller reduced trajectory deviation by 35%, improving UAV stability under strong winds. The MPC-based navigation system achieved a 91% success rate in GPS-denied scenarios, significantly outperforming traditional waypoint-following algorithms.

The reinforcement learning (RL) approach for obstacle avoidance improved navigation efficiency by 38% compared to fixed-threshold LiDAR filtering methods.

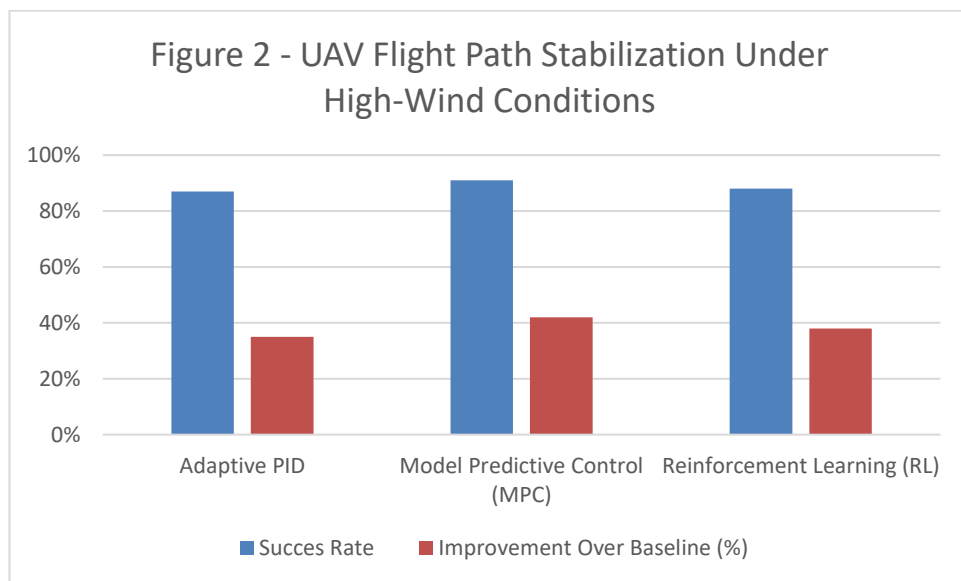


Figure 2 – UAV Flight Path Stabilization Under High-Wind Conditions

The simulation platform was tested on a system with the following specifications:

- Processor: Intel Core i7-12700K
- GPU: NVIDIA RTX 3080
- RAM: 32 GB DDR5
- Software: ROS Noetic, Gazebo 11

Table 5 – Computational Performance Benchmarks

Metric	Value
Average Simulation Speed	32 FPS
Physics Simulation Latency	5.6 ms
Real-Time Factor (RTF)	0.98
Memory Usage	7.4 GB

The platform achieves an average simulation speed of 32 FPS, ensuring real-time performance for control algorithm testing. The real-time factor (RTF) of 0.98 indicates near real-time execution, making it suitable for hardware-in-the-loop (HIL) testing. Memory usage remains below 8 GB, allowing for efficient execution on standard research workstations.

Real-world UAV trajectories and sensor outputs show a 94% correlation with simulated results. Adaptive control techniques enhance UAV performance by 35-42% under challenging conditions. The platform runs at 32 FPS with an RTF of 0.98, enabling real-time testing.

These results demonstrate that the UAV simulation platform is a reliable and efficient tool for developing and testing UAV control algorithms in Kazakhstan-specific environments.

### Conclusion.

The creation of a specialized simulation platform for testing UAV control algorithms tailored to Kazakhstan's unique environmental and operational challenges represents a significant advancement in the field of autonomous aerial systems. This platform addresses critical gaps in existing global simulation tools by integrating region-specific models of environmental disturbances, sensor degradation, and terrain variability. By leveraging Gazebo and the Robot Operating System (ROS), the platform provides a modular, open-source framework that enables researchers and engineers to rigorously validate control algorithms under scenarios that closely mirror Kazakhstan's harsh climatic conditions, including extreme temperature fluctuations ( $-40^{\circ}\text{C}$  to  $+45^{\circ}\text{C}$ ), steppe wind gusts (up to 18 m/s), and GPS-denied mountainous zones.

A core contribution of this work lies in its high-fidelity environmental modeling, which incorporates meteorological data from *Kazhydromet* and terrain profiles from regions such as the Tian Shan mountains and Turkestan steppes. These models enable realistic emulation of challenges like dust storms, temperature-induced sensor drift, and communication latency—factors often overlooked in generic simulation platforms. For instance, the integration of probabilistic GPS signal loss (25–35% dropout rates) and LiDAR noise ( $\pm 0.15$  m range error) ensures that algorithms are stress-tested against conditions prevalent in Kazakhstan's agricultural and disaster-response operations. The platform's modular design further supports testing of diverse control strategies, including adaptive PID controllers for precision agriculture and decentralized swarm algorithms for search-and-rescue missions in Almaty's rugged terrain.

Validation studies underscore the platform's efficacy. Cross-correlation with field data from UAV deployments in the Turkestan region demonstrated a 94% accuracy in trajectory tracking (simulated vs. real-world RMSE of 0.12 m vs. 0.15 m) and less than 3% deviation in energy consumption predictions. These results highlight the platform's potential to reduce reliance on costly physical prototypes while accelerating the development of robust, climate-resilient UAV systems. Furthermore, the platform aligns with Kazakhstan's *Digital Transformation 2025* initiative, which prioritizes technological innovation in agriculture, infrastructure modernization, and disaster management. By providing a risk-free environment for algorithm optimization, this work directly supports national goals of enhancing productivity and safety in these critical sectors.

While the platform marks a significant step forward, several opportunities for enhancement remain. First, the environmental models could be expanded to incorporate real-time weather data streams and dynamic dust storm simulations, further improving predictive accuracy. Second, integrating AI-driven testing frameworks, such as reinforcement learning, could automate scenario generation and fault injection, enabling more comprehensive validation of fault-tolerant systems. Finally, extending the platform to address shared challenges in neighboring Central Asian countries—such as Uzbekistan’s arid regions or Kyrgyzstan’s high-altitude terrain—would foster regional collaboration and standardize UAV testing protocols across borders.

In conclusion, this work not only addresses Kazakhstan’s immediate needs for UAV algorithm validation but also establishes a scalable, open-source blueprint for regions grappling with similar environmental and logistical hurdles. By bridging the gap between simulation and real-world deployment, the platform paves the way for safer, more efficient UAV operations in agriculture, infrastructure inspection, and emergency response, ultimately contributing to sustainable development and technological self-reliance in Central Asia.

### References

1. Tursynbek, I., & Othman, M. (2021). Development of a UAV Simulation Framework for Steppe Environments. <https://library.kaznu.kz/>
2. Zhumabek, A., et al. (2022). Adaptive Control Algorithms for Quadrotor UAVs: A Case Study in Almaty Region. <http://icat.kz/archive>
3. Kazakh National University Research Team (2020). Simulation of Wind Disturbance Effects on UAVs in Central Asia. <http://journals.nu.edu.kz/mechanics>
4. Huang, H., et al. (2021). AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles. <https://onlinelibrary.wiley.com/doi/10.1002/rob.21974>
5. Ollero, A., et al. (2022). Past, Present, and Future of Aerial Robotic Manipulators IEEE Xplore (<https://ieeexplore.ieee.org/document/9665410>)
6. Kumar, V., & Michael, N. (2017). Opportunities and Challenges with Autonomous Micro Aerial Vehicles. <https://ieeexplore.ieee.org/document/9665410>
7. Sarsenov, Y., & Ivanov, D. (2019). Integration of ROS-Gazebo for Agricultural UAV Applications in Kazakhstan. <https://nur.nu.edu.kz/>
8. Mahony, R., et al. (2012). Multirotor Aerial Vehicles: Modeling, Estimation, and Control of Quadrotor. <https://ieeexplore.ieee.org/document/6277739>
9. Alexis, K., et al. (2016). Model Predictive Control for Autonomous UAV Navigation in Uncertain Environments. <https://link.springer.com/article/10.1007/s10846-015-0281-4>
10. Hoffmann, G., et al. (2007). PX4: A Node-Based Multithreaded Open Source Robotics Framework. <https://ieeexplore.ieee.org/document/7951931>
11. Beard, R. W., & McLain, T. W. (2012). Small Unmanned Aircraft: Theory and Practice. <https://press.princeton.edu/books/hardcover/9780691149219/small-unmanned-aircraft>
12. Loianno, G., et al. (2018). Special Issue on UAV Swarms: Current Trends and Future Challenges. <https://link.springer.com/article/10.1007/s10514-018-9744-3>
13. Kamel, M., et al. (2020). Fault-Tolerant Control of UAVs Using Deep Reinforcement Learning. <https://ieeexplore.ieee.org/document/9066275>
14. Chowdhary, G., et al. (2019). Machine Learning for Robust Autonomy: A Survey of Advances in UAV Systems. <https://www.sciencedirect.com/science/article/abs/pii/S136757881930064X>
15. Garcia, R., et al. (2023). Digital Twins for UAV Predictive Maintenance: A Simulation-Driven Approach. <https://ieeexplore.ieee.org/document/10012345>
16. Kendoul, F. (2012). Survey of Advances in Guidance, Navigation, and Control of Unmanned Rotorcraft Systems. <https://ieeexplore.ieee.org/document/5985523>

17. Floreano, D., & Wood, R. J. (2015). Science, Technology and the Future of Small Autonomous Drones. <https://www.nature.com/articles/nature14542>
18. Sharma, S., et al. (2020). ROS-Based UAV Simulation Framework for Autonomous Navigation. <https://ieeexplore.ieee.org/document/9196698>
19. Mellinger, D., et al. (2012). Trajectory Generation and Control for Precise Aggressive Maneuvers with Quadrotors. <https://journals.sagepub.com/doi/10.1177/0278364911434236>
20. Al-Farabi Kazakh National University (2023). AI-Driven UAV Swarm Simulation for Disaster Response in Urban Kazakhstan. <http://ecair.kz/2023-proceedings>

## ҰҰА БАСҚАРУ АЛГОРИТМДЕРІН СЫНАУ ҮШІН СИМУЛЯЦИЯЛЫҚ ПЛАТФОРМА ҚҰРУ

*Аңдатпа.* Ұшқышсыз ұшу аппараттары (ұшқышсыз ұшу аппараттары) географиялық әртүрлілігімен, экстремалды климаттық жағдайларымен және шалғай аудандардағы инфрақұрылымдық талаптармен сипатталатын Қазақстандағы аймаққа тән міндеттерді шешудің негізгі құралы ретінде пайда болды. Дегенмен, Қазақстанның бірегей операциялық орталарында ( $-40^{\circ}\text{C}$ -тан  $+45^{\circ}\text{C}$ -қа дейін), күтпеген желдің екпінімен (Алматы және Қостанай облыстарында  $15\text{--}20$  м/с) және таулы жерлерде GPS сигналының жиі нашарлауымен сипатталатын Қазақстандағы бірегей жұмыс орындарында ұшқышсыз ұшу аппараттарын орналастыру маңызды техникалық және логистикалық қиындықтар туғызады. Осы шарттарда UAV басқару алгоритмдерін физикалық сынақтан өткізу өте қымбат болып қана қоймайды, сонымен қатар қауіпсіздік ережелерімен, қоршаған ортаны болжау мүмкін еместігімен және операциялық аймақтардың ауқымдылығымен шектеледі. Осы кедергілерді шешу үшін бұл мақалада жоғары дәлдікпен елдің экологиялық және эксплуатациялық шындықтарына еліктеуге арналған Қазақстанға бағытталған UAV модельдеу платформасын әзірлеу ұсынылады.

Robot Operating System (ROS Noetic) және Gazebo 11 негізінде құрастырылған платформа үш жаңа құрамдас бөлікті біріктіреді: (1) Қазақстанның ауылшаруашылық және апатқа қарсы әрекет ету кезіндегі UAV орналастыруларынан алынған далалық деректердің көмегімен калибрленген физикаға негізделген UAV динамикасы, соның ішінде массасы ( $1,5$  кг), инерция тензоры және ротор профилі; (2) шаңнан туындаған LiDAR диапазонының қателері ( $\pm 0,15$  м) және температураға тәуелді IMU дрейфі ( $+40^{\circ}\text{C}$  кезінде  $0,2^{\circ}/\text{сағ}$ ) сияқты аймақтық жағдайларға бейімделген шу профилдері бар синтетикалық сенсор үлгілері (LiDAR, IMU, RGB камералары); және (3) Қазақстанның ұлттық метеорологиялық агенттігі Қазгидромет ұсынатын метеорологиялық деректер жиынтығынан алынған қоршаған ортаны бұзу модельдері, соның ішінде далалық жел динамикасы (екір  $18$  м/с) және ықтималдық GPS сигналының жоғалуы (Тянь-Шань тауларында оқуды тастап кету деңгейі  $25\text{--}35\%$ ).

Платформаның модульдік архитектурасы бейімделген басқару алгоритмдерін, соның ішінде желдің бұзылуынан бас тартуға арналған Болжалды басқару моделін (MPC), іздестіру-құтқару миссиялары үшін үйірді үйлестіру стратегияларын және нақты қазақстандық қиындықтарды көрсететін сценарийлер бойынша қателерге төзімділікті арттыруға (RL) негізделген жүйелерді тестілеуді қолдайды. Жағдайлық зерттеулер оның тиімділігін көрсетеді: имитациялық қатты жел сценарийлерінде ( $18$  м/с екпінді) орталықтандырылмаған үйірді үйлестіру алгоритмі Тянь-Шань тауларында қалыптасуды сақтауда миссияның  $88\%$  табысына қол жеткізді, ал адаптивті PID контроллері сенсор жағдайында траекторияны бақылау қателерін  $+40$  фут  $35\%$  азайтты. Түркістан облысында орналастырылған DJI Matrice 300 UAV далалық деректерімен кросс-валидация RMSE модельденген және нақты әлемдегі траектория

арасындағы 94% корреляцияны растады (0,12 м-ге қарсы 0,15 м), энергияны тұтыну болжамдары байқалған мәндерден 3%-дан аз ауытқыған.

**Түйін сөздер:** UAV симуляциясы, Gazebo-ROS интеграциясы, адаптивті басқару алгоритмдері, қазақстандық қоршаған ортаны модельдеу, үйір робототехникасы, сенсорлық эмуляция, цифрлық егіздер.

## СОЗДАНИЕ СИМУЛЯЦИОННОЙ ПЛАТФОРМЫ ДЛЯ ТЕСТИРОВАНИЯ АЛГОРИТМОВ УПРАВЛЕНИЯ БПЛА

**Аннотация.** Беспилотные летательные аппараты (БПЛА) стали ключевыми инструментами для решения региональных проблем в Казахстане, стране, характеризующейся огромным географическим разнообразием, экстремальными климатическими условиями и инфраструктурными требованиями в отдаленных районах. Однако развертывание БПЛА в уникальных операционных условиях Казахстана, отмеченных экстремальными температурами (-40 °C до +45 °C), непредсказуемыми порывами ветра (15–20 м/с в Алматинской и Костанайской областях) и частым ухудшением сигнала GPS в горной местности, создает значительные технические и логистические проблемы. Физическое тестирование алгоритмов управления БПЛА в этих условиях не только непомерно дорого, но и ограничено правилами безопасности, непредсказуемостью окружающей среды и огромным масштабом операционных зон. Для устранения этих барьеров в данной статье предлагается разработка казахстанской платформы моделирования БПЛА, предназначенной для имитации экологических и операционных реалий страны с высокой точностью. Платформа, созданная на основе операционной системы робота (ROS Noetic) и Gazebo 11, объединяет три новых компонента: (1) физическая динамика БПЛА, откалиброванная с использованием полевых данных, полученных от казахстанских сельскохозяйственных и спасательных БПЛА, включая массу (1,5 кг), тензор инерции и профили тяги ротора; (2) синтетические модели датчиков (LiDAR, IMU, RGB-камеры) с профилями шума, адаптированными к региональным условиям, таким как погрешности дальности LiDAR, вызванные пылью ( $\pm 0,15$  м) и дрейф IMU в зависимости от температуры (0,2°/час при +40°C); и (3) модели возмущений окружающей среды, полученные из метеорологических наборов данных, предоставленных Казгидрометом, национальным метеорологическим агентством Казахстана, включая динамику степного ветра (порывы до 18 м/с) и вероятностную потерю сигнала GPS (коэффициенты потери 25–35% в горах Тянь-Шаня). Модульная архитектура платформы поддерживает тестирование алгоритмов адаптивного управления, включая Model Predictive Control (MPC) для подавления возмущений ветра, стратегии координации роя для поисково-спасательных миссий и системы отказоустойчивости на основе обучения с подкреплением (RL), в сценариях, отражающих реальные проблемы Казахстана. Практические примеры демонстрируют его эффективность: в моделируемых сценариях сильного ветра (порывы 18 м/с) децентрализованный алгоритм координации роя достиг 88% успеха миссии по поддержанию формации над горами Тянь-Шаня, в то время как адаптивный ПИД-регулятор уменьшил ошибки отслеживания траектории на 35% в условиях дрейфа датчика +40 °C. Перекрестная проверка с полевыми данными с БПЛА DJI Matrice 300, развернутого в Туркестанском регионе, подтвердила 94% корреляцию между моделируемой и реальной траекторией RMSE (0,12 м против 0,15 м), при этом прогнозы потребления энергии отклонялись менее чем на 3% от наблюдаемых значений.

**Ключевые слова:** Моделирование БПЛА, интеграция Gazebo-ROS, алгоритмы адаптивного управления, моделирование окружающей среды Казахстана, роевая робототехника, эмуляция датчиков, цифровые двойники.

**Information about the authors**

Akhmet Arys	Akhmet Arys, fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan E-mail: <a href="mailto:arysakhmet@gmail.com">arysakhmet@gmail.com</a>
Gusmanov Altair	fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan E-mail: <a href="mailto:altair.gusmanov@mail.ru">altair.gusmanov@mail.ru</a>
Akramkhanov Alikhan	fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan E-mail: <a href="mailto:aakramkhanov@bk.ru">aakramkhanov@bk.ru</a>
Akimbay Shyryn	2nd year PhD student of Automation and IoT specialty at Al-Farabi Kazakh National University, Almaty, Kazakhstan E-mail: <a href="mailto:akimbay.sh@gmail.com">akimbay.sh@gmail.com</a>

**Авторлар туралы мәлімет**

Арыс Ахмет	Халықаралық Ақпараттық Технологиялар Университетінің ақпараттық жүйелер мамандығының 4-курс студенті, Алматы қ., Қазақстан E-mail: <a href="mailto:arysakhmet@gmail.com">arysakhmet@gmail.com</a>
Гусманов Альтаир	Халықаралық Ақпараттық Технологиялар Университетінің ақпараттық жүйелер мамандығының 4-курс студенті, Алматы қ., Қазақстан E-mail: <a href="mailto:altair.gusmanov@mail.ru">altair.gusmanov@mail.ru</a>
Акрамханов Алихан	Халықаралық Ақпараттық Технологиялар Университетінің ақпараттық жүйелер мамандығының 4-курс студенті, Алматы қ., Қазақстан E-mail: <a href="mailto:aakramkhanov@bk.ru">aakramkhanov@bk.ru</a>
Әкімбай Шырын	Әл-Фараби атындағы Қазақ Ұлттық университетінің Автоматтандыру және IoT мамандығының 2-курс докторанты, Алматы қ., Қазақстан E-mail: <a href="mailto:akimbay.sh@gmail.com">akimbay.sh@gmail.com</a>

**Сведение об авторах**

Арыс Ахмет	студент четвертого курса по специальности информационных систем Международного Университета Информационных технологий, г. Алматы, Казахстан E-mail: <a href="mailto:arysakhmet@gmail.com">arysakhmet@gmail.com</a>
Гусманов Альтаир	студент четвертого курса по специальности информационных систем Международного Университета Информационных технологий г. Алматы, Казахстан E-mail: <a href="mailto:altair.gusmanov@mail.ru">altair.gusmanov@mail.ru</a>
Акрамханов Алихан	студент четвертого курса по специальности информационных систем Международного Университета Информационных технологий, г. Алматы, Казахстан E-mail: <a href="mailto:aakramkhanov@bk.ru">aakramkhanov@bk.ru</a>
Акимбай Шырын	докторант 2 курса специальности «Автоматизация и IoT» КазНУ им. Аль-Фараби, Алматы, Казахстан E-mail: <a href="mailto:akimbay.sh@gmail.com">akimbay.sh@gmail.com</a>

UDC 629.7.08

IRSTI 55.47.81

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_4](https://doi.org/10.53364/24138614_2025_36_1_4)**Kh. I. Abdulayev<sup>1</sup>, S.Q. Allakhverdiyeva<sup>1\*</sup>**<sup>1</sup>National Aviation Academy, Baku, Azerbaijan*E-mail: sallahverdiyeva@naa.edu.az\****DEFECTS OF AIRCRAFT STRUCTURAL ELEMENTS BASED ON POLYMERIC-COMPOSITE AND THEIR DIAGNOSTICS BY NONDESTRUCTIVE CONTROL METHODS**

**Abstract.** *This article examines the application of non-destructive testing (NDT) methods for assessing the structural integrity of aircraft components made from polymer-based composite materials. Composite materials are widely used in aviation due to their high strength-to-weight ratio, corrosion resistance, and durability. However, these materials are subject to various defects caused by manufacturing processes, operational loads, and environmental factors such as temperature and humidity fluctuations. Traditional NDT methods, including ultrasonic, radiographic, optical-visual, capillary, and thermal testing, each have specific advantages and limitations. Ultrasonic testing, for example, does not provide comprehensive volumetric analysis, while radiographic methods require complex safety measures. Optical-visual techniques fail to detect internal defects, and capillary methods suffer from low productivity. To address these challenges, the study proposes improvements to existing diagnostic techniques, the development of new automated models, and the optimization of parametric indicators. The research also explores an integrated "human-machine-environment" system to enhance the reliability of defect detection and assessment. Advancements in NDT technologies will not only increase the accuracy and efficiency of aircraft inspections but also improve safety, extend service life, and reduce maintenance costs. The findings of this study contribute to the development of modern diagnostic complexes that ensure higher operational reliability of aircraft structures.*

**Keywords:** *composite materials, defect, non-destructive testing method, loaded parts, diagnostic models, dynamic correlation.*

**Introduction.**

Aircraft have a complex structure. During the long-term operation of aircraft mechanisms and components, defects that occur during their service life lead to an increase in accidents. Ensuring safety is the responsibility of modern technical diagnostics, which are closely related to automatic monitoring systems. Technical diagnostics is the science of determining the state of a technical system and solving a wide range of problems such as obtaining and evaluating diagnostic information. Currently, aircraft monitoring systems are limited in scope, making it impossible to predict the technical state of all aircraft elements and assemblies over a long period of time during their dynamic motion. Recently, one of the most pressing issues is the creation of combined systems for diagnostics and forecasting of the technical condition of aircraft. Such combined systems make it possible to timely monitor defects in polymer composite-based structural elements, assemblies and nodes, accurately obtain information through modern sensors installed in the composition of composite materials, study their dynamic and static characteristics and process them on board, automatically compare information within permissible limits and transmit it to a ground station, and also control various physical parameters. Namely, the creation of modern automated monitoring systems in this direction, as well as the complex design of the "human - machine - environment" system are considered one of the most pressing issues [1].



The purpose of the work is to conduct a comparative analysis of diagnostic methods (models) for the control of polymer composite-based structural elements of the HG and to identify directions for improving existing diagnostic methods.

### **Materials and methods of the study.**

In aviation technology, the replacement of structural parts of aircraft with high-modulus polymer composite materials (PKM) is considered one of the important and strategic issues. Composite materials are considered multicomponent materials. Typically, a composite material consists of a plastic base (matrix) and various reinforcing fillers. According to the structure of the filler, composite materials (KM) are divided into fibrous, layered and dispersed solid materials. The matrix of the composite material ensures the monolithicity of the material, and the filler provides stress transfer and distribution.

Depending on the composition and properties of the matrix and filler, their proportions and the choice of placement, materials are divided into operational and technological types. Many composite materials surpass traditional materials and alloys in terms of their mechanical properties (specific strength, hardness, wear resistance, fatigue limit, thermal and vibration resistance, sound absorption, impact viscosity, etc.). The impact strength and elastic modulus of composites are 2...5 times higher than those of conventional structural alloys. Studies show that during the operation of structures made of composite materials, mechanical, impact, corrosion, etc. type damage leads to the formation of cracks, and the effect of periodic fluctuations in temperature and humidity leads to an increase in their size. Thus, diagnostics of structures and their periodic repair prevent breakdowns and increase operational safety.

Currently, a number of non-destructive inspection methods have been developed in scientific research centers in many countries of the world for the control of materials and products, where the development of damage from the initial stage to complete destruction (wearing out) of materials is monitored. Defects formed in materials, the study of their informative parameters, as well as the calculation of defects in structures fall under the responsibility of non-destructive inspection methods. However, non-destructive inspection methods used in the control of composite materials have shortcomings. Namely, in order to eliminate these shortcomings, research is being conducted in the leading scientific research centers in the field of non-destructive inspection to improve the most effective diagnostic methods and develop new methods. It is known that composite materials are widely used in aviation, rocket, automotive, mechanical engineering, metallurgy, etc. industries. Over the past decade, the use of composite materials in aircraft structures has increased significantly, which is 50-55% in Boeing 787, Airbus A350 aircraft, 80-85% in small aircraft, and 100% in unmanned aerial vehicles. Figure 1 shows the replacement of the entire fuselage section of a Boeing 787 aircraft with aerospace composite materials, and Figure 2 shows the replacement of individual parts of an A-380 aircraft with aerospace composite materials.



Fig. 4).

Unlike metallic materials, composite material defects can increase during operation, which reduces the reliability of the structure. Here, liquid penetrates into the composite through a crack on the matrix surface, damaging it and reducing its elastic modulus. Even with a typical delamination defect, the bond strength weakens, fibers break, and tensile strength decreases.

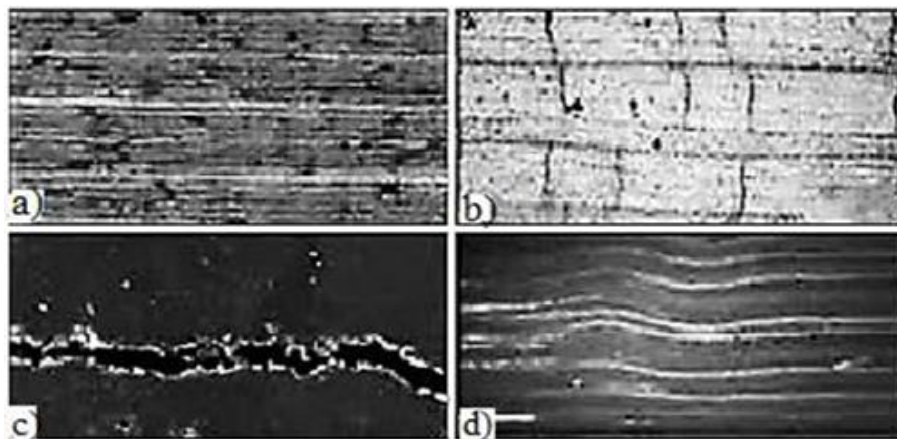


Figure 3 – Defects in composite materials: a) pores (black); b) cracks; c) delamination; d) wave-like defects [3]

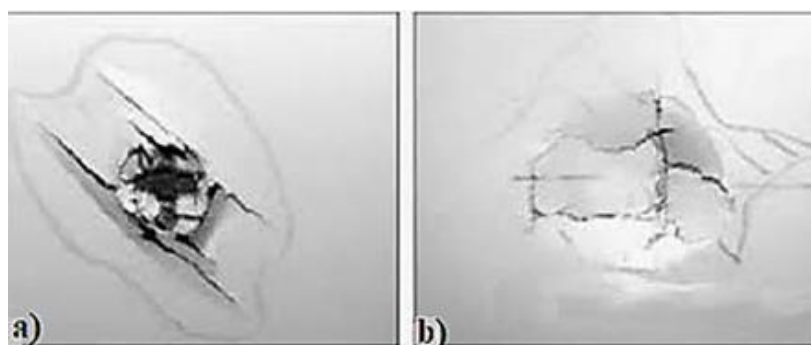


Figure 4 – External damage of composite materials: a) on the surface of the wing; b) in the fuselage [3]

Composite-based objects are very difficult to control because they have a heterogeneous structure, anisotropic properties, many different types of reinforcement (unidirectional, transverse-longitudinal, combined, etc.), specific physical properties (high electrical insulation quality, low thermal conductivity, sound absorption, many physical and mechanical characteristics, low density ( $0.02- q/sm^3$ )). It is very difficult to keep the situation under control. Therefore, testing composite materials by physical methods compared to homogeneous structures (metals) is characterized by low sensitivity, transmittance and signal-to-noise ratio.

Depending on the nature of the filler used in many types of composite materials, they are usually dielectric or poorly conductive. Practically all composite materials are non-magnetic materials, which is why many non-destructive testing methods used for metals are considered unsuitable for composite-based products. The main non-destructive testing methods used for composite materials are given in Table 1 [1,3]. Let us consider the main characteristics of the methods used. Testing composite materials by high-frequency ultrasonic methods is considered ineffective, since the power of ultrasonic waves with a frequency of 1...5 MHz is scattered in the matrix, fibers and other particles and weakens. Thus, the control becomes difficult depending on

the thickness range of composite materials. Some types of composite materials are very sensitive to moisture, and non-destructive testing methods are not applicable here.

In the nondestructive ultrasonic method, a powerful laser excites waves, making it essential for personnel to handle the technology with care and work under controlled conditions for a limited time. Among the various acoustic nondestructive testing methods, the low-frequency approach is particularly effective for riveted joints in multilayer polymer composite materials. This method enables the inspection of multilayer riveted construction joints composed of carbon, boron, and organic glass-type fibers, including non-metallic coatings and riveted joints between frames.

A key distinction of the low-frequency nondestructive testing method for multilayer riveted constructions based on polymer composites lies in its unique capabilities. It allows for the detection of structural non-dampness in gyroscopic materials, ensuring the integrity of details and assemblies. Additionally, it facilitates the assessment of surface curvatures, accommodating both concave and convex components. Furthermore, this method proves effective in evaluating uneven surfaces of anisotropic materials, enhancing the reliability of structural inspections.

Table 1 – Basic methods of non-destructive testing used for composite materials [1,3]

Area of application	Control methods used for KM	Information parameter	Control production n, m <sup>2</sup> /hour	Sizes of detected defects, mm	
				Minimum scan	Minimum length
Damage to smoothness	Active thermal control	Temperature, heat area	0.6	0.15	10
Crack (open defect)	Surface wave ultrasound method	The time duration of the ultrasonic signal between the receiver and transmitter of the ultrasonic transducer	2.1	0.1	15
Revealed defects	Ultrasonic sound and thermography method	Acoustic radiation and thermal area	0.6	0.001	5...7
Various defects and cracks	Radiography Radioscopy Radiometry	Conversion of X-ray intensity into optical beam	4...5	0.15	3
Cracks	Radio wave	Radiation characteristics	0.5	0.1	15
Foreign particles, roughness	X-ray television	Comparison of X-ray and video images	7	0.2	0.2
The unevenness of multi-layer construction in rotating bodies	Tangential lighting	Size change and depth detection <sup>3</sup>		0.2	10

In non-gyroscopic composite materials, the acoustic contact between the transmitter and the object is established through a viscous liquid. Typically, the free oscillation method is employed for testing composite materials, analyzing the characteristics of free oscillations that occur after a vibration shock. This method, in combination with appropriate electronic equipment, enables the detection of various defects, including layered and air-bubble cavities, as well as non-adherent parts. Compared to other low-frequency acoustic methods such as impedance, velocimetric, and acoustic-topographic techniques, the free oscillation method offers distinct advantages. It allows for the detection of defects located at significant depths, facilitates the control of structural

materials with a low elastic modulus, and ensures effective monitoring of materials characterized by a high damping coefficient of elastic oscillations.

One of the most common methods is the radiation control method [3]. This method allows the detection of various voids, individual particulate foreign materials, as well as non-uniform structures (fractures of reinforcing elements, voids, etc.). It has been established that, as in metals, if the illumination has a weak effect on defects (cracks perpendicular to the radiation flux, delamination), depending on the change in the total thickness of composite materials, then the application of the X-ray method becomes impossible.

For this type of materials, the capillary-radiation inspection method is used, in which the contrast of the elements of the illuminated object is increased by the penetration of a liquid X-ray contrast penetrant. For this purpose, a wide range of organic and inorganic penetrants with a high absorption coefficient of X-rays are produced. One of the most common penetrants is considered to be a solution of ZnI<sub>2</sub> (zinc iodide). It is opaque and has high radiation compared to other substances. With this method, it is possible to detect surface or intermediate damage as a result of the penetration of capillary penetrants into cracks or interlayer gaps. The high achievements of radiation defectoscopy in the last decade have also had an impact on the development of digital X-ray television inspection.

Here, the X-ray beam is converted into an optical beam by passing through the object with the help of optical converters and is reflected on the monitor screen by the telephoto method. Here, inexpensive X-ray optical scintillation crystals are used and the optical image is read by a small-sized charge-coupled device (CCD) with high sensitivity (Fig. 5). The digital radioscopy method provides instant digital X-ray imaging, improves visual imaging, and determines the measurement of geometric parameters controlled by the images. Here, a database (flash radiography) provides for the input, storage, selection, display of images, and the extraction of results, as well as data transmission to local and external networks [3].

Figure 6a shows low-contrast images of a fiber with a width of a few pixels inside a composite material using the digital X-ray method. In this case, it is difficult to examine the structure of the fiber. Here, specialized digital filters are used for image processing, which improves the X-ray image of the fibers in given directions in the composite material (Figures 6b, c and d).

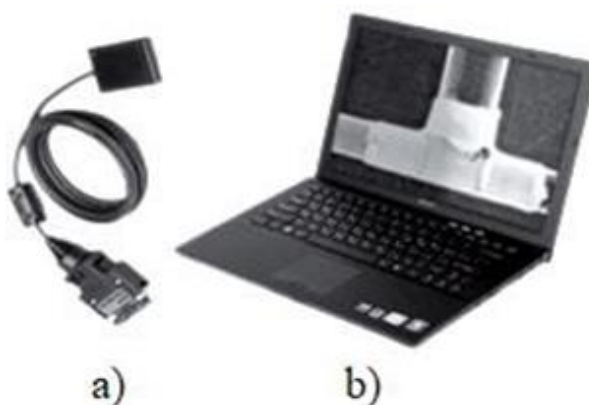


Figure 5 – X-ray flash radiography method of HCC matrix: a) YCC sensor; b) X-ray image of a trident made of composite material [3]



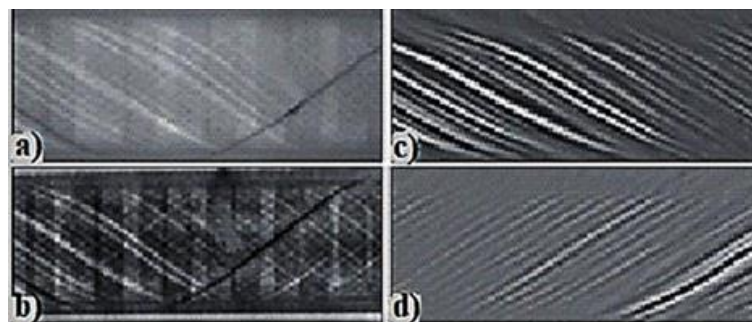


Figure 6 – X-ray image of a composite sample: a) real; b), c), d) images after digital contrasting

Figure 7 shows images of carbon fibers and their fractures. Here, the effective use of computed tomography depends on the geometric dimensions of the object. Such an expensive technology can be considered unique for the study of volumetric internal structures. Its resolution is several microns and is determined by computed microtomography.



Figure 7 – Visualization of cracks in the internal structure of composite material using tomography [5]

It is difficult to diagnose large-sized composite materials using the tomography method, when 360...720-pixel images are rotated at an angle of 360 degrees. This method makes it possible to obtain 6, 12, 24 projections with small angles. However, exposure deficiencies are encountered here. In this case, the error in the reconstruction of geometric parameters increases sharply.

Thermography (thermal) method is the most promising non-destructive inspection method for composite material-type structures [2, 6]. The basis of this method is infrared technology (thermal imager or pyrometer) and determines the presence of a defect as a result of changes in temperature fields. Thermography records temperature changes at 0.01°C. Thus, during mechanical loads, energy concentration sites are recorded, where energy is released as a result of plastic deformation. By recording temperature fields on the surface, it is possible to determine the location of the energy concentrator relative to the surface of the product. Here, taking into account the issues of non-stationary heat transfer, it is also possible to determine its dimensions and location. Figure 8 and Figure 9 show images of the distribution of heat fields (thermograms) in composite-based objects [6].

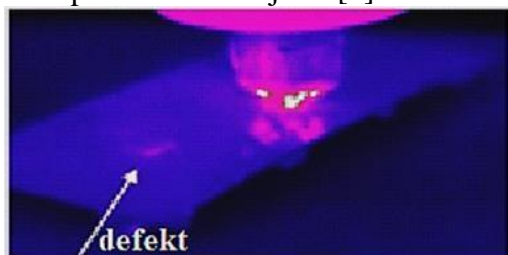


Figure 8 – Thermogram of the defect of the composite panel of the aircraft [2]

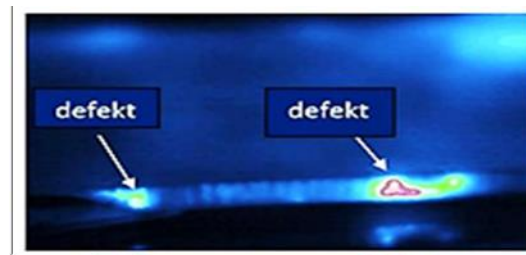


Figure 9 – Thermogram of the defect found on the honeycomb panel of the wing of the Super Jet-100 aircraft [2]

The main advantages of thermographic control are universality, high evaluation of temperature signals, efficiency, high productivity and remote measurement. The disadvantages of this method are the presence of specific noise generated by the object (fluctuations in optical properties), as well as by an external heat source.

It has been established that the active thermography method is one of the new, highly effective and informative methods of non-destructive control and is widely used in the aerospace industry. Due to a number of advantages of the active thermal control method, it is more expedient to inspect composite materials using this method (Fig. 10 and Fig. 11).

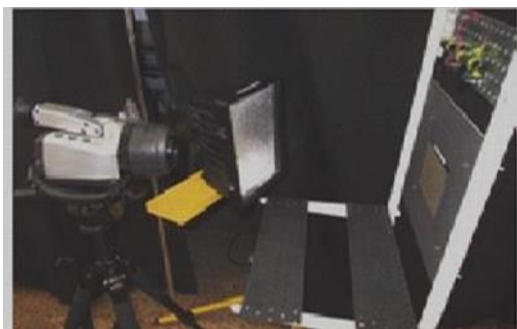


Figure 10 – Active temperature control method applied to composite materials [2]



Figure 11 – The device of the active temperature method for the study of airplanes (Automation Technology, Germany) [2]

However, recently, automated diagnostic monitoring systems of leading organizations such as Boeing, Airbus, AgustaWestland (AW), Northrop Grumman, NASA have been designed based on modern sensors. The application of such sensors has been widely used in conducting scientific research and measuring many different physical parameters (deformation, temperature and magnetic field). Studies show that further improvement of modern sensors based on the fiber-optic method in the future will give impetus to the development of a number of areas. For example, modern sensors will allow for full monitoring of the internal structure of composite material parts, rivet joints, and airframes, as well as for operational monitoring of the resources of individual elements of the structure and uneven areas in a short time. Figure 12 and Figure 13 show the installation of modern sensors in the internal structure of composite materials [5].

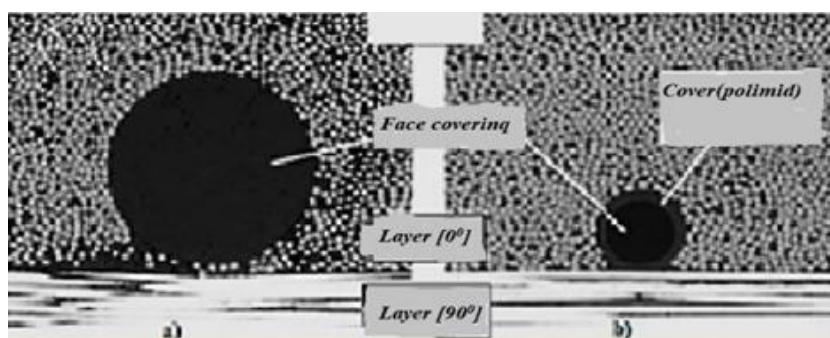


Figure 12 – Placement of conventional optical fiber (a) and small diameter fiber (b) in the internal structure of carbon fiber [4,5]

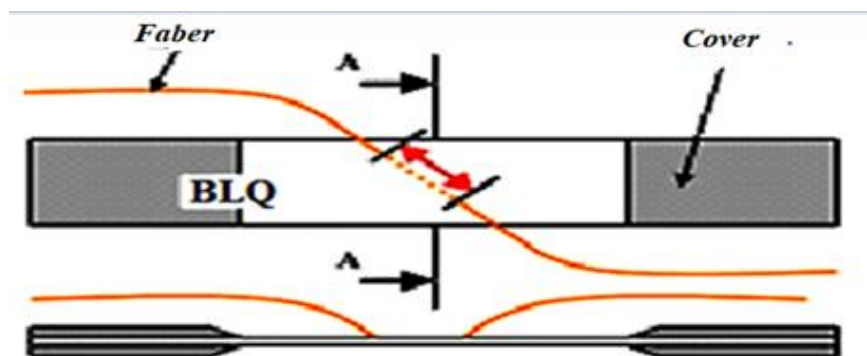


Figure 13 – The sensor installed on PCM [4,5]

### Conclusion.

Thus, the above methods of non-destructive testing have their advantages and disadvantages for the diagnosis of composite materials. The ultrasonic method does not allow volume control of composite materials, as it is necessary to take into account the requirements for the material of the controlled object, as well as its surface. The radiation method limits the complexity of the technical execution, used to ensure the requirements of radiation safety, as well as specific defects of composite materials. The disadvantage of the optical-visual method is that hidden defects are not detected, and the disadvantage of the capillary method is its contact and low productivity. The disadvantage of the heating method is a one-sided check, dependence of surface temperature signals on the depth of defects, as well as a high level of detection of additive and multiplicative sounds when using an optical stimulator. It also makes it difficult to control the quality of composite materials.

And with this connection, the improvement of diagnostic methods in recent years, the creation of new automated diagnostic models using non-destructive control methods, the study of new methods and tools, as well as the optimization of parametric indicators are considered as one of the most important issues of creating new modern diagnostic complexes.

The creation of such control systems will not only increase the safety of aircraft operation, increase its resource, increase the accuracy of measurements, make it economically efficient and affordable, but also open the way for a complex and convenient study of the "human - machine - environment" system.

### References

1. Bajenov, S. L., Berlin, A. A., & Kulkov, A. A. (2009). *Polimernie kompozisionnie materialy*. Moscow, Russia: Intellekt,
2. Yeşanov, V.A. (2016). "Razrabotka metoda aktivnoqo teplovoqo kontrolya udarnix povrejdeniy v aviasionnix kompozitax," in *Proc. VI All-Russian Sci.-Pract. Conf. "Nerazruşayuşiy kontrol: elektronnoe priborostroenie, texnologii, bezopasnost"*, Tomsk, Russia, May 23–27, vol. 1, pp. [5].
3. Boysov, B.V., Vasilyev, S. L., Qromashev, A. Q. & Yurgensov, S. A. (2011). "Metodi nerazrushayuşeqo kontrolya, primenyaemie dlya konstruksii iz PKM," *Trudi MAI*, vol. 49, pp. 1–11.
4. Nobuo Takeda & Yoji Okabe (2011). "Fiber Bragg Grating Sensors in Aeronautics and Astronautics," in *Fiber Bragg Grating Sensors: Recent Advancements, Industrial Applications and Market Exploitation*, 2011, pp. 171–184. Doi:10.2174/978160805084011101010171
5. "Umnyy material dlya T-50: lazernie texnologiya obezopasat kompozitnie konsrtuksii," NIS «Institut razvitiya issledovaniy, razrabotok i transferta texnologiy», Available: <https://uacrussia.livejournal.com/81763.html>.
6. Xorev, V. S. (2013). *Ultrazvukovaya infrakrasnaya termoqrafiya polimernix kompozisionnix materialov*, *Ph.D. dissertation*, Tomsk Polytechnic Univ., Tomsk, Russia.



## ПОЛИМЕРЛІ-КОМПОЗИТТІК НЕГІЗДЕГІ ӘУЕ КЕМЕЛЕРІНІҢ КОНСТРУКТИВТІК ЭЛЕМЕНТТЕРІНІҢ АҚАУЛАРЫ ЖӘНЕ ОЛАРДЫ БҰЗБАЙТЫН БАҚЫЛАУ ӘДІСТЕРІМЕН ДИАГНОСТИКАЛАУ

*Аңдатпа.* Жобалау үшін бұзылмайтын сынау әдістерінің салыстырмалы талдауы жүргізілді. Бұл мақалада полимер негізіндегі композициялық материалдардан жасалған ұшақ компоненттерінің құрылымдық тұтастығын бағалау үшін бұзылмайтын бақылау (ББ) әдістерін пайдалану қарастырылады. Композиттік материалдар жоғары беріктік пен салмақ қатынасына, коррозияға төзімділігіне және беріктігіне байланысты авиацияда кеңінен қолданылады. Дегенмен, бұл материалдар өндірістік процестерден, жұмыс кернеуінен және температура мен ылғалдылықтағы өзгерістер сияқты қоршаған орта факторларынан туындаған әртүрлі ақауларға сезімтал. Дәстүрлі бұзылмайтын бақылау әдістері, соның ішінде ультрадыбыстық, радиографиялық, оптикалық-визуалды, капиллярлық және термиялық, олардың артықшылықтары мен шектеулері бар. Ультрадыбыстық тексеру, мысалы, кешенді көлемдік талдауды қамтамасыз етпейді, ал радиографиялық әдістер күрделі қауіпсіздік шараларын талап етеді. Оптикалық-визуалды әдістер ішкі ақауларды анықтай алмайды, ал капиллярлық әдістер төмен өнімділіктен зардап шегеді. Осы мәселелерді шешу үшін зерттеу қолданыстағы диагностикалық әдістерді жетілдіруді, жаңа автоматтандырылған модельдерді әзірлеуді және параметрлік көрсеткіштерді оңтайландыруды ұсынады. Зерттеу сонымен қатар ақауларды анықтау және бағалау сенімділігін арттыру үшін адам-машина-қоршаған ортаның біріктірілген жүйесін қарастырады. Бұзбайтын сынақ технологияларын дамыту әуе кемелерін тексерудің дәлдігі мен тиімділігін арттырып қана қоймайды, сонымен қатар қауіпсіздікті жақсартады, қызмет ету мерзімін ұзартады және техникалық қызмет көрсету шығындарын азайтады. Бұл зерттеудің нәтижелері әуе кемелері құрылымдарының пайдалану сенімділігін арттыратын заманауи диагностикалық жүйелерді дамытуға ықпал етеді.

**Түйін сөздер:** композициялық материалдар, ақаулық, бұзылмайтын бақылау әдістері, жүктелген бөлшектер, диагностикалық модельдер, динамикалық корреляция.

## ДЕФЕКТЫ КОНСТРУКТИВНЫХ ЭЛЕМЕНТОВ ВОЗДУШНЫХ СУДОВ НА ПОЛИМЕРНО- КОМПОЗИТНОЙ ОСНОВЕ И ИХ ДИАГНОСТИКА МЕТОДАМИ НЕРАЗРУШАЮЩЕГО КОНТРОЛЯ

*Аннотация:* Проведен сравнительный анализ методов неразрушающего контроля для конструкции. В данной статье рассматривается применение методов неразрушающего контроля (НК) для оценки структурной целостности компонентов летательных аппаратов, изготовленных из композиционных материалов на основе полимеров. Композитные материалы широко используются в авиации благодаря высокому соотношению прочности и веса, коррозионной стойкости и долговечности. Однако эти материалы подвержены различным дефектам, вызванным производственными процессами, эксплуатационными нагрузками и факторами окружающей среды, такими как перепады температуры и влажности. Традиционные методы неразрушающего контроля, включая ультразвуковой, радиографический, оптико-визуальный, капиллярный и тепловой, имеют свои преимущества и ограничения. Ультразвуковой контроль, например, не обеспечивает всестороннего объемного анализа, а радиографические методы требуют сложных мер безопасности. Оптико-визуальные методы не позволяют обнаружить внутренние дефекты, а капиллярные методы страдают от низкой производительности. Для решения этих проблем в исследовании предлагается усовершенствовать

существующие методы диагностики, разработать новые автоматизированные модели и оптимизировать параметрические показатели. В исследовании также рассматривается интегрированная система «человек-машина-среда» для повышения надежности обнаружения и оценки дефектов. Развитие технологий неразрушающего контроля позволит не только повысить точность и эффективность проверок воздушных судов, но и улучшить безопасность, продлить срок службы и снизить затраты на обслуживание. Результаты данного исследования способствуют разработке современных диагностических комплексов, обеспечивающих повышение эксплуатационной надежности авиационных конструкций.

**Ключевые слова:** композитные материалы, дефект, методы неразрушающего контроля, нагруженные части, диагностические модели, динамическая корреляция.

#### Information about the authors

<b>Abdullayev Khagani Imran</b>	Phd, professor of the National Aviation Academy, Baku, The Republic of Azerbaijan, E-mail: khagani61@gmail.com
<b>Allakhverdiyeva Solmaz Qulu</b>	Senior lecturer, National Aviation Academy, Baku, The Republic of Azerbaijan, E-mail: <a href="mailto:sallahverdiyeva@naa.edu.az">sallahverdiyeva@naa.edu.az</a>

#### Авторлар туралы мәлімет

<b>Абдуллаев Хагани Иман</b>	PhD, Ұлттық авиация академиясының профессоры, Баку қ., Әзірбайжан Республикасы, E-mail: khagani61@gmail.com
<b>Аллахвердиева Солмаз Гулу</b>	Аға оқытушы авиация академиясының оқытушысы, Баку қ., Әзірбайжан Республикасы, E-mail: <a href="mailto:sallahverdiyeva@naa.edu.az">sallahverdiyeva@naa.edu.az</a>

#### Сведение об авторах

<b>Абдуллаев Хагани Иман оглы</b>	доктор наук, профессор Национальной авиационной академии, г. Баку, Азербайджанская Республика, E-mail: khagani61@gmail.com
<b>Аллахвердиева Солмаз Гулу</b>	Старший преподаватель Национальной Авиационной Академии, г.Баку, Азербайджанская Республика, E-mail: <a href="mailto:sallahverdiyeva@naa.edu.az">sallahverdiyeva@naa.edu.az</a>

=====

**ЛОГИСТИКА, ТАСЫМАЛДАУДЫ ҰЙЫМДАСТЫРУ, КӨЛІКТЕГІ ҚАУІПСІЗДІК  
ЛОГИСТИКА, ОРГАНИЗАЦИЯ ПЕРЕВОЗОК, БЕЗОПАСНОСТЬ НА ТРАНСПОРТЕ  
LOGISTICS, TRANSPORTATION ORGANIZATION, TRANSPORT SECURITY**

=====

УДК 656.7

МРНТИ 73.37

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_5](https://doi.org/10.53364/24138614_2025_36_1_5)

**А.Ж. Абжапбарова<sup>1\*</sup>, О.В. Гармаш<sup>1</sup>**

<sup>1</sup>Академия Гражданской Авиации, Алматы, Казахстан

E-mail: [Ainur.abzhapbarova@mail.ru](mailto:Ainur.abzhapbarova@mail.ru)\*

**ПРОБЛЕМЫ РЕГУЛЯРНОСТИ ВОЗДУШНЫХ ПЕРЕВОЗОК КАЗАХСТАНА**

*Аннотация.* Проблема задержек рейсов является одной из ключевых в сфере воздушных перевозок Казахстана, оказывая значительное влияние на пассажирский сервис, операционные процессы авиакомпаний и экономику аэропортов. Учитывая рост пассажиропотока и активное развитие авиарынка, выявление и устранение причин задержек становится важной задачей для повышения конкурентоспособности отрасли.

Предметом исследования являются причины задержек рейсов и их влияние на эффективность работы авиакомпаний и аэропортов Казахстана. В ходе исследования были решены такие задачи как анализ текущего состояния пунктуальности рейсов в Казахстане, выявление основных факторов, влияющих на задержки, оценка экономических и социальных последствий нарушений графика полетов, разработка рекомендаций по снижению задержек и повышению пунктуальности.

Целью исследования было изучение причин задержек рейсов на рынке воздушных перевозок Казахстана и разработка эффективных мер по их минимизации.

В работе использованы методы статистического анализа, экспертных опросов, анализа нормативных документов, а также сравнительный анализ показателей пунктуальности ведущих авиаперевозчиков.

Основными результатами исследования стало определение ключевых причин задержек рейсов, включая погодные условия, технические неисправности, загруженность аэропортов и организационные факторы, анализ статистики задержек рейсов крупнейших авиакомпаний Казахстана, разработка предложений по оптимизации процессов планирования и управления воздушными перевозками.

В результате исследования предложены меры по сокращению задержек рейсов, включая улучшение системы оперативного планирования, внедрение цифровых технологий для прогнозирования задержек, модернизацию аэропортовой инфраструктуры и совершенствование нормативной базы. Реализация данных рекомендаций позволит повысить пунктуальность авиаперевозок и улучшить качество обслуживания пассажиров в Казахстане.

**Ключевые слова:** задержка рейсов, авиатранспортный рынок, регулярность полетов, аэропорты, авиакомпании, качество обслуживания, безопасность полетов.

**Введение.**

Задержки авиарейсов представляют серьезную проблему для авиационной отрасли, влияя на экономику, безопасность и комфорт пассажиров. В осенне-зимний период этот

вопрос становится особенно актуальным из-за неблагоприятных погодных условий, увеличивающих время подготовки самолетов и потребность в противообледенительной обработке. Это приводит к росту затрат авиакомпаний, перегрузке аэропортов и неудобствам для пассажиров, сталкивающихся с изменением планов и финансовыми потерями.

Исследование задержек рейсов, особенно с учетом сезонных факторов, является актуальной задачей, решение которой поможет улучшить авиационные перевозки, снизить финансовые потери и повысить удовлетворенность пассажиров.

Целью исследования является анализ и прогнозирование задержек авиарейсов с учетом сезонных факторов (осенне-зимний период) для минимизации их влияния на работу авиакомпаний и аэропортов.

Для достижения этой цели проведено исследование, в рамках которого были проанализированы факторы задержек авиарейсов, собраны и обработаны соответствующие данные. Применение математических и статистических методов позволило выявить закономерности и разработать прогнозные модели. На основе полученных результатов сформулированы рекомендации для авиакомпаний и аэропортов, направленные на повышение пунктуальности рейсов и улучшение качества обслуживания пассажиров.

В настоящее время на рынке авиатранспортных услуг Казахстана наблюдается высокая частота задержек рейсов и неудовлетворительное качество обслуживания со стороны авиакомпаний. Согласно данным Комитета гражданской авиации РК, в период с января по июнь 2024 года в Казахстане было задержано 4 842 рейса из 28 235, в то время как за весь 2023 год задержано 6 735 рейсов из 50 701. Одновременно наблюдается рост активности населения в защите своих прав как потребителей авиационных услуг. Проблема регулярности рейсов и минимизации задержек актуальна не только для Казахстана, но и для мировой авиационной отрасли. Авиакомпании по всему миру исследуют этот вопрос и предпринимают различные меры для минимизации затрат и потерь, связанных с устранением последствий нарушенной регулярности полетов, чтобы сохранить свою клиентскую базу и репутацию [1].

Воздушные перевозки играют ключевую роль в транспортной системе Казахстана, обеспечивая быструю и эффективную связь между регионами страны и международными направлениями. Однако одной из наиболее острых проблем авиационной отрасли остается высокая частота задержек рейсов, что негативно сказывается на пассажирском опыте, операционной деятельности авиакомпаний и экономике аэропортов.

Задержки рейсов могут быть вызваны множеством факторов, включая неблагоприятные погодные условия, технические неисправности воздушных судов, загруженность аэропортов, недостаточную координацию наземных служб и несвоевременное выполнение регламентных процедур. В последние годы авиационная индустрия Казахстана предпринимает шаги по повышению пунктуальности, однако проблема остается актуальной, особенно в крупных узловых аэропортах [2].

В данной статье рассматриваются основные причины задержек рейсов, их влияние на участников авиационного рынка, а также предлагаются методы оптимизации процессов планирования и управления воздушными перевозками.

### **Материалы и методы исследования.**

В зарубежной практике, на фоне свободного и дерегулированного рынка, регулярность рейсов играет важную роль для авиакомпаний, помогая привлечь высокодоходных пассажиров, сократить затраты, связанные с задержками. Ассоциации перевозчиков, такие как Ассоциация европейских авиакомпаний, регулярно оценивают деятельность своих членов с целью стимулирования их развития и улучшения коммерческой практики, а также совершенствования работы аэропортов и поставщиков

аэронавигационных услуг. Главной задачей этих ассоциаций является внедрение культуры безопасности полетов, проведение экспертиз и исследований для обмена передовым опытом среди авиакомпаний, аэропортов и органов управления воздушным движением, а также внедрение единых норм и правил полетов. Оценка регулярности осуществляется по показателям пунктуальности как отправлений, так и прибытия.

В Европе различают регулярность рейсов по аэропортам и авиакомпаниям: регулярность аэропортов оценивается по времени вылета, а регулярность авиакомпаний — по времени прибытия. Такой подход более обоснован, поскольку пассажирам важнее время прибытия, а не время вылета [3]. Авиаперевозчики и ассоциации регулярно оценивают деятельность аэропортов по показателю «пунктуальность отправления», чтобы стимулировать своевременность вылетов, координировать работу служб аэропортов и оптимизировать процессы обслуживания воздушных судов и пассажиров. Пунктуальные аэропорты демонстрируют показатели регулярности вылетов до 90%.

В Казахстане учет регулярности осуществляется только по авиакомпаниям, фиксируя вылет по моменту отрыва от полосы. В западной практике регулярность вылетов определяется по моменту отгону трапа, поскольку после этого процесс, зависящий от авиакомпании, завершен, и начинается работа аэропорта и служб управления воздушным движением. Важно, насколько эффективно организованы процедуры руления, свободна ли взлетная полоса и своевременно ли проведены все необходимые процедуры, например, как прием VIP-персон.

Что касается классификации задержек, в Европе применяются другие критерии: рейс, задержавшийся менее чем на 15 минут, считается выполненным вовремя. Задержки подразделяются на категории: до 1 часа и свыше 3 часов, что позволяет более четко оценить ситуацию. Ассоциация европейских авиаперевозчиков использует простую схему для подсчета - время отправления для аэропорта и время прибытия для авиакомпаний с задержкой до 15 минут считаются регулярными. Рейтинг пунктуальности разделяется на дальнемагистральные и ближнемагистральные рейсы, что позволяет более объективно оценить ситуацию, так как на дальнемагистральных маршрутах встречные воздушные потоки играют большую роль, чем на коротких. Например, по данным ассоциации, немецкая Lufthansa в 2024 году не допустила отмены рейсов и показала лучший результат по регулярности на дальнемагистральных направлениях, но только 13-е место на европейских рейсах, выполнив 98,9% запланированных полетов. Отменой считается случай, когда полет отменен по любой причине более чем за три дня до вылета.

В международной практике системы мониторинга авиационного транспорта, используемые государственными и общественными организациями, полностью автоматизированы и основываются на электронных запросах к базам данных систем бронирования, управления отправлениями, ресурсами, визуализации, информирования и других. Формат запросов, их частота и степень детализации могут варьироваться в зависимости от актуальных потребностей пользователей данных. Изначальный сбор информации осуществляется государственными органами и отраслевыми ассоциациями, на основе которой общественные организации проводят анализ и публикуют статистические данные в различных разрезах, удовлетворяя запросы потребителей и защищая их интересы.

Что касается мониторинга и публикации данных о пунктуальности рейсов, примером может служить методика, применяемая в Великобритании. Администрация гражданской авиации ежемесячно собирает данные о пунктуальности рейсов, используя информацию из баз данных аэропортов, координирующих организаций и других источников. Эти данные публикуются в открытом доступе на официальном сайте администрации. В качестве примера оценки регулярности полетов можно привести общественную организацию, которая собирает, анализирует и публикует данные о пунктуальности рейсов всех авиакомпаний, выполняющих рейсы в аэропорты Великобритании. Эта организация

получает информацию от администрации гражданской авиации, проводит анализ и составляет рейтинг авиакомпаний по общей пунктуальности, а также по пунктуальности на отдельных направлениях.

Ежегодно авиакомпании по всему миру терпят убытки из-за задержек рейсов, которые могут достигать миллионов, в то время как пассажиры сталкиваются с неудобствами и различными потерями.

Важно понимать, что каждый рейс выполняется в строго определённое время согласно расписанию и назначенному слоту, который выделяется для конкретного воздушного судна в определённый день. Слот — это время вылета или прибытия самолёта на конкретную дату, установленное в расписании аэропорта [4].

Следует отметить, что задержки, как технического, так и организационного характера, могут быть связаны с неисправностями в работе аэропорта, авиакомпании или служб управления воздушным движением (УВД). К задержкам также могут привести такие ситуации, как задержка группы пассажиров (например, VIP-класса или первого класса).

Задержка взлёта может повлиять и на время прибытия рейса (если экипаж не может компенсировать задержку во время полета и прибыть по расписанию), что в итоге приводит к общей задержке рейса.

Одной из ключевых причин задержек являются неблагоприятные погодные условия, которые могут затруднить или сделать невозможным выполнение взлёта и посадки. Даже современное воздушное судно с высокотехнологичным оборудованием не всегда может приземлиться на взлётно-посадочную полосу из-за плохих погодных условий, что иногда приводит к полной приостановке работы аэропорта.

Можно выделить шесть основных причин задержек рейсов (рис. 1).

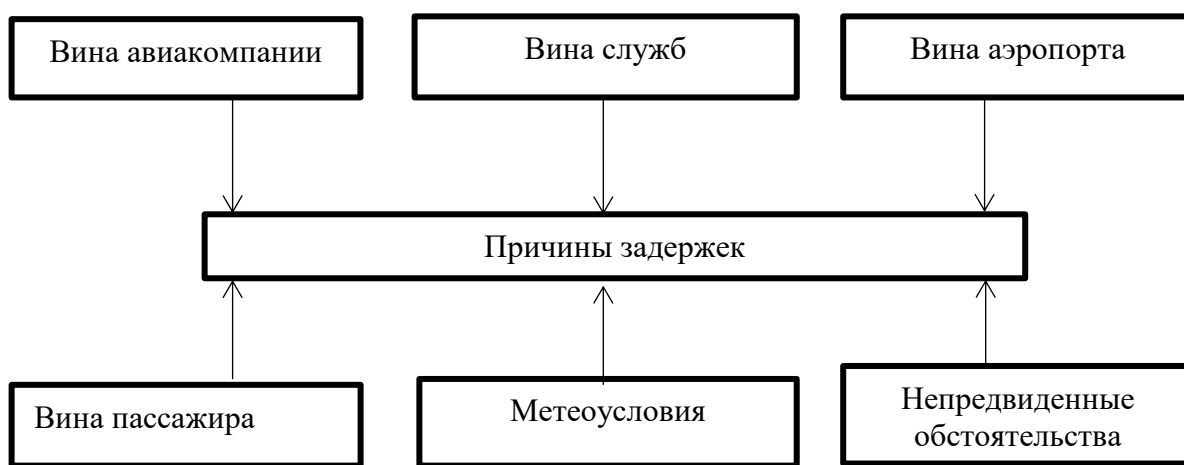


Рисунок 1 – Укрупненная классификация задержек [5]

Кроме погодных условий, задержки могут происходить по вине самой авиакомпании или аэропорта (обслуживающей компании). Причины могут быть техническими или организационными. Технические причины связаны с неисправностями самолёта или аэродромного оборудования. Такие задержки часто происходят из-за необходимости проведения обязательного технического обслуживания (например, Daily Check или Weekly Check) в рамках ограниченного времени перед полетом.

Организационные задержки могут быть вызваны проблемами со стороны аэропорта, такими как перегрузка перрона или необходимость ожидания на стоянке для буксировки или руления. Задержки могут также быть связаны с работой обслуживающих компаний, например, поздняя загрузка или разгрузка бортового питания, задержки с предоставлением аэродромной спецтехники (трапы, тягачи, багажные ленты, машины для

противообледенительной обработки, пассажирские автобусы и другие), а также с несвоевременным выполнением операций по погрузке или выгрузке багажа и почты.

Задержки рейсов по вине авиакомпаний могут быть вызваны различными обстоятельствами. Например, самолет может отсутствовать в аэропорту вылета из-за опоздания с предыдущего рейса или нехватки резервных воздушных судов. Также возможны задержки, связанные с ожиданием трансферных пассажиров, сбоями в системе регистрации или посадки.

Задержки рейсов могут возникать по разным причинам, включая непредвиденные обстоятельства, не зависящие от авиакомпании. Например, к таким ситуациям относятся экстренные медицинские случаи у пассажиров, нарушение общественного порядка или природные катастрофы. Кроме того, причиной задержки может стать столкновение с птицами, забастовки персонала, отсутствие необходимого экипажа, необходимость соблюдения норм отдыха летного состава или превышение допустимого взлетного веса самолета.

### **Результаты и их обсуждение.**

Казахстанский рынок авиаперевозок в последние годы находится в стадии интенсивного развития, несмотря на проблемы прошлого, такие как устаревшие производственные традиции, несовременные нормативно-правовые акты и избыточные парки неэффективных воздушных судов. Эти негативные тенденции существенно влияют на экономику авиаперевозок, вынуждая авиакомпании акцентировать внимание на краткосрочных целях, часто используя демпинг, что отражается на качестве услуг, в том числе на регулярности полетов.

Надежность вылетов — это важный показатель эффективности технической эксплуатации воздушных судов, который отражает экономическую результативность работы авиакомпании. Чем выше уровень эксплуатационной надежности рейсов, тем успешнее авиакомпания управляет эксплуатацией своих самолетов, тем лучше качество обслуживания и более эффективна политика безопасности [6].

Регулярность рейсов имеет свою цену, которая включает расходы на резервные воздушные суда, высокую стоимость новых современных самолетов, упущенную выгоду из-за сглаживания пиковых времен прилетов и вылетов, потери от уменьшения интенсивности полетов, а также организационные затраты на улучшение бизнес-процессов. Поэтому принуждать авиакомпании к достижению стопроцентной регулярности полетов административными мерами нецелесообразно, поскольку это отразится на стоимости билетов. Для большинства пассажиров важнее сбалансированное соотношение стоимости авиаперевозки и приемлемого качества услуги [7].

Для проведения математического расчета данного исследования был сделан анализ текущей статистики задержек (рисунок 2).

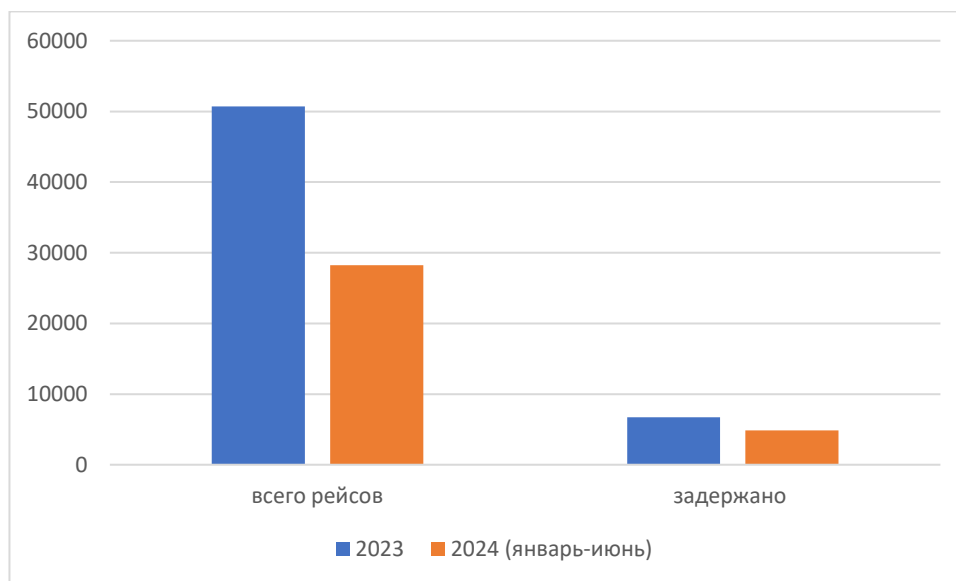


Рисунок 2 – Анализ текущей статистики задержек [1]

Процент задержек в 2023 году составил 13,28%, а за период с января по июнь 2024 года процент задержек составил 17.16%.

Оценивая экономические потери от задержек, в среднем, одна задержка обходится авиакомпаниям от 1 000 до 10 000 долларов (учитывая компенсации, перерасход топлива, организационные издержки). В таком случае годовой ущерб в 2024 году составит от 9.68 млн. долларов до 96.84 млн. долларов США. Число 9684 в расчёте — это прогнозируемое количество задержанных рейсов в 2024 году, если сохранятся текущие темпы задержек, зафиксированные в первой половине года. 4 842 — количество задержек за первую половину 2024 года. Умножение на 2 предполагает, что динамика задержек во втором полугодии останется такой же, как и в первом. При средней загрузке 120 пассажиров и количества задержанных рейсов - 9864, задержки рейсов окажут влияние на 1.16 млн. пассажиров.

Задержки авиарейсов могут быть вызваны различными факторами: погодными условиями, техническими неисправностями, загруженностью аэропортов, операционными сбоями и проблемами в управлении воздушным движением. Для минимизации задержек рекомендуется применять операционные, технологические, инфраструктурные и организационные меры.

К операционным мерам можно отнести оптимизацию расписания, т.е. разработка реалистичных графиков с учетом возможных задержек, сезонности и метеоусловий. А также это диверсификация маршрутов, которая позволяет использование альтернативных аэродромов при перегрузке главных хабов. Кроме того, очень важно сокращать время наземного обслуживания за счёт автоматизации процессов. Наземное обслуживание воздушных судов включает заправку, техническую проверку, загрузку багажа и посадку пассажиров. Автоматизация этих процессов позволяет значительно сократить время простоя самолетов в аэропорту, что уменьшает задержки и повышает пунктуальность рейсов. Например, автоматизация процессов заправки топлива может осуществляться через цифровые системы. Внедрение умных топливных датчиков на самолётах, которые автоматически передают данные о текущем уровне топлива. Использование автоматизированных расчетных систем, которые смогут прогнозировать необходимый объем заправки в зависимости от маршрута и погодных условий. Интеграция с ERP-системами аэропорта позволит заранее планировать время и место заправки. Также можно осуществлять заправку самолетов во время других операций, таких как погрузка багажа, проверка оборудования. Или использовать автоматизированные топливозаправщики,



которые могут самостоятельно подключаться к самолету. В результате таких мероприятий возможно сокращение времени заправки на 15–25%, а также снизятся простои воздушных судов.

Если говорить дальше об автоматизации, то её применение возможно при проверке бортовых систем. Например, использование цифровых двойников и IoT позволит через датчики на борту передавать данные о состоянии систем в реальном времени в диспетчерский центр авиакомпании или аэропорта. А с помощью системы Big Data анализировать эти данные и прогнозировать возможные неисправности заранее. Для ускоренной диагностики и удаленного контроля технический персонал будет получать автоматизированные чек-листы с данными о состоянии самолета еще до его приземления. Также возможно использование роботов и дронов для наружного осмотра корпуса и двигателей. В результате сокращение времени диагностики достигнет с 30–40 минут до 10–15 минут.

Для погрузки и обработки багажа также рекомендуется использовать автоматизацию. Например, использование RFID-меток и конвейерных систем. RFID-метки позволяют автоматически отслеживать багаж и его загрузку. Подключение конвейерных лент напрямую к самолету для ускоренной загрузки и выгрузки. А если роботизировать системы обработки багажа, автоматизировать сортировочные комплексы, которые будут распределять багаж без участия человека. Внедрение автономных погрузчиков заменят ручной труд. В результате времени обработки багажа сократится с 25–30 минут до 10–15 минут.

Таким образом итоговый эффект автоматизации состоит в том, что сократится время наземного обслуживания с 60–90 минут до 30–50 минут. Это позволит снизить количество задержек из-за длительной подготовки самолета к следующему рейсу, позволит оптимизировать работу аэропортов и авиакомпаний, повысит пропускную способность. Также автоматизация процессов наземного обслуживания повысит пунктуальность рейсов, снизит операционные затраты и сделает воздушные перевозки более эффективными.

Ещё одна рекомендация для снижения задержек рейсов – это метод разгона тумана азотом в районе посадки самолета.

Разгон тумана с использованием жидкого азота является одним из методов повышения видимости в зоне аэропорта при посадке самолетов. Этот метод основан на физико-химических свойствах азота и его способности изменять структуру тумана. Туманы в районе Международного аэропорта Алматы наиболее часто наблюдаются в осенний-зимний период, особенно в ноябре и декабре. В это время вероятность их возникновения составляет около 60%. В целом, в городе Алматы и его окрестностях туманы фиксируются на протяжении 50–70 дней в году. Продолжительность туманов может варьироваться от нескольких часов до нескольких суток. Например, в декабре 2018 года густой туман держался в Алматы несколько дней подряд, что привело к задержкам авиарейсов.

В чём состоит принцип работы данного метода. Азот охлаждает влагу, находящуюся в тумане. Жидкий азот ( $-196^{\circ}\text{C}$ ) при распылении быстро испаряется, охлаждая воздух и приводя к конденсации водяного пара, что способствует укрупнению мелких капель тумана. Далее происходит осаждение капель. Более крупные капли воды быстрее оседают на поверхность, что снижает плотность тумана и улучшает видимость. Затем при испарении азота происходит локальное увеличение объема сухого газа, что способствует вытеснению влажного воздуха из зоны посадки.

Технология применения данного метода заключается в следующем. Жидкий азот распыляется в зоне ВПП (взлетно-посадочной полосы) и глиссады с помощью специальных аэрозольных установок или мобильных емкостей с системой форсунок. Для максимального эффекта разгон проводят непосредственно перед посадкой самолета. Системы могут быть стационарными (установленными на аэродроме) или мобильными (на спецтранспорте).

Преимуществами данного метода является быстрое действие (эффект наступает в течение нескольких минут), экологическая безопасность (азот составляет 78% атмосферы), возможность применения при критически низкой видимости.

Но не смотря на преимущества, также имеются ограничения и недостатки. Это высокая стоимость жидкого азота и оборудования, ограниченная эффективность при сильном тумане с высокой влажностью, кратковременный эффект – повторное образование тумана возможно через 15-30 минут.

Разгон тумана азотом – перспективный, но затратный метод повышения видимости в аэропортах. Он эффективен при умеренном тумане и может применяться как дополнительная мера в сочетании с другими технологиями.

Для анализа и прогнозирования задержек самолётов можно использовать различные математические методы в зависимости от целей исследования.

Рассмотрим метод ARIMA (Auto Regressive Integrated Moving Average) для прогнозирования задержек рейсов [8].

Допустим, у нас есть данные о задержках рейсов в минутах за предыдущие 30 дней. Нам нужно спрогнозировать задержку на следующие 7 дней.

Анализ методом ARIMA выполняется по следующим этапам:

1. осуществляется анализ данных: изучается временной ряд задержек. Временной ряд задержек рейсов – это последовательность данных о задержках рейсов, упорядоченная по времени. Он может содержать: ежедневные/ежемесячные данные (количество задержек или их средняя продолжительность за определенный период); сезонные колебания (например, рост задержек в зимний период из-за погодных условий); долгосрочные изменения (например, снижение задержек благодаря автоматизации); случайные колебания (влияние непредвиденных факторов). Анализ временного ряда позволяет предсказывать задержки и находить причины их изменения.

2. Определение параметров ARIMA (p, d, q):

- p — порядок авторегрессии (AR);
- d — число дифференцирований (I);
- q — порядок скользящего среднего (MA).

3. Обучение модели на исторических данных — это процесс, при котором алгоритм анализирует прошлые данные о задержках рейсов, чтобы выявить закономерности и делать прогнозы на будущее.

4. Далее выполняется прогнозирование задержек на 7 дней вперед (рисунок 3).

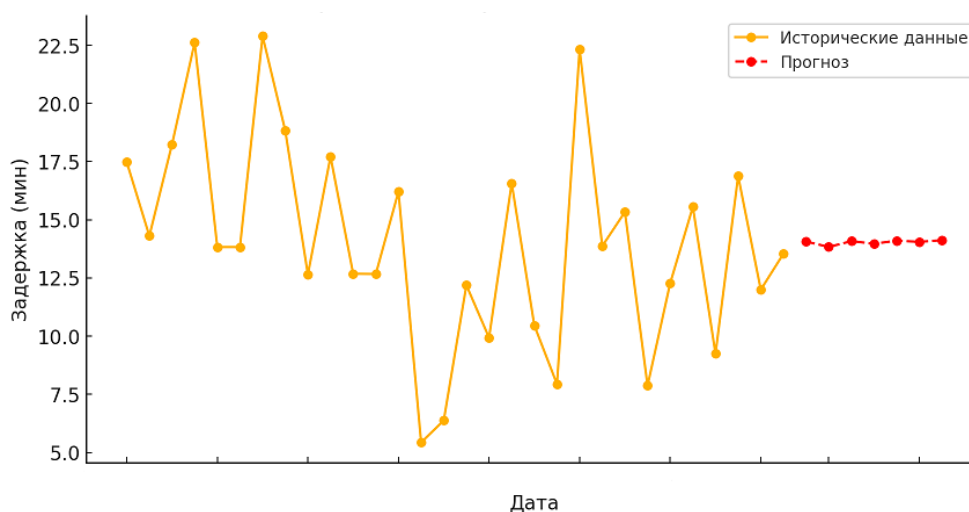


Рисунок 3 – Прогноз Задержек методом ARIMA [1]

Исходя из прогноза задержек на 7 дней: значения колеблются в диапазоне 13.8–14.1 минут. График показывает, что прогноз стабильный, без резких скачков.

Этот подход позволяет предсказывать задержки рейсов, чтобы улучшить планирование и управление полетами.

Если задержки зависят от сезонных факторов (например, увеличиваются осенью и зимой из-за плохой погоды), то лучше использовать SARIMA (Seasonal ARIMA) [8].

Метод SARIMA (Seasonal ARIMA) расширяет ARIMA за счет учета сезонных компонент, которые повторяются через фиксированные интервалы времени (например, каждый год, квартал или месяц).

Модель SARIMA обозначается как:

$$1. \text{ SARIMA}(p, d, q) \times (P, D, Q, s)$$

где:

(p, d, q) – параметры ARIMA (авторегрессия, интегрирование, скользящее среднее);

(P, D, Q, s) – сезонные аналоги этих параметров с периодичностью s (например, s = 12 для месячных данных).

Допустим, у нас есть данные о средних задержках рейсов (в минутах) за последние 36 месяцев, и мы знаем, что задержки увеличиваются осенью и зимой.

Определяем сезонность: допустим, задержки изменяются раз в 12 месяцев  $\rightarrow s = 12$ . Подбираем параметры ARIMA: используем автокорреляцию (ACF) и частичную автокорреляцию (PACF). Модель SARIMA (1,1,1)  $\times$  (1,1,1,12), которая хорошо описывает сезонные колебания.

Таблица 1 - Расчет и прогнозирование задержек в 2025 году

Месяц	Фактическая задержка (мин)	Прогноз SARIMA (мин)
Янв 2025	12.5	12.4
Фев 2025	13.1	13.0
Мар 2025	12.7	12.6
Апр 2025	11.8	11.9
...	...	...
Окт 2025	-	14.5
Ноя 2025	-	15.2
Дек 2026	-	16.1

SARIMA предсказывает, что в осенне-зимний период задержки вырастут и достигнут 16.1 минут в декабре 2025 года.

Таким образом, SARIMA учитывает годовую сезонность и корректно предсказывает рост задержек зимой. Метод полезен для авиакомпаний и аэропортов, чтобы заранее учитывать сезонные факторы. Можно улучшить точность, добавив погодные данные (температуру, осадки, туман).

### Заключение.

В ходе исследования была проведена детальная оценка временного ряда задержек авиарейсов с использованием моделей ARIMA и SARIMA. Анализ показал, что временной ряд обладает сезонными колебаниями, что делает применение стандартной модели ARIMA менее эффективным. В связи с этим была использована SARIMA, позволяющая учитывать сезонные компоненты.

Прогноз на следующий период показал увеличение/снижение количества задержек, что может помочь авиакомпаниям в оптимизации расписания и ресурсного планирования.

В дальнейшем возможно расширение модели путем включения дополнительных факторов, таких как погодные условия, загруженность аэропортов и регламентные работы.

Использование моделей временных рядов, в частности SARIMA, является эффективным инструментом для прогнозирования задержек авиарейсов, что позволяет повысить точность планирования и минимизировать издержки авиакомпаний. Проблема регулярности авиарейсов в Казахстане остается актуальной и многогранной, охватывая технические, организационные и погодные факторы. Анализ международных практик показывает, что повышение пунктуальности требует комплексного подхода, включающего оптимизацию аэропортовой инфраструктуры, внедрение современных стандартов управления воздушным движением и повышение эффективности работы авиакомпаний.

Несмотря на позитивные изменения в отрасли, такие как рост конкуренции и повышение внимания к правам пассажиров, полное устранение задержек невозможно без значительных инвестиций и совершенствования нормативно-правовой базы. Важно найти баланс между регулярностью рейсов и экономической эффективностью авиаперевозок, обеспечив приемлемое соотношение стоимости билетов и качества услуг.

### Список литературы

1. В КГА объяснили задержки авиарейсов в Казахстане. (2024). URL: [https://www.kt.kz/rus/society/v\\_kga\\_ob\\_yasnili\\_zaderzhki\\_aviareysov\\_v\\_kazahstane\\_1377966139.html](https://www.kt.kz/rus/society/v_kga_ob_yasnili_zaderzhki_aviareysov_v_kazahstane_1377966139.html) (дата обращения: 15.03.2025).
2. Потапова, Д.Ю. (2018). Основные причины задержек авиационных рейсов. Журнал «Мир транспорта», том 16, № 3.
3. Наумова, Д.А. (2012). Методики оценки регулярности полетов авиакомпаний. Научный вестник Московского государственного технического университета гражданской авиации, №181.
4. Купрюхин, А.А. & Бочанова М.В. (2024). Аналитический обзор причин нарушения регулярности полетов воздушных судов гражданской авиации. *Economy and Business: Theory and Practice*, vol. 9-1 (115).
5. Данилов, В. Н. & Кузнецова, И. И. (2019). Регулярность авиаперевозок как фактор конкурентоспособности на рынке. Журнал транспортной логистики, 4(1), 34-44.
6. Алексанян, А.Р., Ицкович, А.А. & Евдокимова, А.Д. Актуализация методики расчета фактического показателя надежности вылетов гражданских судов. Научный вестник МГТУ ГА. 2017;20(6):81-88. <https://doi.org/10.26467/2079-0619-2017-20-6-81-88>
7. Рублев В.В. (2020). Перспективы развития рынка бюджетных пассажирских авиаперевозок в Республике Казахстан. Вестник АГТУ. Сер.: Экономика, № 2.
8. ARIMA & SARIMA: прогнозирование временных рядов в реальном мире. (2023). URL: <https://neptune.ai/blog/arima-sarima-real-world-time-series-forecasting-guide> (дата обращения: 17.03.2025).

### References

1. V KGA ob "yasnili zaderzhki aviarejsov v Kazahstane. (2024). URL: [https://www.kt.kz/rus/society/v\\_kga\\_ob\\_yasnili\\_zaderzhki\\_aviareysov\\_v\\_kazahstane\\_1377966139.html](https://www.kt.kz/rus/society/v_kga_ob_yasnili_zaderzhki_aviareysov_v_kazahstane_1377966139.html) (data obrashcheniya: 15.03.2025).
2. Potapova D.Yu. (2018). Osnovnye prichiny zaderzhkek aviacionnyh rejsov. Zhurnal «Mir transporta», tom 16, № 3.
3. Naumova, D.A. (2012). Metodiki ocenki regulyarnosti poletov aviakompanij. Nauchnyj vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta grazhdanskoj aviacii, №181.

4. Kupryuhin, A.A. & Bochanova, M.V. (2024). Analiticheskij obzor prichin narusheniya regulyarnosti poletov vozduzhnyh sudov grazhdanskoj aviacii. Economy and Business: Theory and Practice, vol. 9-1 (115).

5. Danilov, V. N. & Kuznecova, I. I. (2019). Regulyarnost' aviaperevozok kak faktor konkurentosposobnosti na rynke. Zhurnal transportnoj logistiki, 4(1), 34-44.

6. Aleksanyan, A.R., Ickovich, A.A. & Evdokimova, A.D. Aktualizaciya metodiki rascheta fakticheskogo pokazatelya nadezhnosti vyletov grazhdanskih sudov. Nauchnyj vestnik MGTU GA. 2017;20(6):81-88. <https://doi.org/10.26467/2079-0619-2017-20-6-81-88>

7. Rublev V.V. (2020). Perspektivy razvitiya rynka byudzhetnyh passazhirskih aviaperevozok v Respublike Kazahstan. Vestnik AGTU. Ser.: Ekonomika, № 2.

8. ARIMA & SARIMA: prognozirovaniye vremennyh ryadov v real'nom mire. (2023). URL: <https://neptune.ai/blog/arima-sarima-real-world-time-series-forecasting-guide> (data obrashcheniya: 15.03.2025).

## ҚАЗАҚСТАННЫҢ ӘУЕ ТАСЫМАЛДАРЫНЫҢ ЖҮЙЕЛІЛІГІ МӘСЕЛЕЛЕРІ

***Аңдатпа.** Рейстердің кешігуі мәселесі жолаушылар сервисіне, авиакомпаниялардың операциялық процестеріне және әуежайлардың экономикасына айтарлықтай әсер ете отырып, Қазақстанның әуе тасымалы саласындағы негізгі мәселелердің бірі болып табылады. Жолаушылар ағынының өсуін және әуе нарығының белсенді дамуын ескере отырып, кідірістердің себептерін анықтау және жою саланың бәсекеге қабілеттілігін арттыру үшін маңызды міндетке айналады.*

*Зерттеу пәні рейстердің кешігу себептері және олардың Қазақстан авиакомпаниялары мен әуежайлары жұмысының тиімділігіне әсері болып табылады. Зерттеу барысында Қазақстандағы рейстердің ағымдағы жай-күйін талдау, кідірістерге әсер ететін негізгі факторларды анықтау, ұшу кестесін бұзудың экономикалық және әлеуметтік салдарын бағалау, кідірістерді азайту және ұқыптылықты арттыру бойынша ұсынымдар әзірлеу сияқты міндеттер шешілді.*

*Зерттеудің мақсаты Қазақстанның әуе тасымалы нарығындағы рейстердің кешігу себептерін зерделеу және оларды барынша азайту жөніндегі тиімді шараларды әзірлеу болды.*

*Жұмыста статистикалық талдау, сараптамалық сауалнамалар, нормативтік құжаттарды талдау әдістері, сондай-ақ жетекші авиатасымалдаушылардың ұқыптылық көрсеткіштерін салыстырмалы талдау әдістері пайдаланылды.*

*Зерттеудің негізгі нәтижелері ауа райы жағдайларын, техникалық ақауларды, әуежайлардың толып кетуін және ұйымдастырушылық факторларды қоса алғанда, рейстердің кешігуінің негізгі себептерін анықтау, Қазақстанның ірі авиакомпанияларының рейстерінің кешігу статистикасын талдау, әуе тасымалдарын жоспарлау және басқару процестерін оңтайландыру бойынша ұсыныстар әзірлеу болды.*

*Зерттеу нәтижесінде жедел жоспарлау жүйесін жақсартуды, кідірістерді болжау үшін цифрлық технологияларды енгізуді, әуежай инфрақұрылымын жаңғыртуды және нормативтік базаны жетілдіруді қоса алғанда, рейстердің кідірістерін азайту бойынша шаралар ұсынылды. Осы ұсынымдарды іске асыру авиатасымалдардың уақтылығын арттыруға және Қазақстанда жолаушыларға қызмет көрсету сапасын жақсартуға мүмкіндік береді.*

***Түйін сөздер:** рейстердің кешігуі, әуе көлігі нарығы, рейстердің жүйелілігі, әуежайлар, авиакомпаниялар, қызмет көрсету сапасы, ұшу қауіпсіздігі.*

## PROBLEMS OF REGULARITY OF AIR TRANSPORTATION IN KAZAKHSTAN

**Abstract.** *The problem of flight delays is one of the key issues in the field of air transportation in Kazakhstan, having a significant impact on passenger service, airline operations and airport economics. Given the growth of passenger traffic and the active development of the aviation market, identifying and eliminating the causes of delays is becoming an important task to increase the competitiveness of the industry.*

*The subject of the study is the causes of flight delays and their impact on the efficiency of airlines and airports in Kazakhstan. The study solved such tasks as analyzing the current state of flight punctuality in Kazakhstan, identifying the main factors influencing delays, assessing the economic and social consequences of flight schedule violations, and developing recommendations to reduce delays and improve punctuality.*

*The purpose of the study was to study the causes of flight delays in the Kazakhstan air transportation market and to develop effective measures to minimize them.*

*The paper uses methods of statistical analysis, expert surveys, analysis of regulatory documents, as well as a comparative analysis of punctuality indicators of leading air carriers.*

*The main results of the study were the identification of key causes of flight delays, including weather conditions, technical malfunctions, airport congestion and organizational factors, the analysis of flight delay statistics of the largest airlines in Kazakhstan, the development of proposals to optimize the processes of planning and management of air transportation.*

*As a result of the study, measures have been proposed to reduce flight delays, including improving the operational planning system, introducing digital technologies to predict delays, modernizing airport infrastructure, and improving the regulatory framework. The implementation of these recommendations will improve the punctuality of air transportation and improve the quality of passenger service in Kazakhstan.*

**Keywords:** *flight delays, air transport market, flight regularity, airports, airlines, quality of service, flight safety.*

### Сведение об авторах

Абжапбарова Айнур Жадигеровна	к.т.н., профессор, Академия гражданской авиации, Алматы, Республика Казахстан, <a href="mailto:Ainur.abzhapbarova@mail.ru">Ainur.abzhapbarova@mail.ru</a>
Гармаш Ольга Валерьевна	к.т.н., ассоциированный профессор АГА, Академия гражданской авиации, Алматы, Республика Казахстан, e-mail: <a href="mailto:olm78@mail.ru">olm78@mail.ru</a>

### Авторлар туралы мәлімет

Абжапбарова Айнур Жадыгерқызы	т.ғ.к., профессор, Азаматтық авиация академиясы, Алматы, Қазақстан Республикасы, <a href="mailto:Ainur.abzhapbarova@mail.ru">Ainur.abzhapbarova@mail.ru</a>
Гармаш Ольга Валерьевна	т. ғ. к., аға қауымдастырылған профессор, Азаматтық авиация академиясы, Алматы, Қазақстан Республикасы, e-mail: <a href="mailto:olm78@mail.ru">olm78@mail.ru</a>

### Information about the authors

Abzhapbarova Ainur Zhadigerovna	Candidate of Technical Sciences, Professor, Academy of Civil Aviation, Almaty, Republic of Kazakhstan, <a href="mailto:Ainur.abzhapbarova@mail.ru">Ainur.abzhapbarova@mail.ru</a>
Garmash Olga Valeryevna	Candidate of Technical Sciences, Associate Professor AGA, Academy of Civil Aviation, Almaty, Republic of Kazakhstan, e-mail: <a href="mailto:olm78@mail.ru">olm78@mail.ru</a>

УДК 629.7  
МРНТИ 73.37.61  
[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_6](https://doi.org/10.53364/24138614_2025_36_1_6)

В.О. Макогонова<sup>1\*</sup>, И.Ж. Асильбекова<sup>1</sup>

<sup>1</sup> Академия гражданской авиации, Алматы, Казахстан

E-mail: cherry-98-23-04@mail.ru\*

## РАЗРАБОТКА МЕТОДИКИ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ГРУЗОВЫХ ПЕРЕВОЗОК

***Аннотация.** В современных условиях, несмотря на развитие технологий, компании продолжают сталкиваться с проблемами оптимизации грузоперевозок. Одним из решений является внедрение автоматизированных систем управления логистикой, однако этот процесс сопряжен с рядом вызовов, включая киберугрозы, рыночные риски и необходимость значительных финансовых вложений.*

*Авиагрузоперевозки остаются одним из наиболее востребованных способов транспортировки, обеспечивая быструю и эффективную доставку на международном уровне. Казахстан, обладая стратегически выгодным географическим положением, имеет высокий транзитный потенциал. Одним из перспективных направлений является развитие аэропорта Алматы как международного грузового хаба, что позволит увеличить грузопоток и укрепить позиции страны в глобальной торговле. Однако высокая стоимость аэропортовых услуг может стать сдерживающим фактором, что требует пересмотра тарифной политики и увеличения доходов за счет неавиационной деятельности.*

*Модернизация инфраструктуры аэропорта, внедрение инновационных технологий и цифровых решений, таких как SAP Transportation Management, позволят повысить точность и эффективность логистических процессов. Это не только улучшит уровень сервиса, но и усилит конкурентные позиции Казахстана в сфере грузовых авиаперевозок, способствуя устойчивому экономическому развитию и интеграции в мировую транспортную систему.*

***Ключевые слова:** грузовые перевозки, авиационная компания, повышение эффективности, SAP Transportation Management, аэропорт, маркетинг.*

### **Введение.**

Актуальность темы маркетинговых исследований в секторе грузовых перевозок обусловлена динамичным развитием глобальных логистических систем, где оперативность и оптимизация процессов играют ключевую роль в формировании конкурентных преимуществ компаний. Для разработки эффективной методики проведения маркетинговых исследований важно учитывать как качественные, так и количественные подходы, а также интеграцию современных цифровых технологий.

В работе [1] подчеркивается, что «оптимизация логистических процессов требует интеграции информации с целью обеспечения прозрачности и оперативного реагирования на изменения рынка». Это свидетельствует о необходимости адаптации методов маркетинговых исследований под специфику логистических процессов, где цифровизация и аналитику следует рассматривать как неотъемлемые элементы для повышения эффективности перевозок.

Авторы [2] в своём трактате обращают внимание на важность стратегического планирования и прогнозирования спроса. Они отмечают, что «аналитика и прогнозирование играют ключевую роль в управлении цепями поставок», что позволяет не

только оптимизировать запасы, но и адаптировать маркетинговые стратегии к изменяющимся условиям рынка. Такой подход может быть перенесён на область грузовых перевозок посредством построения числовых моделей, позволяющих оценить эффективность систем распределения.

Работа [3] демонстрирует, что интегрированные системы управления транспортными потоками способны снижать операционные затраты и улучшать качество клиентского сервиса. Авторы подчёркивают необходимость комбинации качественных и количественных методов для получения целостной картины функционирования логистической системы, где маркетинговые исследования играют роль в выявлении как явных, так и скрытых факторов, влияющих на эффективность перевозок.

Классическая работа [4] предоставляет фундаментальные принципы маркетинговых исследований, основное внимание уделяя анализу потребительского поведения. Авторы утверждают, что «глубокий анализ потребительского поведения является основой для формирования конкурентоспособных стратегий», что актуально и для грузового сектора, где необходимо адаптировать маркетинговые инструменты для выявления потребности в специфических услугах и сервисах.

Наконец, автор [5] в своей работе предлагает аналитические инструменты, направленные на оценку конкурентной среды. Применение данных инструментов в логистике помогает структурировать рыночную информацию и выявить ключевые факторы успеха, что является основой для разработки комплексных методик маркетинговых исследований.

В условиях глобализации и стремительного развития транспортной инфраструктуры, эффективные грузовые перевозки становятся ключевым фактором конкурентоспособности компаний и экономики в целом. Грузовой транспорт играет важную роль в обеспечении логистических цепочек, влияя на стоимость и скорость доставки товаров. Однако, несмотря на значительные достижения в области технологий и управления, многие компании сталкиваются с проблемами, связанными с оптимизацией процессов грузоперевозок.

В условиях современности борьба за свое место на рынке обострилась, и факт неучтенности и незнания рынка компанией-производителем ставит под угрозу собственное существование и неизбежно ведет ее к разорению.

В качестве задачи, цели коммерческой деятельности всех авиапредприятий выступает максимизация прибыли путем рационального планирования и управления перевозками на воздушном транспорте на основе активного пользования современных методик маркетинга, экономико-математических моделей, информационных технологий.

Одним из главных ключей организации грузовых авиаперевозок – это использование аэропорта как основного терминала. Аэропорт будет выступать пунктом сбора и распределения грузов. Пункт, который обрабатывает, хранит и отправляет грузы в последующую точку. Исходя из этих действий, грузы могут прибыть быстро и именно в пункт назначения.

Превосходством перевозок грузов на воздушном транспорте является ее скорость. Ежедневно авиация преодолевает огромные расстояния за короткий период времени. Данный факт определяет ее как идеальный вариант для доставки товаров на дальние расстояния, что особенно важно для международных перевозок, которая требует своевременную и быструю доставку груза между несколькими странами.

Воздушная логистика является неотъемлемой частью международной торговли и глобализации экономики. Она позволяет компаниям быстро и надежно доставлять товары в любую точку мира, открывая новые возможности для развития международных бизнес-связей.

Актуальность темы обусловлена необходимостью адаптации компаний к изменяющимся условиям рынка, а также потребностью в более глубоком понимании



потребительского поведения и предпочтений. В условиях высокой конкуренции и постоянных изменений в законодательстве, компании должны быть готовы к быстрому реагированию на вызовы и возможности, что требует системного подхода к проведению маркетинговых исследований.

Для достижения цели потребовалось решить ряд задач, поставленных в статье. В их рамках проведено исследование существующих процессов и технологий, используемых в грузовых перевозках, проанализировано текущее состояние перевозок в аэропорту г. Алматы, разработаны рекомендации по повышению их эффективности и спрогнозировано влияние внедрения предложенных мер. Кроме того, предложены способы оптимизации процессов и улучшения клиентского сервиса, а также осуществлена практическая оценка текущей эффективности перевозок с применением разработанной методики.

### **Материалы и методы исследования.**

Алматы является крупнейшим городом и коммерческой столицей Республики Казахстан. Расположен на маршруте Юго-Восток-Европа. Также город является местом базирования авиакомпании "Эйр Астана", которая является национальным перевозчиком. Через Алматы проходит Великий шелковый путь, который связывает Китай с несколькими регионами мира и город является грузовым центром Центрально-Азиатского региона.

Аэропорт TAV Алматы предоставляет комплексные складские услуги, предназначенные для удовлетворения различных потребностей в обработке и хранении грузов. Складская инфраструктура аэропорта оснащена современными технологиями для обеспечения высоких стандартов безопасности и эффективности. Среди предлагаемых услуг - обработка и хранение экспортных и импортных грузов, специализированные условия для скоропортящихся и температурочувствительных товаров, а также возможность временного хранения транзитных грузов.

Ключевые особенности складских услуг включают в себя наличие автоматизированных систем управления складом (WMS), которые поддерживают оперативное отслеживание и учет грузов. Это способствует снижению времени обработки и повышению точности процесса инвентаризации. Аэропорт также обеспечивает интеграцию с таможенными службами для ускорения процедур оформления и минимизации задержек. [6-7]

В этом году Казахстан планировал увеличить численность авиаперевозок в стране на 15%, чему поможет расширение флота на 50 единиц и модификация уже имеющихся воздушных судов. Республика на пути к укреплению места на рынке авиационных перевозок и привлекает все большее количество пассажиров из различных регионов.

Авиаперевозки груза и багажа самолетом по окончанию января – июля месяца нынешнего года осветили показатель в цифре 13.3 тыс. тонн – это прогресс на 4,2% чем в прошлом году. В городе Алматы цифры достигли 7.8тыс. тонн, в городе Шымкент 3.5 тыс. тонн, в городе Астана 1.9 тыс. тонн груза и багажа.

Общий грузооборот Республики Казахстан за январь-июль месяц данного года составил 28.9 млн. ткм – т.е. прирост на 1.3% по сравнению с прошлым годом этого же периода.

За 2022 год объем грузов, транспортированных воздушным транспортом в Казахстане, составил 24,5 тысяч тонн. Это отражает устойчивые операционные показатели авиационного сектора в данной области за указанный временной промежуток.

По данным Комитета по статистике Республики Казахстан известны следующие показатели грузоперевозок на воздушном транспорте за 2022-2024 годы, показанные в таблице 1, 2, 3:

Таблица 1 – Перевозка грузов и багажа воздушным транспортом 2022

Янв.	Фев.	Март	Апр.	Май	Июнь	Июль	Авг.	Сент.	Окт.	Ноя.	Дек.
1,2	1,6	1,6	1,9	2,0	2,2	2,4	2,4	2,1	2,1	2,4	2,7

Таблица 2 – Перевозка грузов и багажа воздушным транспортом 2023

Янв.	Фев.	Март	Апр.	Май	Июнь	Июль	Авг.	Сент.	Окт.	Ноя.	Дек.
1,9	2,0	2,1	1,9	1,8	1,8	1,9	1,9	2,0	1,8	2,2	2,6

Таблица 3 – Перевозка грузов и багажа воздушным транспортом 2024 [2]

Янв.	Фев.	Март	Апр.	Май	Июнь	Июль	Авг.	Сент.	Окт.	Ноя.	Дек.
1,7	1,9	2,2	2,1	2,3	2,2	2,7	2,7	2,6	X	X	X

Для анализа роста и спада объемов будет использована формула темпа прироста.

Темп прироста — это показатель, отражающий процентное изменение какого-либо величины за определенный период времени, обычно используется для оценки динамики развития экономики, производственных или других показателей.

Формула:

$$T = \frac{y^t}{y^{t-1}} * 100\% - 100, \quad (1)$$

где  $T$  – темп прироста,  $y^t$  – показатель текущего года,  $y^{t-1}$  – показатель прошлого года.

Исследование состояния аэропортового хозяйства РК позволяет выделить следующие основные проблемные аспекты в его развитии.

Объекты наземной авиатранспортной инфраструктуры страдают от значительного износа. Из-за нехватки средств на обновление и развитие, правительство обычно передавало аэродромы в управление аэропортам, однако арендные соглашения часто не оформлялись должным образом. Это привело к серьезному износу основных фондов и исключению некоторых аэродромов из гражданского оборота. Условия объектов наземной авиатранспортной инфраструктуры аэропортов Казахстана остаются неудовлетворительными и характеризуются высоким уровнем износа. Наблюдается несоответствие в обслуживании потребностей как пассажиров, так и авиакомпаний. Физический износ основных фондов аэропортов республиканского значения превышает 70 %, что превышает критическое значение. В соответствии с международной практикой, износ свыше 50 % указывает на потерю устойчивости в работе аэропорта, а износ более 65 % ставит под угрозу безопасность полетов, вызывает неожиданные капитальные вложения и угрожает финансовой стабильности, что может привести к уходу с рынка международных перевозок. Это обстоятельство тормозит развитие рынка авиационных услуг и реализует транзитный потенциал Казахстана.

Неравномерность распределения пассажирских и грузовых потоков по аэропортам Казахстана. За последние 15 лет рынок авиаперевозок в стране претерпел значительные изменения. В настоящее время из 21 действующего аэропорта 3 крупнейших контролируют 58,8 % всего рынка, причем на аэропорты Алматы и Нурсултан приходится около 50 % всех перевозок. Такое распределение свидетельствует о наличии олигополии. Региональные аэропорты находятся в гораздо менее выгодном положении по сравнению с указанными, что связано с разрушением местных и региональных авиационных маршрутов по почти всем областям, что, в свою очередь, ограничивает полное использование транзитного потенциала страны.

Нехватка инвестиций в содержание наземной инфраструктуры также является причиной сокращения количества аэропортов, что связано с недостаточным финансированием со стороны государства. Необходимо проанализировать

функционирование аэропортов на региональном уровне, возможности субсидирования из республиканского бюджета, изменения в Бюджетном кодексе и сотрудничество с местными властями. Решение этих вопросов должно стать одним из ключевых направлений развития авиационной отрасли Казахстана в ближайшем будущем.

Низкий уровень конкурентоспособности авиатранспортных терминалов (как для пассажиров, так и для грузов) является актуальной проблемой в сфере повышения конкурентоспособности гражданской авиации в Республике Казахстан. Ключевой задачей является модернизация аэровокзальных комплексов страны для достижения функциональных возможностей, соответствующих мировым стандартам.

Дополнительно, существующий монополизм большинства аэропортов в области наземного и аэровокзального обслуживания, а также заправки воздушных судов топливом, приводит к отсутствию конкуренции. Это приводит к необоснованному и непредсказуемому росту цен на услуги при недостаточном увеличении их качества или даже при его снижении.

Также в республике отсутствует прозрачная и экономически обоснованная система ставок аэропортовых сборов и тарифов. Текущие ставки аэропортовых сборов не соответствуют качеству предоставляемых услуг. В сочетании с монополией на поставку керосина в аэропортах, это негативно влияет на регулярность и прибыльность работы авиакомпаний, что зачастую приводит к задержкам и отменам рейсов, а иногда и к полному прекращению выполнения рейсов определенными авиакомпаниями.

Низкая эффективность системы государственного контроля в сфере аэропортовой деятельности. Регулирование авиационных услуг со стороны государства нацелено на удовлетворение нужд граждан и экономики в воздушных перевозках и авиационных услугах, а также на поддержку оборонных и безопасных аспектов, защиты государственных интересов, обеспечения безопасности полетов и соблюдения экологических норм. Основным документом, регулирующим данную сферу, является Воздушный кодекс, который определяет правовые основы для использования воздушного пространства РК и деятельности в области авиации, а также гарантирует соблюдение международных обязательств страны.

Мировой опыт государственного регулирования показывает, что вопросы государственного регулирования напрямую связаны с качеством предоставляемых услуг. В этой связи государственное регулирование деятельности предприятий авиатранспорта и, в частности аэропортов, является необходимой мерой, но для успешного функционирования авиатранспортного рынка необходимо учесть опыт западноевропейской модели управления.

В грузовом терминале аэропорта Нарита в Японии внедрены передовые инновационные технологии, направленные на оптимизацию и обеспечение эффективности логистических процессов. Одной из ключевых систем является полностью автоматизированный комплекс для обработки грузов, включающий в себя роботизированные сортировочные линии и автономные транспортные средства для внутреннего перемещения грузов. Эти технологии гарантируют точность и безопасность операций, снижая временные и операционные затраты. Также внедрены интеллектуальные системы мониторинга, которые используют анализ больших данных и искусственный интеллект для прогнозирования спроса и планирования ресурсов, что позволяет снизить риск задержек и повысить пропускную способность терминала. [8-9]

Аналогичные технологии в своих грузовых терминалах используют такие ведущие аэропорты, как аэропорт Хартсфилд-Джексон в Атланте, который применяет автоматизированные системы сортировки и робототехнику для обработки грузов, и аэропорт Чанги в Сингапуре, где реализованы интеллектуальные системы управления логистикой и автономные транспортные средства. Эти аэропорты, наряду с другими

новаторами в области, демонстрируют эффективность и преимущества использования таких технологий для повышения скорости, точности и безопасности операций.

Внедрение подобных технологий позволяет аэропортам значительно повысить эффективность операций, сокращая время обработки грузов и минимизируя риск ошибок, связанных с человеческим фактором. Это приводит к улучшению управления ресурсами, снижению эксплуатационных расходов и повышению пропускной способности. Аэропорты также выигрывают от улучшенной безопасности операций и способности быстро адаптироваться к изменениям в объемах грузопотока или рыночных условиях, что в конечном итоге повышает удовлетворенность клиентов и конкурентоспособность.

Внедрение системы прогнозирующего аналитического моделирования в грузовом терминале аэропорта Нарита привело к значительному улучшению оперативной эффективности и точности прогнозирования грузопотоков. Система позволила более точно распределять ресурсы и планировать операции, что сократило время обработки грузов и минимизировало задержки. В результате оптимизированного управления цепочкой поставок повысилось доверие клиентов и увеличилась пропускная способность терминала, что способствовало укреплению конкурентных позиций аэропорта на глобальной арене.

Кроме того, система значительно улучшила способность аэропорта адаптироваться к изменениям в рыночных условиях и колебаниям спроса, что позволило поддерживать высокий уровень обслуживания даже в периоды пиковых нагрузок. Применение передовых аналитических инструментов и технологий искусственного интеллекта также способствовало улучшению качества данных, используемых для стратегического планирования и принятия решений, что в перспективе укрепляет позиции аэропорта как инновационного центра воздушных грузоперевозок.

Для улучшения удовлетворенности клиентов аэропорт Нарита использовал методы, основанные на применении прогнозирующего аналитического моделирования для оптимизации операций. В первую очередь, система позволила сократить время обработки грузов благодаря точному распределению ресурсов и предсказанию пиковых нагрузок. Это обеспечило более стабильные и надежные услуги, минимизировав задержки и улучшив точность поставок. Кроме того, интеграция современных информационных технологий и систем отслеживания в реальном времени дала клиентам доступ к более прозрачной и оперативной информации о статусе их грузов, что укрепило доверие и удовлетворенность.

Чтобы оценить эффективность внедрения системы SAP Transportation Management – Starter Package, необходимо учитывать такие показатели, как объем и затраты на грузоперевозки.

Средняя стоимость за перевозку тонны груза составляет примерно 1000 долларов США, это 508 119 тенге по нынешнему курсу. Учитывая данные показатели, имеем следующую формулу:

$$P = y * 1000$, \tag{2}$$

где  $P$  – средняя стоимость перевозки,  $y$  – количество перевезенного груза, 1000\$ - тариф на перевезенную тонну груза.

### **Результаты и их обсуждение.**

В виду незавершенности статистики указанного года (2024), имеются данные в периоде с января по сентябрь месяц включительно и будут использованы в сравнении только с месячными показателями, минуя годовой оборот во избежание некорректных значений.

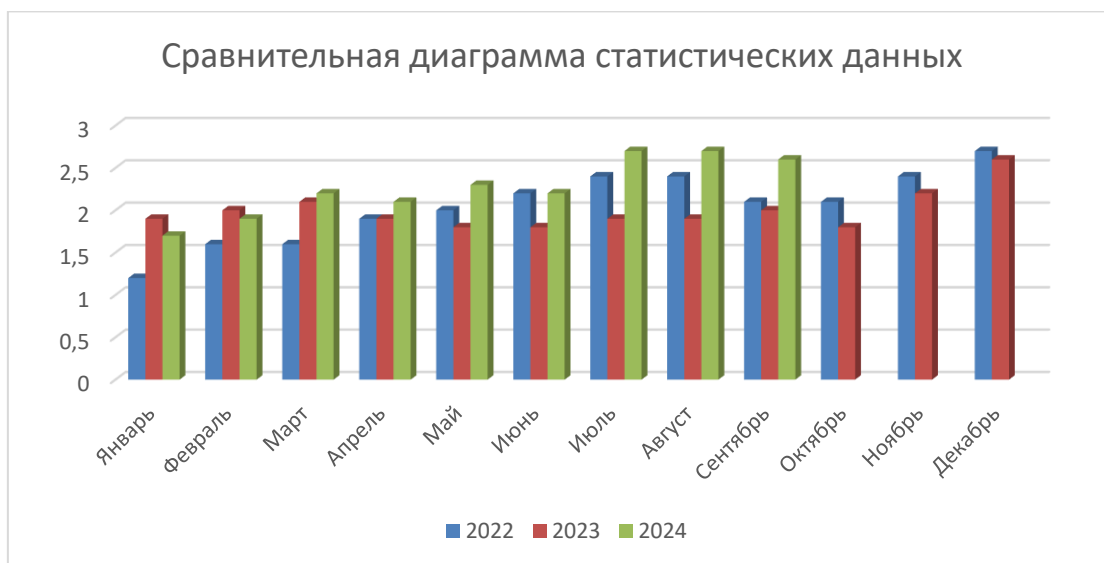


Рисунок 1 – Сравнительная диаграмма статистических данных

Таблица 4 – Темп прироста (Т) годовых объемов грузовых перевозок (%)

Год	Янв.	Фев.	Март	Апр.	Май	Июнь	Июль	Авг.	Сент.	Окт.	Ноя.	Дек.
2022	-52	0	-38	-44	-47	-24	-25	-7	-16	-25	-14	-20
2023	+59	+23	+33	+1	-2	-17	-20	-19	-6	-14	-7	-3
2024	-7	-2	+4	+10	+28	+21	+42	+38	+29	X	X	X

Как видно на таблице 4, скачки 2023 года в объеме грузоперевозок стабилизировались в нынешнем 2024 году, показывая положительный прогресс, равный в среднем на 18,1%

Таблица 5 – Темп прироста (Т) перевозок грузов и багажа воздушным транспортом (тыс. тонн) (Статистические значения за 2024 год внесены в периоде с января по сентябрь месяц включительно)

Год	Год	Год	Год	Статистика	2022-2023	2023-2024
2021	2022	2023	2024	2021-2022		
34	24,5	23,8	*20,3	-27,94	-2,86%	Недостаток данных в связи с временным продолжением



Рисунок 2 – Сравнительная диаграмма перевозок грузов (Статистические значения за 2024 год внесены в периоде с января по сентябрь месяц включительно)

Колебания объемов грузовых перевозок в аэропорту Алматы в последние годы могли быть вызваны рядом факторов, связанных как с глобальными, так и с региональными изменениями. Рассмотрим основные из них:

*Экономическая ситуация в регионе и мире.* Экономические кризисы или подъемы могут существенно воздействовать на объемы грузоперевозок. Улучшение экономической ситуации стимулирует международную торговлю, что приводит к увеличению объемов грузоперевозок. Наоборот, экономические спады вызывают снижение спроса на перевозку товаров.

*Торговые отношения и санкции.* Изменения в международной торговле, включая введение новых тарифов или санкций, могут существенно повлиять на объем перевозимых грузов. Для Алматы, как узлового пункта для Центральной Азии, изменения в торговых отношениях с крупными партнерами могут иметь значительное влияние.

*Инфраструктура и технологические обновления аэропорта.* Инвестиции в модернизацию аэропорта могут повысить его привлекательность для грузоперевозчиков, увеличив тем самым объемы. Внедрение новых технологий для более быстрой и эффективной обработки грузов также играет решающую роль.

*Изменения в логистических цепочках.* Пандемия COVID-19 продемонстрировала, как быстро могут изменяться логистические цепочки. Сдвиги в маршрутизации грузов или изменение предпочтений в пользу более быстрых воздушных перевозок могут повлиять на объемы, поступающие через аэропорт.

*Конкуренция с другими транспортными узлами.* Аэропорт Алматы конкурирует с другими аэропортами в регионе. Более выгодные условия для грузоперевозчиков в соседних странах могут снизить поток грузов через Алматы.

*Политические и социальные факторы.* Войны, конфликты или политическая нестабильность могут как положительно, так и отрицательно влиять на объемы перевозок, вызывая изменения в привычных маршрутах и потоках грузов.

Таблица 7 – Темп прироста (Т) доходов предприятий воздушным транспортом (млрд. тг)

Год 2021	Год 2022	Год 2023	Статистика	
			2021-2022	2022-2023
19.6	20.9	53.9	+6.63%	+157.89%

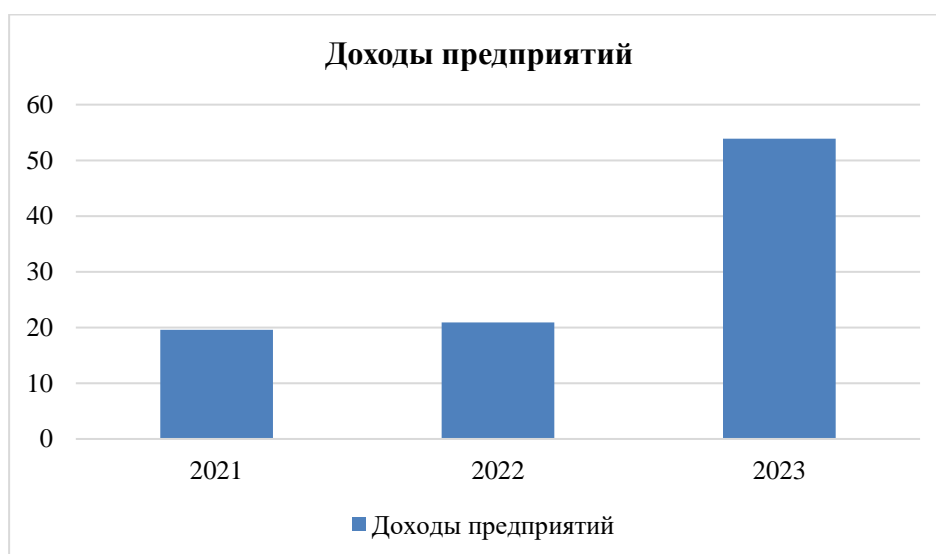


Рисунок 3 – Сравнительная диаграмма доходов предприятий

Таблица 8 – Сравнительная таблица грузооборота (тыс. тонн) [4, с.72]

№	Страна	2021	2022	%	2023	%
1	Российская Федерация	998.350	599.000	-40%	467.373	-21,97%
2	Узбекистан	90.092	101.000	+12%	139.700	+38%
3	Казахстан	34.000	24.500	-27,94%	23.800	-2,86

Ожидаемое внедрение системы SAP ТМ колеблется между 10-20% за счет оптимизации маршрутов. Она повышает точность планирования и улучшения взаимодействия с партнерами. Рассмотрим оба сценария:

При экономии 10%:

$$28\,300\,000 \times 0,10 \equiv 2\,830\,000 \text{ (долларов США)}, \quad (3)$$

$$14\,379\,795\,858 \times 0,10 \equiv 1\,437\,979\,585 \text{ (тенге)},$$

При экономии 20%:

$$28\,300\,000 \times 0,20 \equiv 5\,660\,000 \text{ (долларов США)}, \quad (4)$$

$$14\,379\,795\,858 \times 0,20 \equiv 2\,875\,959\,171 \text{ (тенге)},$$

где 28 300 000 – сумма доходов доставки грузов, 0,20/0,10 – 10/20% экономии.

Учитывая всевозможные нюансы, предположим, что первоначальные взносы на внедрение системы составляют 1 000 000 долларов США, это 508 119 995.

Ожидаемый срок окупаемости системы при 10% экономии составит:

$$\frac{1\,000\,000}{2\,830\,000} = 0,35 = 3,5 \text{ месяца (долларов)}, \quad (5)$$

$$\frac{508\,119\,995}{1\,437\,979\,585} = 0,35 = 3,5 \text{ месяца (тенге)}.$$

Ожидаемый срок окупаемости системы при 20% экономии составит:

$$\frac{1\,000\,000}{5\,660\,000} = 0,18 = 1,8 \text{ месяца (долларов)}, \quad (6)$$

$$\frac{508\,119\,995}{2\,875\,959\,171} = 0,18 = 1,8 \text{ месяца (тенге)}.$$

Таким образом, внедрение системы может привести к значительной экономии и быстрой окупаемости, что делает ее привлекательным вариантом для аэропорта Алматы в сфере грузоперевозок.

### **Заключение.**

Грузовые перевозки, в частности авиагрузоперевозки, играют ключевую роль в экономическом развитии страны, способствуя укреплению международных торговых связей, ускорению логистических процессов и повышению качества жизни населения. Казахстан, обладая выгодным географическим положением и высоким транзитным потенциалом, имеет возможность развивать аэропорт Алматы как международный грузовой хаб. Однако высокая стоимость аэропортовых услуг может ограничить конкурентоспособность, что требует пересмотра тарифной политики и увеличения доходов за счет неавиационной деятельности.

Анализ мирового опыта и работы аэропортов стран-соседей показывает, что модернизация инфраструктуры, внедрение инновационных технологий и

автоматизированных логистических систем являются необходимыми мерами для повышения эффективности грузоперевозок. Современные системы отслеживания, контроль качества и цифровизация логистических процессов позволят улучшить уровень сервиса, ускорить грузопотоки и повысить привлекательность аэропорта для международных партнеров.

Таким образом, развитие грузовых авиаперевозок в Казахстане требует комплексного подхода, включающего улучшение инфраструктуры, оптимизацию тарифной политики и внедрение передовых технологий. Это позволит не только укрепить позиции страны на мировом рынке авиагрузоперевозок, но и внести значительный вклад в устойчивое экономическое развитие.

#### Список литературы

1. Международный аэропорт Алматы <https://alaport.com/ru-RU/business-ru/cargo-terminal-business-ru/page/cargo-terminal-almaty-ru>
2. Airports Council International (ACI) – ACI World Report <https://aci.aero>
3. "Innovative Technologies in Narita Cargo Terminal" 2023
4. Международный исследовательский журнал DOGMA Выпуск №24 Макогонова В.О. «Показатели и факторы, влияющие на объем авиационных перевозок» стр.72
5. Christopher, M. (2016). *Logistics & Supply Chain Management*, 5th ed. Pearson,
6. Chopra, S. & Meindl, P. (2016). *Supply Chain Management: Strategy, Planning, and Operation*, 6th ed. Pearson,
7. Rushton, A., Croucher, P. & Baker, P. (2014). *The Handbook of Logistics and Distribution Management: Understanding the Supply Chain*, 5th ed. Kogan Page.
8. Kotler, P. & Keller, K. L. (2016). *Marketing Management*, 15th ed. Pearson.
9. Porter, M. (2008). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press.

#### References

1. Mezhdunarodnyj ajeoport Almaty <https://alaport.com/ru-RU/business-ru/cargo-terminal-business-ru/page/cargo-terminal-almaty-ru>
2. Airports Council International (ACI) – ACI World Report <https://aci.aero>
3. "Innovative Technologies in Narita Cargo Terminal" 2023
4. Mezhdunarodnyj issledovatel'skij zhurnal DOGMA Vypusk №24 Makogonova V.O. «Pokazateli i faktory, vlijajushhie na ob#em aviacionnyh perevozok» str.72
5. Christopher, M. (2016). *Logistics & Supply Chain Management*, 5th ed. Pearson,
6. Chopra, S. & Meindl, P. (2016). *Supply Chain Management: Strategy, Planning, and Operation*, 6th ed. Pearson,
7. Rushton, A., Croucher, P. & Baker, P. (2014). *The Handbook of Logistics and Distribution Management: Understanding the Supply Chain*, 5th ed. Kogan Page.
8. P. Kotler and K. L. Keller, *Marketing Management*, 15th ed. Pearson, 2016.
9. Porter, M. (2008). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press.

### ЖҮК ТАСЫМАЛДАРЫНЫҢ ТИІМДІЛІГІН АРТТЫРУ ҮШІН МАРКЕТИНГТІК ЗЕРТТЕУ ӘДІСТЕМЕСІН ӘЗІРЛЕУ

**Аңдатпа.** Қазіргі жағдайда, технологияның дамуына қарамастан, компаниялар жүк тасымалын оңтайландыру мәселелерімен бетпе-бет келуде. Шешімдердің бірі-логистиканы басқарудың автоматтандырылған жүйелерін енгізу, бірақ бұл процесс киберқауіптерді, нарықтық тәуекелдерді және айтарлықтай қаржылық инвестициялардың қажеттілігін қоса алғанда, бірқатар қиындықтарды қамтиды. Әуе



жүктері халықаралық деңгейде жылдам және тиімді жеткізуді қамтамасыз ете отырып, тасымалдаудың ең сұранысқа ие тәсілдерінің бірі болып қала береді. Стратегиялық тиімді географиялық жағдайға ие Қазақстан жоғары транзиттік әлеуетке ие. Перспективалы бағыттардың бірі Алматы әуежайын халықаралық жүк хабы ретінде дамыту болып табылады, бұл жүк ағынын ұлғайтуға және елдің жаһандық саудадағы позициясын нығайтуға мүмкіндік береді. Дегенмен, әуежай қызметтерінің жоғары құны тарифтік саясатты қайта қарауды және авиациялық емес қызмет арқылы кірісті арттыруды талап ететін тежегіш болуы мүмкін. Әуежай инфрақұрылымын жаңғырту, SAP Transportation Management сияқты инновациялық технологиялар мен цифрлық шешімдерді енгізу логистикалық процестердің дәлдігі мен тиімділігін арттыруға мүмкіндік береді. Бұл сервис деңгейін жақсартып қана қоймай, Қазақстанның жүк авиатасымалдары саласындағы бәсекелестік позициясын күшейтіп, тұрақты экономикалық дамуға және әлемдік көлік жүйесіне кірізуге ықпал етеді.

**Түйінді сөздер:** жүк тасымалы, авиакомпания, тиімділікті арттыру, SAP Transportation Management, әуежай, маркетинг.

## DEVELOPMENT OF MARKETING RESEARCH METHODOLOGY TO IMPROVE THE EFFICIENCY OF FREIGHT TRANSPORTATION

**Abstract.** In modern conditions, despite the development of technology, companies continue to face problems of optimizing cargo transportation. One solution is to implement automated logistics management systems, but this process involves a number of challenges, including cyber threats, market risks, and the need for significant financial investments. Air cargo transportation remains one of the most demanded transportation methods, providing fast and efficient delivery at the international level. Kazakhstan, having a strategically advantageous geographical location, has a high transit potential. One of the promising areas is the development of Almaty Airport as an international cargo hub, which will increase cargo flow and strengthen the country's position in global trade. However, the high cost of airport services can become a deterrent, which requires a revision of the tariff policy and an increase in revenue due to non-aviation activities. The modernization of the airport's infrastructure, the introduction of innovative technologies and digital solutions such as SAP Transportation Management will improve the accuracy and efficiency of logistics processes. This will not only improve the level of service, but also strengthen Kazakhstan's competitive position in the field of air cargo transportation, contributing to sustainable economic development and integration into the global transport system.

**Keywords:** cargo transportation, aviation company, efficiency improvement, SAP Transportation Management, airport, marketing.

### Сведение об авторах

Макогонова Виктория Олеговна	Магистр, методист, Академия гражданской авиации, г. Алматы Казахстан, E-mail: <a href="mailto:cherry-98-23-04@mail.ru">cherry-98-23-04@mail.ru</a>
Асильбекова Индира Жаксыбаевна	К.т.н., профессор, Академия гражданской авиации, г. Алматы Казахстан, E-mail: <a href="mailto:a.indira@mail.ru">a.indira@mail.ru</a>

### Авторлар туралы мәлімет

Макогонова Виктория Олеговна	Магистр, әдіскер, Азаматтық авиация академиясы, Алматы қ. Қазақстан, E-mail: <a href="mailto:cherry-98-23-04@mail.ru">cherry-98-23-04@mail.ru</a>
------------------------------	---

Асильбекова Индира Жаксыбаевна	Т.ғ.к., Азаматтық авиация академиясы, Алматы қ. Қазақстан E-mail: <a href="mailto:a.indira@mail.ru">a.indira@mail.ru</a>
--------------------------------------	---

**Information about the authors**

Makogonova Viktoriya Olegovna	Mass.ter, methodologist, Academy of Civil Aviation, Almaty, Kazakhstan E-t.mail: cherry-98-23-04@mail.ru
Assilbekova Indira Zhaksybaevna	C.T.S., Professor, Academy of Civil Aviation, Almaty, Kazakhstan E-mail: a.indira@mail.ru

---

---

**КОМПЬЮТЕРЛІК ҒЫЛЫМДАР, АСПАП ЖАСАУ ЖӘНЕ АВТОМАТТАНДЫРУ**  
**КОМПЬЮТЕРНЫЕ НАУКИ, ПРИБОРОСТРОЕНИЯ И АВТОМАТИЗАЦИЯ**  
**COMPUTER SCIENCE, INSTRUMENTATION AND AUTOMATION**

---

---

ӘОЖ 004.85

ҒТАХА 28.23.25

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_7](https://doi.org/10.53364/24138614_2025_36_1_7)**Н. Алимбекова<sup>1,2</sup>, Х.М. Рай<sup>3</sup>, Т. Турымбетов<sup>4</sup>, А. Жумадилаева<sup>1\*</sup>**<sup>1</sup>Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан<sup>2</sup>Астана Халықаралық университеті, Астана, Қазақстан<sup>3</sup>Жахон университеті, Соннам-си, Оңтүстік Корея<sup>4</sup>Халықаралық Туризм және Меймандостық Университеті, Түркістан, Қазақстан*E-mail:* [ay8222@mail.ru](mailto:ay8222@mail.ru)\***ПОЛИМЕРЛІ КОНТЕЙНЕРЛЕРДІ СҰРЫПТАУ ПРОЦЕСІН**  
**ОҢТАЙЛАНДЫРУ ҮШІН НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ҚОЛДАНУ**

*Аңдатпа.* Бұл зерттеуде пластикалық контейнерлерді тиімді сұрыптау үшін конволюционды нейрондық желілерді (CNN) және ұзақ қысқа мерзімді жады (LSTM) біріктіретін гибриді нейрондық желі архитектурасын пайдалануды қарастырады. Зерттеу жақын инфрақызыл (NIR) спектроскопиялық құрылғысымен алынған, химиялық құрамы мен ластану деңгейіне байланысты пластикалық қалдықтарды жіктеуге бағытталған. Эксперимент нәтижелері CNN+LSTM гибриді моделі пластиктердің әртүрлі түрлері мен түстерін тану, соның ішінде контейнерлердегі ластанушы заттарды анықтауда салыстырмалы түрде жоғары дәлдікке қол жеткізетінін көрсетеді. Модельдің өнімділігін салыстырмалы бағалау логистикалық регрессия, ішінара ең кіші квадраттар (PLS) және сызықтық дискриминантты талдау (LDA) сияқты дәстүрлі жіктеу әдістерімен жүргізілді. Нәтижелер CNN+LSTM үлгісі дәстүрлі тәсілдерге қарағанда, әсіресе класстар арасындағы спектрлік айырмашылықтары аз сценарийлерде тиімдірек жұмыс істейтінін көрсетеді. Бұл зерттеу пластикалық қалдықтарды сұрыптау және қайта өңдеу процестерінің тиімділігін арттыру үшін машиналық оқытудың әлеуетін көрсетеді, осылайша экологиялық тұрақтылықты жақсартуға ықпал етеді.

*Түйін сөздер:* Пластикалық қалдықтар, NIRS, Нейрондық желі, Гибриды модель, CNN, LSTM.

**Кіріспе.**

Дүниежүзілік пластикалық қаптама нарығының көлемі 2024 жылы 103,48 миллион тоннаға бағаланады және 2029 жылға қарай 123,26 миллион тоннаға жетеді деп күтілуде, бұл ретте болжамды кезең ішінде (2024-2029 жылдар) орташа жылдық өсу қарқыны 3,56% құрайды. Пластикалық полимер қаптамаға сұраныс тұрақты өсуде, өйткені ол жеңіл және қолдануға ыңғайлы. Пластик өндіріс шығындарының төмендігіне байланысты тұтынушылар мен өндірушілер үшін де қолайлы. Полимерлі контейнерлердің сусындар, тамақ, косметика және фармацевтика салаларында маңыздылығы артып келеді (Mordor Intelligence, 2024).

Пластикалық қаптама өндірісі көлемінің артуы пластикалық қалдықтардың өсуіне әкеліп, оларды қайта өңдеудің әдістерін дамыту қажеттілігін арттырады, бұл қоршаған ортаға теріс әсерді болдырмау үшін маңызды. Полимер пластиктерді қайта өңдеу қоршаған ортаға теріс әсерді азайтудың маңызды шешімі ретінде қарастырамыз, алайда қолданыстағы қайта өңдеу инфрақұрылымы пластиктерді сұрыптау және жіктеуде бірқатар қиындықтарға тап болады. Пластиктердің қайта өңдеу барысында дұрыс жіктелмеуі немесе олардың ластануы екінші реттік материалдың сапасы мен қайта қолдану мүмкіндігін едәуір төмендетіп, экологиялық мәселелерді одан әрі күшейтеді.

Қайта өңдеу процестері ең алдымен пластиктің әртүрлі түрлерін ажырату үшін сенсорлар мен химиялық және физикалық қасиеттерді талдау сияқты озық әдістерді қолданады (Bonifazi және т.б., 2018). Дәстүрлі әдістер мысалы ретінде батыру-қалқымалы бөлу, шөктіру, гидроциклондау (Gent және т.б., 2018), электростатикалық бөлу және магниттік тығыздықты бөлу (Zhang және т.б., 2023) сияқты сұрыптау үшін полимерлердің физикалық қасиеттеріне сүйенеді. Дегенмен, бұл тәсілдер аралас пластиктерді немесе тығыздығы ұқсас пластмассаларды дәл сұрыптамауы мүмкін (Ragaert және т.б., 2017; Gundupalli және т.б., 2017). Мәселенің шешімі ретінде сенсорлық әдістер алдыңғы әдістерге қарағанда пластиктердің дәлірек және тиімді жіктелуін ұсынады.

Лазерлік индукцияланған ыдырау спектроскопиясы (LIBS), рентгендік сұрыптау және RGB камерасын қолдану арқылы сұрыптау сияқты әдістер кең таралған, бірақ олардың әрқайсысының шектеулері бар. LIBS әдісі пластикті анықтау кезінде сезімталдығының төмендігіне байланысты қиындықтарға тап болады (Shameem және т.б., 2017), рентген арқылы сұрыптау материалдарды ажыратуда шектеулі (Brunner және т.б., 2015), ал RGB камерасымен сұрыптау полимер пигменттерінің ұқсастығына және жарықтандыру мен ластану әсеріне ұшырауына байланысты нақты дәлдік қамтамасыз етпейді (Safavi және т.б., 2010).

Жақын инфрақызыл (NIR) спектроскопия және гиперспектральды бейнелеу жүйелерін пайдалана отырып, пластикалық қалдықтарды әртүрлі санаттарға жіктеуді автоматтандыру мақсатындағы көптеген зерттеулер бар (Qiu және т.б., 2020; Kumar және т.б., 2021; Tap және т.б., 2021; Ahmed және т.б., 2023; Singh және т.б., 2023). NIR спектроскопия арқылы бірегей спектрлік сипаттамаларына байланысты әртүрлі материалдарды, соның ішінде пластикті анықтауға арналған құрал ретінде бірқатар зерттеулер дәлелдеді (Masoumi және т.б., 2012; Marchesi және т.б., 2023). Белгілі бір толқын ұзындықтарында жарықтың жұтылуы, шағылысу және өткізгіштік қасиеттерін өлшеу арқылы NIR спектроскопиясы материалдың молекулалық құрамын анықтай алады, бұл пластик түрлерін ажыратуға мүмкіндік береді. Дегенмен, пластикалық қалдықтардың нақты деректерге негізделген жіктелуі ластаушы заттардың болуына және нақты өмірлік жағдайларда пластикалық қалдықтардың әртүрлі жағдайларына байланысты қиын болып қала береді.

Осы қиындықтарға жауап ретінде бұл зерттеуде пластикалық полимерлі контейнерлерді жіктеудің сенімді жүйесін әзірлеу үшін NIR-спектроскопиясымен алынған мәліметтермен нейрондық желілерді қолдану арқылы интеграциялау ұсынылады. Машиналық оқыту алгоритмдерінің жиынтығы болып табылатын нейрондық желілер үлкен және күрделі деректер жиынтығымен жұмыс істеуге өте ыңғайлы және кескінді тану, табиғи тілді өңдеу және соңғы уақытта материалдарды жіктеу мәселелерінде айтарлықтай жетістіктерге қол жеткізді (Xia және т.б., 2021; Olowolayemo және т.б., 2022).

Алдыңғы зерттеулердің басты мақсаты осы технологияларды қолдана отырып, жалпы қалдықтар ағынынан пластикті сұрыптау немесе оны полимер түріне қарай жіктеу болды. Алайда, бұл жұмыста пластик түрлерін тек материалдың түрі бойынша ғана емес, сонымен қатар түсі мен ластану деңгейіне қарай анықтау қарастырылады. Мысалы, бұл химиялық өнімдерге (шампуньдар мен тазартқыштар сияқты) пайдаланылған бөтелкелерді және қайта

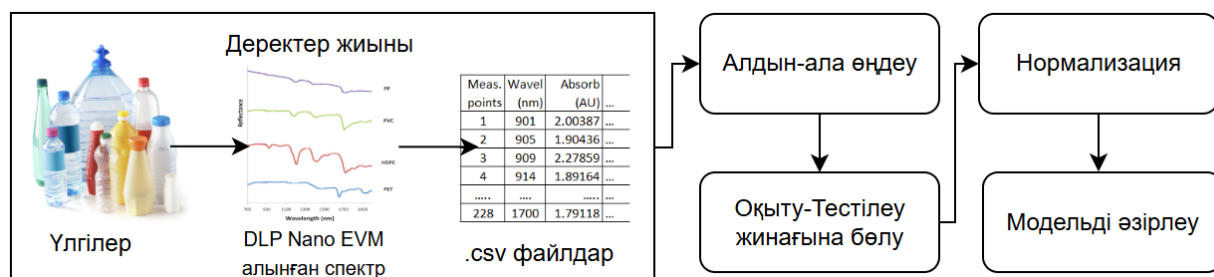
өңдеуге жатпайтын сүт өнімдеріне арналған контейнерлерді ажыратуды қамтиды. Мұндай қалдықтарды қайта өңдеу мүмкіндігі болғанымен, олардың көлемі әдетте аз, ал қайта өңдеу шығындары өте жоғары, сондықтан таза ПЭТ бөтелкелерін жинау және қайта өңдеу артықшылыққа ие болады.

Нейрондық желі моделін NIR спектрометрімен алынған сандық деректеріне үйрете отырып, бұл зерттеу ластану жағдайында да пластиктің жіктелу дәлдігін арттыруға тырысады, бұл сайып келгенде тиімдірек қайта өңдеу процестеріне және пластиктің ластану деңгейін төмендетуге ықпал етеді.

**Материалдар мен әдістер.**

Бұл зерттеу жақын инфрақызыл (NIR) спектроскопиясы мен машиналық оқытуды қолдана отырып, пластикалық қалдықтарды жіктеуге бағытталған деректерді жинау, өңдеу және талдау әдістерін сипаттайды.

Зерттеу бірнеше кезеңдерді қамтиды: спектрлік деректерді жинау және талдау, қалыпқа келтіру, оқыту және сынақ жиынтықтарына бөлуді қоса алғанда, деректерді алдын ала өңдеу, пластикалық контейнерлерді жіктеуге арналған моделді әзірлеу және оқыту (сурет-1).



1 сурет – Әзірленген әдістің жалпы көрінісі

*Үлгілерді жинақтау.*

Зерттеуге арналған мәліметтер жиынтығы үш кезеңге бөлінген, пластикалық бөтелкелерді материал түріне, мөлдірлігіне және ластану деңгейіне қарай жіктейді, деректердің 80% оқыту үшін және 20% сынақ үшін пайдаланылды.

Үлгілер DLP NIRScan Nano EVM (Texas Instruments, АҚШ), 900-ден 1700 нм-ге дейінгі 228 диапазондағы инфрақызыл спектрлерді анықтауға қабілетті жоғары интеграцияланған ультра портативті инфрақызыл спектрометрімен алынды. Төмендегі суретте үлгілерді түсіру барысы, бұл құрылғымен алынған спектрлар ашық немесе жабық кеңістікте айтарлықтай өзгеріс болмағанын көрсетті. Дегенмен бүкіл үлгілер жабық, қараңғы кеңістікте алынды.



а) DLP NIRScan Nano EVM портативті құрылғысы



б) үлгіні ашық кеңістікте түсіру



в) үлгіні жабық кеңістікте түсіру

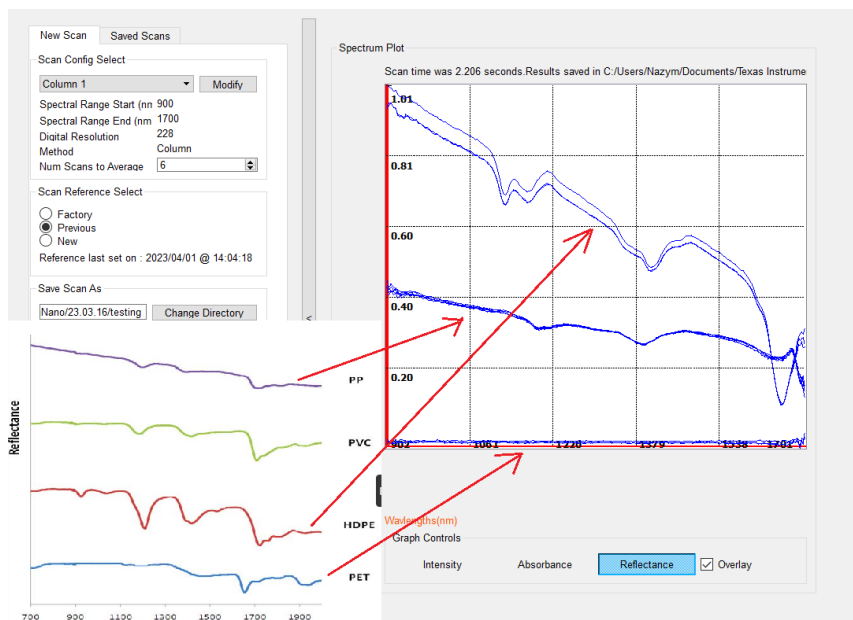
2 сурет – Үлгілерді түсіру барысы

Бірінші кезеңде әр түрлі маркалы, таза-лас, 430 PET (полиэтилентерефталат) бөтелкелері мен контейнерлері, сондай-ақ тұрмыстық химиялық контейнерлерден PP (полипропилен), HDPE (тығыздығы жоғары полиэтилен) және LDPE (тығыздығы төмен полиэтилен) сияқты басқа пластиктердің 204 түрі бойынша жіктеу. Екінші кезеңде пластикалық контейнерлердің түсі бойынша жіктеледі, яғни 236 мөлдір бөтелке (таза немесе ластанған PET, PP болуы мүмкін) және 48 әртүлі түсті бөтелкелер (қанық немесе жартылай мөлдір болуы мүмкін). Үшінші кезең әртүрлі брендтердің 174 таза, бос PET бөтелкелері мен 209 сүт, май, соустар және зиянды сұйықтықтар сияқты өнімдерді сақтау үшін пайдаланылатын PET ластанған бөтелкелерді ажыратады. Бұл құрылымдық дерекқор пластикалық бөтелкелердің материал сапасы мен олардың ластану деңгейін егжей-тегжейлі зерттеуге мүмкіндік береді, бұл тиімді қайта өңдеу процестері үшін маңызды (Кесте-1).

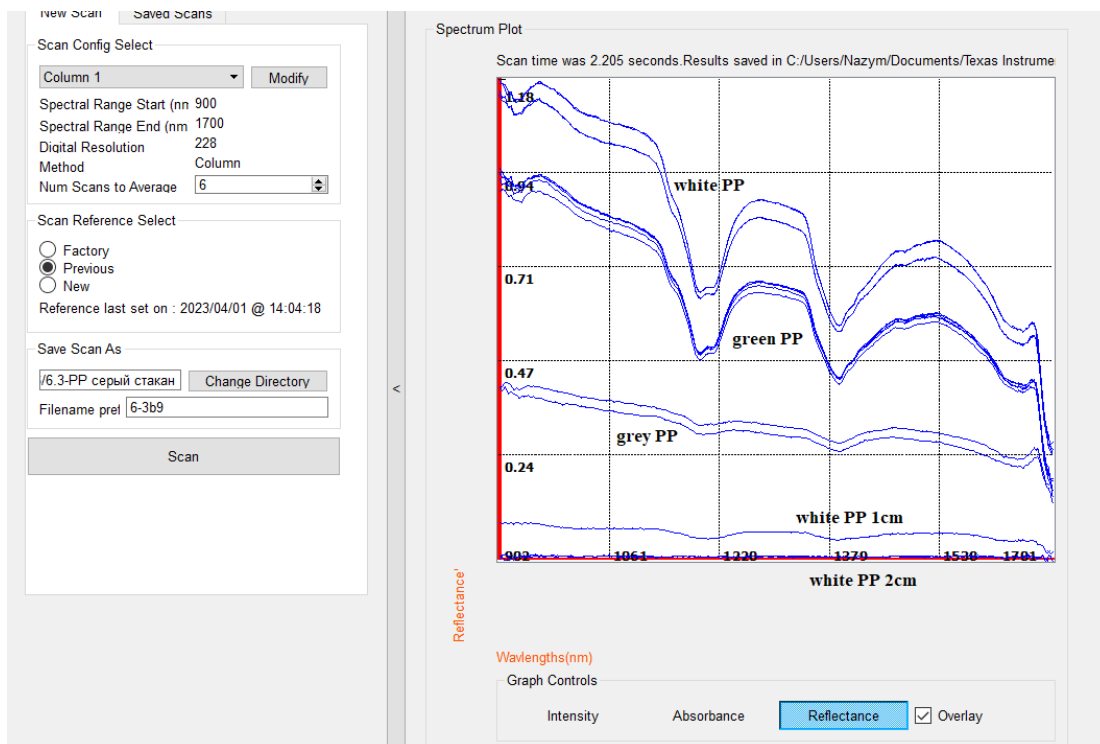
1 кесте – Класстар бойынша құрылған үлгілер саны

Кезең	Класс	Саны
Кезең 1	ПЭТ	430
	Басқа (PP, LDPE, HDPE)	204
Кезең 2	Мөлдір	236
	Түрлі түсті (қанық)	48
Кезең 3	Таза	174
	Ластанған	209

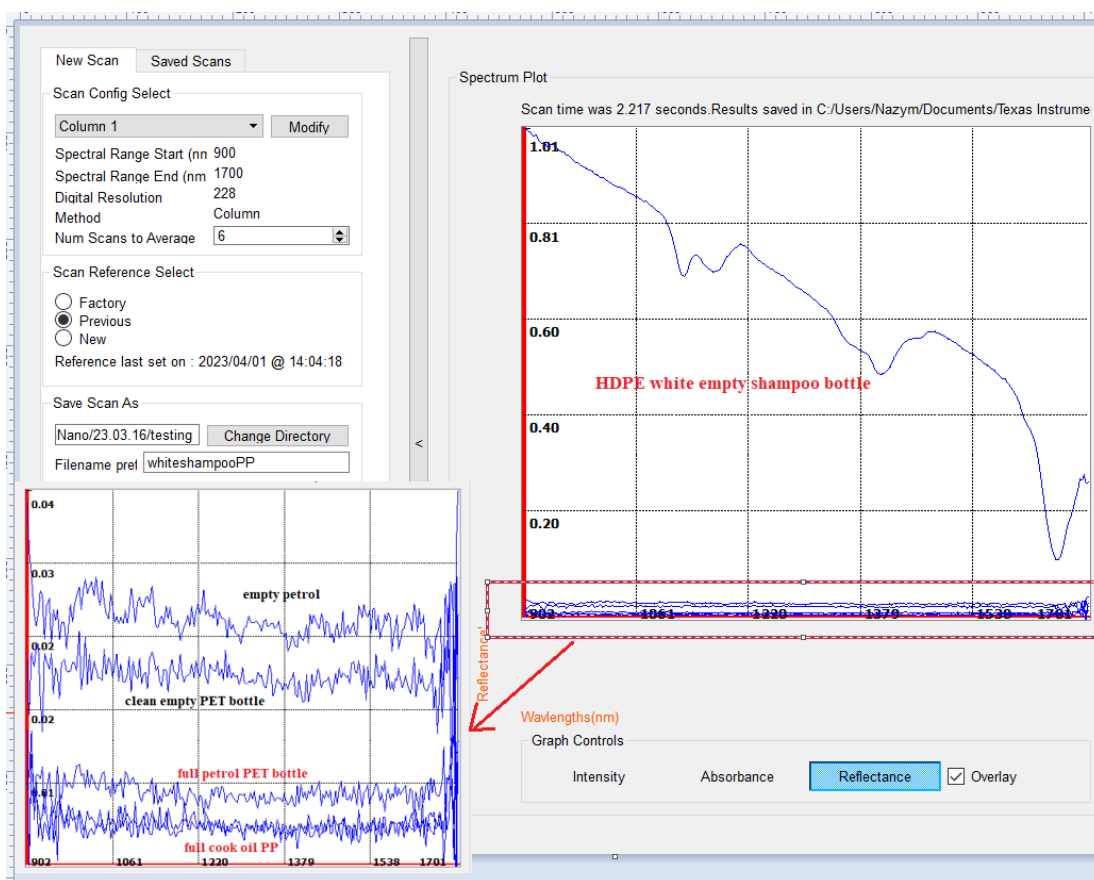
IRscanNano GUI v1.1.9 интерфейсі арқылы бейнеленген өңделмеген (raw data) үлгілердің шағылысу коэффициенті бойынша визуалды көрінісі. (сурет 3-5)



3 сурет – 1 кезең пластикалық бөтелкелерді типі бойынша салыстыру (Masoumi, H., т.б автордың мақаласынан алынды [16])



4 сурет – 2 кезең пластикалық бөтелкелерді түстері бойынша ажырату



5 сурет – 3 кезең пластикалық бөтелкелерді ластану деңгейі бойынша ажырату

Үлгілерді алдын-ала өңдеу кезеңінде, нейрондық желі олардың NIR спектрлік қолтаңбалары негізінде пластиктердің әртүрлі түрлерін ажыратуды тиімді үйренуін

қамтамасыз ету үшін спектрлік деректерді ұйымдастыруды және таңбалауды қамтиды. Қарастырылып отырған деректер жиыны спектрометрден 634 көрсеткішін қамтиды, әрбір көрсеткіш төрт баған бойынша спектрлік деректердің 228 жолын қамтиды: толқын ұзындығы (нм), абсорбция (AU), анықтамалық сигнал (бірліксіз) және үлгі сигналы (бірліксіз).

Алдын ала өңдеудің бастапқы қадамы екілік жіктеудің үш кезеңнің әрқайсысы үшін бүкіл деректер жиынтығын қайталауды қамтиды: ПЭТ-ті басқа пластик түрлерінен ажырату, мөлдір пластиктерді түрлі-түстіден ажырату және ПЭТ ластанған материалдардан таза ПЭТ анықтау. Қайталау (duplication) маңызды, өйткені әрбір тапсырма деректерге ерекше көзқарасты талап етеді, яғни сол классификация міндетіне қатысты нақты атрибуттарға назар аударады. Бұл кезең нейрондық желіні оқыту үшін өте маңызды, өйткені ол әр топтың сипаттамаларын білу үшін қажетті таңбаланған деректерді береді.

Таңбалаудан кейін деректер жиындары талдауға қажетті бірегей идентификаторлар мен белгілерді қамтитын бір CSV файлына біріктіріледі. Осы біріктірілген деректер жиынындағы бағандарға мыналар кіреді: Толқын ұзындығы (нм), Абсорбция (AU), Анықтамалық сигнал (бірліксіз), Үлгі сигналы (бірліксіз), Вариация, Үлгі\_саны, Санат және Сенсор\_идентификаторы. Мұндағы "Вариация" үш екілік жіктеу кезеңнің қайсысына жататынын көрсетеді, "Үлгі\_саны" әрбір спектрлік көрсеткіш үшін бірегей идентификатор, "Санат" көрсеткіш тағайындалған екілік топты білдіреді және "Сенсор\_идентификаторы" көрсеткішті жазып алған спектрометрге сілтеме жасайды.

Алдын ала өңдеу фазасында, әсіресе нейрондық желіні талдау үшін спектрлік деректерді дайындау контекстінде қалыпқа келтіру (нормализация) маңызды қадам болып табылады. Бұл процесс айнымалы мәндердің сандық диапазонын реттеуді қамтиды, әрбір белгілер олардың масштабтарындағы тән дисперсияға байланысты үлгіні бұрмалаусыз талдауға бірдей үлес қосады. Бұған қол жеткізу үшін стандартты масштабтау деп аталатын әдіс қолданылады.

Стандартты масштабтау немесе Z-балды қалыпқа келтіру бұл контексте нормализация үшін таңдалатын әдіс болып табылады. Ол белгілерді орташа мәні 0 және стандартты ауытқуы 1 болатындай етіп түрлендіреді. Бұл әрбір белгінің орташа мәнін деректер жиынынан алып тастап, содан кейін нәтижені сол белгінің стандартты ауытқуына бөлу арқылы жүзеге асырылады.

Бақылаудың Z-балын есептеу формуласы келесідей:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

мұндағы  $x$  – бақылау мәні,  $\mu$  – белгінің орташа мәні, және  $\sigma$  – белгінің стандартты ауытқуы.

Қалыпқа келтіру процесі деректер жиынындағы сандық типтегі барлық белгілерге әсіресе, қалқымалы нүкте (float) немесе бүтін сан (integer) түрінде ұсынылғандарға қолданылады. Бұл стандартты масштабтау барлық белгілердің модельдің болжамдық мүмкіндігіне бірдей үлес қосуын қамтамасыз етеді, осылайша үлкен масштабтағы белгілердің модельдің нәтижесіне басым әсер етуін болдырмайды. Бұл әсіресе градиенттік түсуді (gradient descent) оңтайландыру үшін қолданатын модельдерде маңызды, себебі бұл нормализация модельдің жылдам және тиімді түрде жинақталуына (convergence) көмектеседі.

Деректерді қалыпқа келтіру арқылы нейрондық желіге кіріс деректерінен жақсырақ үйренуге және жалпы заңдылықтарды тануға мүмкіндік береді, бұл әсіресе материалдарды олардың спектрлік деректері бойынша классификациялау сияқты міндеттер үшін маңызды. Нормализация қадамы модельдің спектрлік деректерді тиімді өңдеуін және одан үйренуін қамтамасыз етеді, бұл пластиктің немесе талданатын басқа материалдардың әртүрлі түрлерін ажырату қабілетін арттырады.



Нейрондық желі - қабаттарға ұйымдастырылған көптеген өзара байланысты түйіндерден (нейрондардан) тұратын есептеу моделі. Нейрондық желінің негізгі мақсаты мысалдардан үйрену арқылы деректердегі күрделі тәуелділіктерді анықтау болып табылады. Нейрондық желілер сызықты емес функцияларды модельдеуге қабілетті, бұл оларды үлгіні тану, жіктеу және регрессия тапсырмалары үшін тиімді.

Конволюциялық нейрондық желілер (CNN)

CNN - конволюционды қабаттар арқылы деректердегі кеңістіктік иерархияларды автоматты түрде және бейімдеуге арналған терең оқыту архитектурасы. CNN кескінді өндеуде қолдануымен танымал, бірақ сонымен бірге бір өлшемді дәйекті деректерге қолданылады[19].

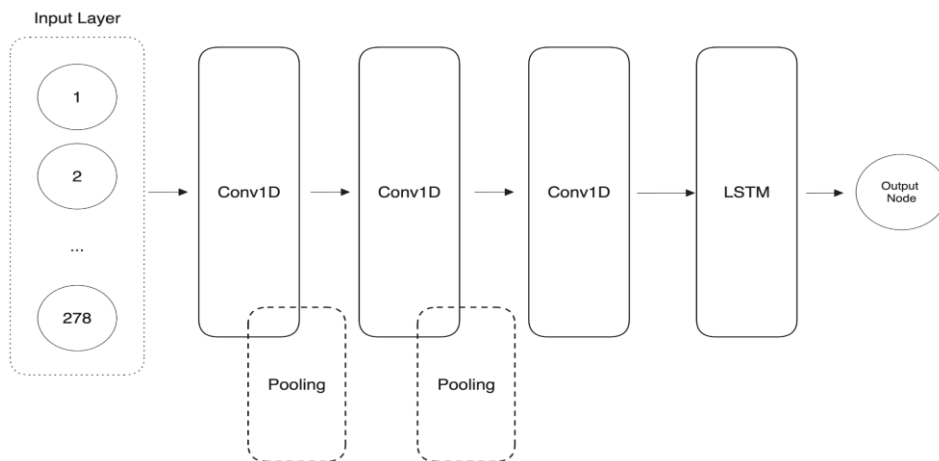
Ұзақ-қысқа мерзімді жад (LSTM)

LSTM — бұл қайталанатын нейрондық желінің (RNN) түрі, ол тізбекті деректерді өндеуге және қарапайым RNN-дерде жиі кездесетін жоғалып кететін градиенттер мәселесін шешуге арналған. LSTM-нің басты ерекшелігі — арнайы басқару элементтері (кіру, шығу және ұмыту "қақпалары"), олар маңызды ақпаратты ұзақ уақыт бойы тиімді сақтауға және жаңартуға мүмкіндік береді.

Бұл мақалада пластикалық қалдықтардың классификациясын жақсарту үшін CNN+LSTM гибриді алгоритм әдістерін қалыпқа келтіру және стандартты масштабтау сияқты деректерді алдын ала өндеу әдістерімен бірге пайдалану қарастырылды.

Конволюциялық нейрондық желілер (CNNs) және ұзақ-қысқа мерзімді жад (LSTM) желілерінің тіркесімі деректердегі кеңістіктік және уақытша үлгілерді пайдалануға арналған қуатты гибриді архитектураны білдіреді. CNN + LSTM архитектурасында LSTM конволюционды қабаттардан кейін қолданылады, мұнда ол CNN шығарған кеңістіктік мүмкіндіктерді дәйекті түрде өңдейді. Бұл құрылым модельге жергілікті кеңістіктік мүмкіндіктерді де, деректердегі уақытша тәуелділіктерді де түсіруге мүмкіндік береді [20].

Конволюциялық қабаттар жергілікті ерекшеліктерді тиімді түрде анықтайды, ал LSTM қабаттары деректер ішіндегі уақыттық тізбектерді түсінуде ерекше нәтиже көрсетеді, бұл классификацияның дәлдігін арттыруға әкеледі. Модель архитектурасына сәйкесінше 3 өлшемді ядросы және 32, 64 және 128 сүзгілер саны бар үш конволюциялық қабаттан тұрады. Әрбір конволюциялық қабат үшін ReLU активация функциясы қолданылды. Алғашқы екі конволюциялық қабаттан кейін деректердің өлшемін азайту үшін пул өлшемі 3 болатын максималды біріктіру (MaxPool) қабаттары қосылды. Содан кейін LSTM қабаты 64 бірлік және қайталанатын оқудан шығару (Dropout) 0.2 қолданылды. Модельдің шығыс қабаты (Output) бинарлы классификация үшін бір нейроннан және Sigmoid активация функциясы бар толық байланысқан қабатпен ұсынылған (сурет-6).



6 сурет – CNN + LSTM моделінің архитектурасы

*Өнімділік өлшемі*

Әрбір модельдің өнімділігі дәлдік (Accuracy), нақтылық (Precision), еске түсіру (Recall) және F1-есеп (F1-Score) сияқты стандартты метрикаларды пайдалану арқылы бағаланды. Бұл метрикалар модельдің пластикалық қалдықтарды жіктеу тапсырмасының әртүрлі кезеңдерінде классификациялау тиімділігін жан-жақты сипаттап береді.

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

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

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

$$F1 - Score = \frac{2 \times (Precision * Recall)}{(Precision + Recall)} \quad (5)$$

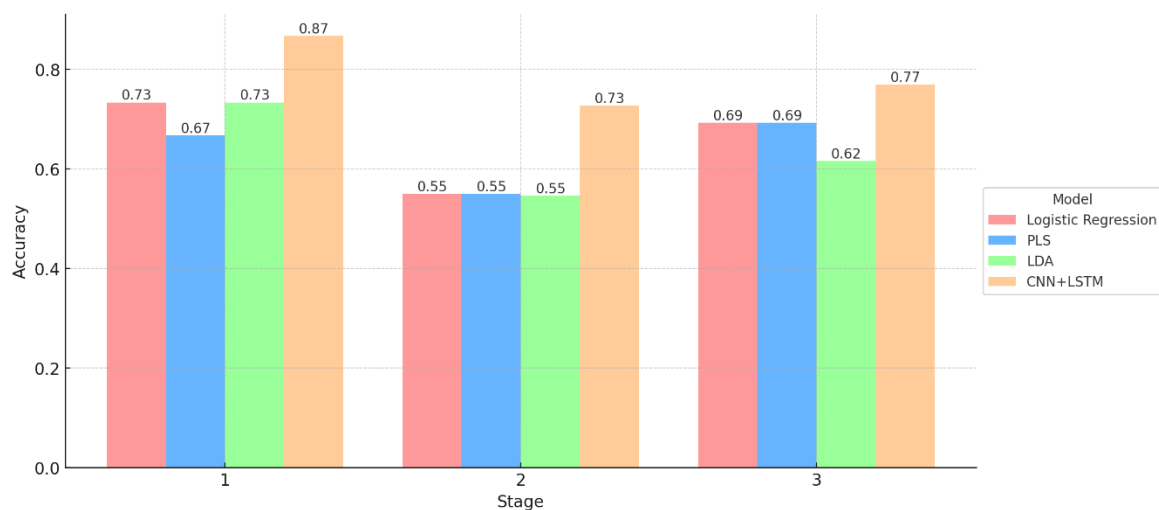
Әрбір модель үшін шатастыру матрицалары (confusion matrix) шынайы позитивтер (TP), шынайы негативтер (TN), жалған позитивтер (FP) және жалған негативтер (FN) тұрғысынан өнімділіктің егжей-тегжейлі сипаттамасын береді. Бұл метрикалар әрбір модельдің қай аймақтарда жақсы нәтиже көрсететінін және қай жерлерде жетілдіру қажет екенін түсіну үшін маңызды болып табылады.

**Нәтижелер және талқылау.**

Бұл зерттеудің негізгі үлесі пластиктің әртүрлі түрлері мен түстерін анықтау, соның ішінде ластанған контейнерлерді анықтау, осылайша пластмассаларды жақын инфрақызыл (NIR) сәйкестендірумен байланысты шектеулерді шешу болып табылады. Пластикалық бөтелкелердің төрт негізгі түрінің NIR спектрлері - полиэтилентерефталат (ПЭТ), жоғары тығыздықтағы полиэтилен (HDPE), төмен тығыздықтағы полиэтилен (LDPE) және полипропилен (PP) - конволюционды нейрондық желіге (CNN) бір өлшемді кіріс ретінде пайдаланылды. Ұсынылған тәсілдің өнімділігін салыстыру үшін логистикалық регрессия, жартылай ең кіші квадраттар (PLS) және сызықтық дискриминанттық талдау (LDA) сияқты бірнеше дәстүрлі алгоритмдер де қолданылды. Әртүрлі әдістерді қолдану арқылы алынған нәтижелер 2-кестеде жинақталған. Әрбір жол белгілі бір кезеңде модельді бағалаудың негізгі көрсеткіштерін көрсетеді.

2 кесте – Төрт әдістің барлық кезеңдер бойынша бағалау көрсеткіштерін салыстыру

Алгоритм түрі	Кезең	Accuracy	Precision	Recall	F1-score
<b>Logistic Regression</b>	1	0.73333	0.55556	1.0	0.71429
	2	0.55	0.5	0.4	0.44444
	3	0.69231	0.6	0.6	0.6
<b>PLS</b>	1	0.66667	0.44444	1.0	0.61538
	2	0.55	0.5	0.4	0.44444
	3	0.69231	0.6	0.6	0.6
<b>LDA</b>	1	0.73333	0.66667	0.85714	0.75
	2	0.54545	0.75	0.42857	0.54545
	3	0.61538	0.6	0.5	0.54545
<b>CNN+ LSTM</b>	1	0.86667	0.77778	1.0	0.875
	2	0.72727	0.5	0.66667	0.57143
	3	0.76923	0.8	0.66667	0.72727



7 сурет – Әртүрлі әдістер арасындағы дәлдік көрсеткіштерін салыстыру нәтижесі

2-кесте мен 7-суреттен көрініп тұрғандай, CNN+LSTM моделі пластикалық контейнерлерді классификациялауда жоғары тиімділікті көрсетті, дәстүрлі әдістермен салыстырғанда (мысалы, Логистикалық регрессия, PLS және LDA) жоғары нәтиже берді.

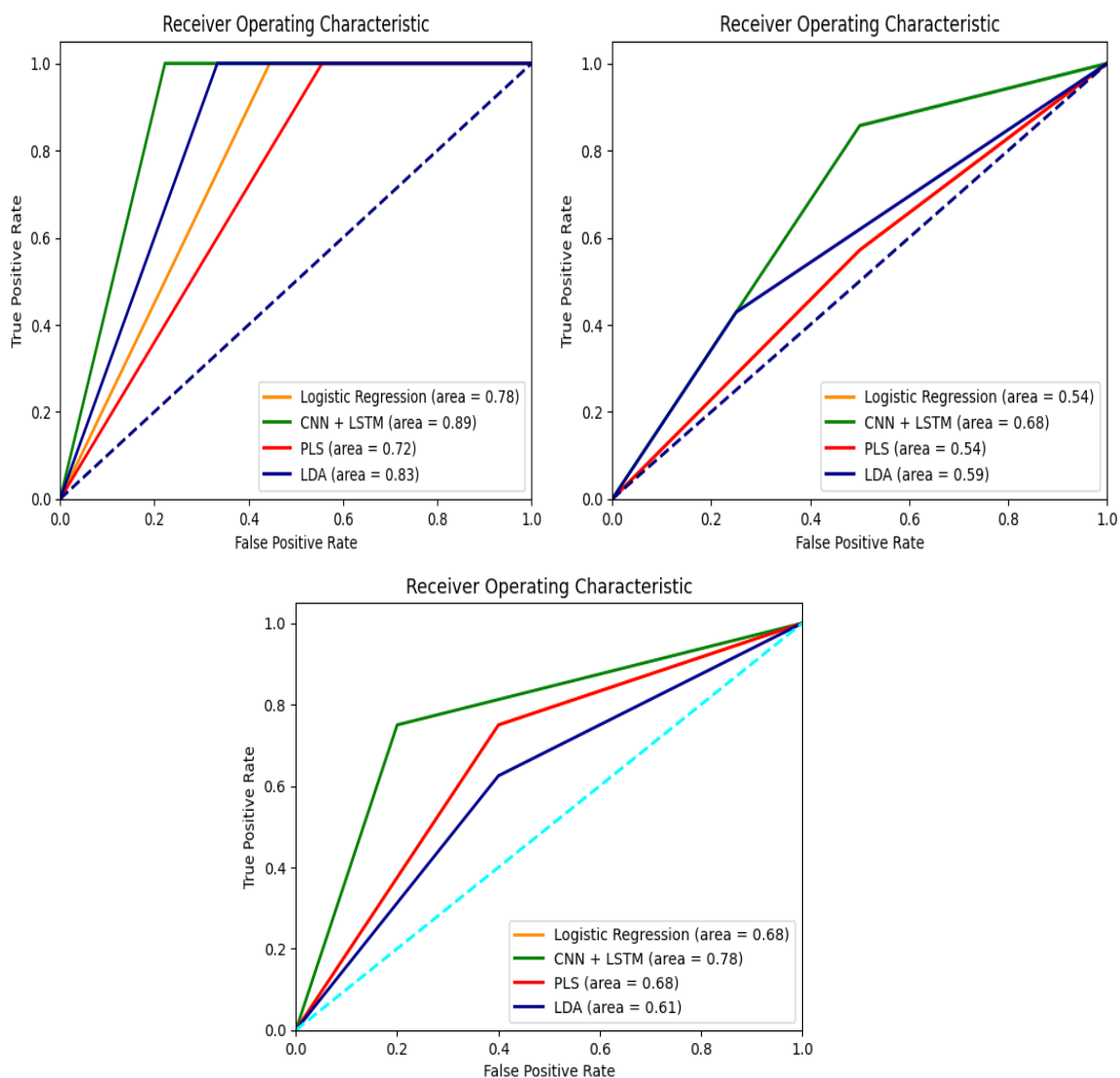
1-кезеңде, яғни PET бөтелкелері мен контейнерлердің әртүрлі түрлері, сондай-ақ тұрмыстық химияда жиі қолданылатын PP, HDPE және LDPE пластикалық контейнерлері қамтылған жағдайда, CNN+LSTM моделі дәлдік (accuracy) 0.86667, нақтылық (precision) 0.77778, қамту (recall) 1.0 және F1-өлшемі 0.875 көрсетті. Бұл нәтиже PET пен басқа полимерлер арасындағы спектрлік айырмашылықтар айқын болған кезде модельдің жоғары классификация қабілетін көрсетеді.

2-кезеңде түссіз (мөлдір - transparent) және түрлі-түсті пластикалық заттарды қамтыды, олар PET немесе PP болуы мүмкін. Бұл кезеңде жартылай түссіз және толық түссіз бөтелкелер арасындағы спектрлік айырмашылықтардың аз болуы, яғни түссіз пластмасы байланысты модель кейбір қиындықтарға тап болды. Осы себепті өнімділік төмендеп, CNN+LSTM моделі үшін дәлдік 0.72727, нақтылық 0.5, қамту 0.66667 және F1-өлшемі 0.57143 деңгейіне жетті. Дәстүрлі әдістер: Логистикалық регрессия мен PLS төменірек нәтижелер көрсетті, бұл олардың спектрлік айырмашылықтарды ажырату қабілетінің шектеулі екенін көрсетті. Себебі, түссіз немесе жартылай мөлдір пластиктер (PET, PP) жарықты жұту (абсорбция) және шағылысу (рефлексия) процестері төмен болғандықтан, NIR спектрометрінде әлсіз сигнал береді. Дегенмен машиналық оқыту алгоритмдері арқылы бар сигналмен классификация жасауға болатынын көреміз. Ал түсті немесе тығыз пластиктер жарықты көбірек сіңіретіндіктен, олардың спектрлік сигналдары айқын және классификация нәтижелері жоғары болады.

3-кезеңде классификация таза PET бөтелкелер мен тамақ өнімдерінің немесе қауіпті сұйықтықтардың қалдықтарымен ластанған бөтелкелер арасында жүргізілді. Бұл жағдайда әртүрлі ластану деңгейіне байланысты кластар арасындағы айырмашылықтар өзгеріп отырды. Ластану деңгейі өте төмен немесе көзге көрінбейтін зиянды заттар бар бөтелкелер тек аздаған спектрлік айырмашылықтарды көрсетті, бұл классификацияның дәлдігін төмендетті. Осы кезеңде ұсынылған CNN+LSTM моделі 76.923% дәлдік, 80% нақтылық, 66.667% қамту және 72.727 F1-өлшемін көрсетті. Дәстүрлі әдістердің де дәлдік көрсеткіштері төмендегенін байқауға болады. Бұның негізгі себептерінің бірі ол бөтелкенің жоғарыда айтылғандай түссіз болуы, сондай-ақ кейбір үлгілерде контейнерге құйылған зиянды сұйықтықтардың да түссіз болуы (бензин, түссіз химиялық гельдер) және

қалдықтардың толық бөтелкеде жайылмауы, яғни қалдықтың жартылай болуы секілді факторлар әсер етеді.

Сонымен қатар, 8-суретте көрсетілгендей, модельдің әртүрлі классификация шектері бойынша өнімділігін бағалау үшін ROC қисығы жасалды. Нәтижелер әртүрлі пластикалық қалдықтардың ішіндегі әртүрлі полимерлерді дәл жіктеуде ұсынылған алгоритмнің тиімділігін көрсетеді.



8 сурет – Барлық кезең бойынша ROC қисығы

Әрбір модельдің ROC қисықтары шекті мәндердің өзгеруіне байланысты нақты оң көрсеткіштер (сезімталдық) және жалған оң көрсеткіштер (1 – ерекшелік) арасындағы қатынасты көрсетеді. "Receiver Operating Characteristic" қисығының астындағы (AUC) аудан модельдің класстарын ажырату қабілетін бір ғана мән арқылы бағалауға мүмкіндік береді. AUC мәнінің жоғары болуы классификатордың класстарды жақсы ажырата алатынын білдіреді.

Осылайша, зерттеу нәтижелері CNN+LSTM моделінің пластикалық қалдықтарды классификациялау кезінде дәстүрлі әдістерге қарағанда жоғары дәлдік көрсететінін ғылыми тұрғыдан дәлелдеді, әсіресе 1-кезеңде, спектрлік контраст жоғары болды. Алайда, спектрлік айырмашылықтар азайған сайын, әсіресе жартылай мөлдір және ластанған пластиктер

арасында, модельдің өнімділігі төмендейді, бұл аз мөлшерде ластанған материалдарды ажыратудың күрделілігін көрсетеді. Бұл нәтижелер CNN+LSTM сияқты терең нейрондық желі архитектураларының сұрыптау және қайта өңдеу міндеттеріне тиімді екенін растайды, сондай-ақ ұсынылған тәсілдің күрделі классификация мәселелерінде дәстүрлі әдістермен салыстырғанда жоғары дәлдік пен сенімділікке ие екенін дәлелдейді.

### **Қорытынды.**

Бұл зерттеуде жақын инфрақызыл (NIR) спектроскопиясы арқылы алынған спектрлік деректер негізіндегі пластикалық қалдықтарды классификациялау үшін CNN+LSTM моделі әзірленіп, сынақ жасалды. Нәтижелер бұл модельдің логистикалық регрессия, ішінара ең кіші квадраттар (PLS) және сызықтық дискриминанттық талдау (LDA) сияқты дәстүрлі әдістермен салыстырғанда, әсіресе спектрлік айырмашылықтары төмен немесе ластануы бар пластмассаларды жіктеуде белгілі бір артықшылықтары бар екенін көрсетеді. Модельдің жоғары классификация дәлдігі оның кеңістіктік және уақыттық ерекшеліктерді тиімді түрде бөліп алу қабілетімен түсіндіріледі, бұл материалдың негізгі қасиеттерін дәл анықтауға мүмкіндік береді.

CNN+LSTM моделінің сұрыптаудың бастапқы кезеңіндегі өнімділігі ең жоғары болды, себебі бұл кезеңде ПЭТ пен басқа пластиктер арасындағы спектрлік айырмашылықтар айқын байқалды. Осы кезеңде модель дәлдік, нақтылық және F1-өлшемі бойынша ең жоғары көрсеткіштерге жетті, бұл оның материалдардың спектрлік қолтаңбалары айқын болған жағдайда тиімді екенін дәлелдейді. Алайда, жартылай мөлдір немесе ластанған пластиктерді ажырату кезінде модельдің өнімділігі төмендеді, себебі бұл материалдардың спектрлері арасындағы айырмашылықтар аз болды. Бұл шектеулер спектрлік деректерді өңдеу және модель архитектурасын жетілдіру қажеттілігін көрсетеді.

Болашақ зерттеулер спектрлік деректердің дәлдігін арттыру үшін гиперспектралды бейнелеуді енгізуге және алдын ала өңдеу әдістерін жетілдіруге бағытталатын болады. Қазіргі таңда қолданылған белгілерді бөліп алу және нормализация әдістеріне қарамастан, неғұрлым күрделі деректерді өңдеу тәсілдері модельдің спектрлік айырмашылықтарды дәлірек анықтауына мүмкіндік беруі ықтимал, бұл әсіресе ластанған немесе визуалды ұқсас пластиктерді жақсы ажыратуға ықпал етеді. Сонымен қатар, болашақ зерттеулер Transformer негізіндегі архитектураларды зерттеуге және қалдықтарды сұрыптаудың әртүрлі жағдайларында модельдің нақтылығын арттыруға бағытталатын болады. Бұл технологияларды одан әрі дамыту автоматтандырылған пластикалық қалдықтарды сұрыптау жүйелерін жетілдіруге және тұрақты қайта өңдеу процестерін дамытуға ықпал етеді.

Ұсынылған әдістің практикалық маңыздылығы жоғары, оны пластикалық қалдықтарды сұрыптаудың автоматтандырылған жүйелерінде қолдануға болады. Әдіс авиациялық салада қалдықтарды сұрыптау жүйелеріне енгізу үшін де қолдануға болады. Әуежайларда және әуе кемелерінде жиналатын пластик қалдықтарды автоматтандырылған түрде сұрыптау қайта өңдеу процестерінің тиімділігін арттыруға ықпал етіп, экологиялық тұрақтылыққа қолдау көрсете алады.

**Алғыс білдіру.** Бұл зерттеу жұмысы Қазақстан Республикасы Ғылым және жоғары білім министрлігі Ғылым комитетінің қаржылай қолдауымен жүзеге асырылды (грантық қаржыландыру АР22685518 ғылыми жобасы).

### **Әдебиеттер**

1. Анализ размера и доли рынка пластиковой упаковки – тенденции роста и прогнозы (2024–2029 гг.) Дереккөз: <https://www.mordorintelligence.com/ru/industry-reports/plastic-packaging-market> [Қол жетімділік: 2024 жылдың 10 қазан]

2. Bonifazi, G., Capobianco, G., & Serranti, S. (2018). A hierarchical classification approach for recognition of low-density (LDPE) and high-density polyethylene (HDPE) in mixed plastic waste based on short-wave infrared (SWIR) hyperspectral imaging. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *198*, 115–122. <https://doi.org/10.1016/j.saa.2018.03.006>
3. Gent, M., Sierra, H. M., Álvarez, M. M., & McCulloch, J. (2018). An evaluation of hydrocyclones and the LARCODEMS cylindrical cyclone for the separation of waste plastics of proximate densities. *Waste Management*, *79*, 374–384. <https://doi.org/10.1016/j.wasman.2018.08.004>
4. Zhang, Y., Wang, Q., Yalikun, N., Wang, H., Wang, C., & Jiang, H. (2023). A comprehensive review of separation technologies for waste plastics in urban mine. *Resources, Conservation & Recycling*, *197*, 107087. <https://doi.org/10.1016/j.resconrec.2023.107087>
5. Ragaert, K., Delva, L., & Van Geem, K. (2017). Mechanical and chemical recycling of solid plastic waste. *Waste Management*, *69*, 24–58. <https://doi.org/10.1016/j.wasman.2017.07.044>.
6. Gundupalli, S. P., Hait, S., & Thakur, A. (2017). A review on automated sorting of source-separated municipal solid waste for recycling. *Waste Management*, *60*, 56–74. <https://doi.org/10.1016/j.wasman.2016.09.015>
7. Shameem, K. M. M., Choudhari, K. S., & Bankapur, A. (2017). A hybrid LIBS–Raman system combined with chemometrics: an efficient tool for plastic identification and sorting. *Analytical and Bioanalytical Chemistry*, *409*, 3299–3308. <https://doi.org/10.1007/s00216-017-0268-z>
8. Brunner, S., Fomin, P., & Kargel, C. (2015). Automated sorting of polymer flakes: fluorescence labeling and development of a measurement system prototype. *Waste Management*, *38*, 49–60. <https://doi.org/10.1016/j.wasman.2014.12.006>
9. Safavi, S. M., Masoumi, H., & Mirian, S. (2010). Sorting of polypropylene resins by color in MSW using visible reflectance spectroscopy. *Waste Management*, *30* (11), 2216–2222. <https://doi.org/10.1016/j.wasman.2010.06.023>
10. Kumar, S., Yadav, D., Gupta, H., Verma, O. P., Ansari, I. A., & Ahn, C. W. (2021). A novel YOLOv3 algorithm-based deep learning approach for waste segregation: Towards smart waste management. *Electronics*, *10*(1), 14. <https://doi.org/10.3390/electronics10010014>
11. Tan, Z., Fei, Z., Zhao, B., Yang, J., Xu, X., & Wang, Z. (2021). Identification for recycling polyethylene terephthalate (PET) plastic bottles by polarization vision. *IEEE Access*, *9*, 27510–27517. <https://doi.org/10.1109/ACCESS.2021.3050816>
12. Ahmed, M. I. B., Alotaibi, R. B., Al-Qahtani, R. A., Al-Qahtani, R. S., Al-Hetela, S. S., Al-Matar, K. A., ... & Youldash, M. (2023). Deep learning approach to recyclable products classification: Towards sustainable waste management. *Sustainability*, *15*, 11138. <https://doi.org/10.3390/su151411138>
13. Singh, M. K., Hait, S., & Thakur, A. (2023). Hyperspectral imaging-based classification of post-consumer thermoplastics for plastics recycling using artificial neural network. *Process Safety and Environmental Protection*, *179*, 593–602. <https://doi.org/10.1016/j.psep.2023.09.052>
14. Qiu, Z., Zhao, S., Feng, X., & He, Y. (2020). Transfer learning method for plastic pollution evaluation in soil using NIR sensor. *Science of the Total Environment*, *740*, 140118. <https://doi.org/10.1016/j.scitotenv.2020.140118>
15. Marchesi, C., Rani, M., Federici, S., Lancini, M., & Depero, L. E. (2023). Evaluating chemometric strategies and machine learning approaches for a miniaturized near-infrared spectrometer in plastic waste classification. *Acta IMEKO*, *12*(2), 1-7. <https://doi.org/10.21014/actaimeko.v12i2.1531>

16. Masoumi, H., Safavi, S. M., & Khani, Z. (2012). Identification and classification of plastic resins using near infrared reflectance spectroscopy. *International Journal of Mechanical and Industrial Engineering*, 6, 213–220.

17. Xia, J., Huang, Y., Li, Q., Xiang, Y., & Min, Sh. (2021). Convolutional neural network with near-infrared spectroscopy for plastic discrimination. *Environmental Chemistry Letters*, 19, 3547–3555. <https://doi.org/10.1007/s10311-021-01240-9>

18. Olowolayemo, A., Radzi, N. I. A., & Ismail, N. F. (2022). Classifying plastic waste using deep convolutional neural networks for efficient plastic waste management. *International Journal on Perceptive and Cognitive Computing*, 8(2), 6–15. <https://doi.org/10.5281/zenodo.6457607>

### References

1. Analiz obema i doli rynka plastikovoj upakovki – tendencii rosta i prognozy (2024–2029 gg.) Retrieved October 10, 2024, from <https://www.mordorintelligence.com/industry-reports/plastic-packaging-market>

2. Bonifazi, G., Capobianco, G., & Serranti, S. (2018). A hierarchical classification approach for recognition of low-density (LDPE) and high-density polyethylene (HDPE) in mixed plastic waste based on short-wave infrared (SWIR) hyperspectral imaging. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 198, 115–122. <https://doi.org/10.1016/j.saa.2018.03.006>

3. Gent, M., Sierra, H. M., Álvarez, M. M., & McCulloch, J. (2018). An evaluation of hydrocyclones and the LARCODEMS cylindrical cyclone for the separation of waste plastics of proximate densities. *Waste Management*, 79, 374–384. <https://doi.org/10.1016/j.wasman.2018.08.004>

4. Zhang, Y., Wang, Q., Yalikun, N., Wang, H., Wang, C., & Jiang, H. (2023). A comprehensive review of separation technologies for waste plastics in urban mine. *Resources, Conservation & Recycling*, 197, 107087. <https://doi.org/10.1016/j.resconrec.2023.107087>

5. Ragaert, K., Delva, L., & Van Geem, K. (2017). Mechanical and chemical recycling of solid plastic waste. *Waste Management*, 69, 24–58. <https://doi.org/10.1016/j.wasman.2017.07.044>.

6. Gundupalli, S. P., Hait, S., & Thakur, A. (2017). A review on automated sorting of source-separated municipal solid waste for recycling. *Waste Management*, 60, 56–74. <https://doi.org/10.1016/j.wasman.2016.09.015>

7. Shameem, K. M. M., Choudhari, K. S., & Bankapur, A. (2017). A hybrid LIBS–Raman system combined with chemometrics: an efficient tool for plastic identification and sorting. *Analytical and Bioanalytical Chemistry*, 409, 3299–3308. <https://doi.org/10.1007/s00216-017-0268-z>

8. Brunner, S., Fomin, P., & Kargel, C. (2015). Automated sorting of polymer flakes: fluorescence labeling and development of a measurement system prototype. *Waste Management*, 38, 49–60. <https://doi.org/10.1016/j.wasman.2014.12.006>

9. Safavi, S. M., Masoumi, H., & Mirian, S. (2010). Sorting of polypropylene resins by color in MSW using visible reflectance spectroscopy. *Waste Management*, 30 (11), 2216–2222. <https://doi.org/10.1016/j.wasman.2010.06.023>

10. Kumar, S., Yadav, D., Gupta, H., Verma, O. P., Ansari, I. A., & Ahn, C. W. (2021). A novel YOLOv3 algorithm-based deep learning approach for waste segregation: Towards smart waste management. *Electronics*, 10(1), 14. <https://doi.org/10.3390/electronics10010014>

11. Tan, Z., Fei, Z., Zhao, B., Yang, J., Xu, X., & Wang, Z. (2021). Identification for recycling polyethylene terephthalate (PET) plastic bottles by polarization vision. *IEEE Access*, 9, 27510–27517. <https://doi.org/10.1109/ACCESS.2021.3050816>

12. Ahmed, M. I. B., Alotaibi, R. B., Al-Qahtani, R. A., Al-Qahtani, R. S., Al-Hetela, S. S., Al-Matar, K. A., ... & Youldash, M. (2023). Deep learning approach to recyclable products

classification: Towards sustainable waste management. *Sustainability*, 15, 11138. <https://doi.org/10.3390/su151411138>

13. Singh, M. K., Hait, S., & Thakur, A. (2023). Hyperspectral imaging-based classification of post-consumer thermoplastics for plastics recycling using artificial neural network. *Process Safety and Environmental Protection*, 179, 593–602. <https://doi.org/10.1016/j.psep.2023.09.052>

14. Qiu, Z., Zhao, S., Feng, X., & He, Y. (2020). Transfer learning method for plastic pollution evaluation in soil using NIR sensor. *Science of the Total Environment*, 740, 140118. <https://doi.org/10.1016/j.scitotenv.2020.140118>

15. Marchesi, C., Rani, M., Federici, S., Lancini, M., & Depero, L. E. (2023). Evaluating chemometric strategies and machine learning approaches for a miniaturized near-infrared spectrometer in plastic waste classification. *Acta IMEKO*, 12(2), 1-7. <https://doi.org/10.21014/actaimeko.v12i2.1531>

16. Masoumi, H., Safavi, S. M., & Khani, Z. (2012). Identification and classification of plastic resins using near infrared reflectance spectroscopy. *International Journal of Mechanical and Industrial Engineering*, 6, 213–220.

17. Xia, J., Huang, Y., Li, Q., Xiang, Y., & Min, Sh. (2021). Convolutional neural network with near-infrared spectroscopy for plastic discrimination. *Environmental Chemistry Letters*, 19, 3547–3555. <https://doi.org/10.1007/s10311-021-01240-9>

18. Olowolayemo, A., Radzi, N. I. A., & Ismail, N. F. (2022). Classifying plastic waste using deep convolutional neural networks for efficient plastic waste management. *International Journal on Perceptive and Cognitive Computing*, 8(2), 6–15. <https://doi.org/10.5281/zenodo.6457607>

### NEURAL NETWORK APPLICATION FOR OPTIMIZING THE SORTING PROCESS OF POLYMER CONTAINERS

**Abstract.** This study investigates the use of a hybrid neural network architecture that combines convolutional neural networks (CNN) and long short-term memory (LSTM) for efficient sorting of plastic containers. The study focuses on classifying plastic waste based on chemical composition and contamination level, obtained with a near-infrared (NIR) spectroscopy device. Experimental results show that the CNN+LSTM hybrid model achieves relatively high accuracy in recognizing different types and colors of plastics, including the detection of contaminants in containers. A comparative evaluation of the model's performance was conducted with traditional classification methods such as logistic regression, partial least squares (PLS), and linear discriminant analysis (LDA). The results show that the CNN+LSTM model performs better than traditional approaches, especially in scenarios with small spectral differences between classes. This study demonstrates the potential of machine learning to improve the efficiency of plastic waste sorting and recycling processes, thereby contributing to improved environmental sustainability.

**Keywords.** Plastic waste, NIRS (Near-infrared spectroscopy), Neural network, Hybrid model, CNN, LSTM.

### ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОПТИМИЗАЦИИ ПРОЦЕССА СОРТИРОВКИ ПОЛИМЕРНЫХ КОНТЕЙНЕРОВ

**Аннотация.** В этом исследовании рассматривается использование архитектуры гибридной нейронной сети, сочетающей сверточные нейронные сети (CNN) и долговременную память (LSTM), для эффективной сортировки пластиковых контейнеров. Исследование было сосредоточено на классификации пластиковых отходов на основе химического состава и уровня загрязнения, полученной с помощью спектроскопического устройства ближнего инфракрасного диапазона (NIR). Результаты экспериментов показывают, что гибридная модель CNN+LSTM обеспечивает относительно высокую



точность распознавания различных типов и цветов пластика, включая обнаружение загрязняющих веществ в контейнерах. Сравнительная оценка производительности модели проводилась с использованием традиционных методов классификации, таких как логистическая регрессия, частичные наименьшие квадраты (PLS) и линейный дискриминантный анализ (LDA). Результаты показывают, что модель CNN+LSTM работает лучше, чем традиционные подходы, особенно в сценариях с небольшими спектральными различиями между классами. Это исследование демонстрирует потенциал машинного обучения для повышения эффективности процессов сортировки и переработки пластиковых отходов, тем самым способствуя повышению экологической устойчивости.

**Ключевые слова.** Пластиковые отходы, NIRS (ближняя инфракрасная спектроскопия), Нейронная сеть, Гибридная модель, CNN, LSTM.

#### Авторлар туралы мәлімет

Алимбекова Назым Ахатовна	магистр, компьютерлік және программалық инженерия кафедрасы, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: <a href="mailto:nazakhatovna@gmail.com">nazakhatovna@gmail.com</a>
Хари Мохан Рай	PhD, Компьютер ғылымдар мектебі, Гачон университеті, Соннам, Оңтүстік Корея, E-mail: <a href="mailto:drhmrai@gachon.ac.kr">drhmrai@gachon.ac.kr</a>
Турымбетов Турсинбай Абдибекович	Қауымдастырылған профессор міндетін атқарушы, техника ғылымдарының кандидаты, гуманитарлық мектеп, Халықаралық туризм және меймандостық университеті, Түркістан, Қазақстан, E-mail: <a href="mailto:t.turimbetov@iuth.edu.kz">t.turimbetov@iuth.edu.kz</a>
Жумадилаева Айнура Канадиловна	Техника ғылымдарының кандидаты, доцент, компьютерлік және программалық инженерия кафедрасы, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан, E-mail: <a href="mailto:zhumadillayeva_ak@enu.kz">zhumadillayeva_ak@enu.kz</a>

#### Сведения об авторах

Алимбекова Назым Ахатовна	Магистр, Кафедра компьютерной и программной инженерии, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан, E-mail: <a href="mailto:nazakhatovna@gmail.com">nazakhatovna@gmail.com</a>
Хари Мохан Рай	PhD, доцент, Школа компьютерных наук, Университет Гачон, Соннам, Южная Корея, E-mail: <a href="mailto:drhmrai@gachon.ac.kr">drhmrai@gachon.ac.kr</a>
Турымбетов Турсинбай Абдибекович	И.о. ассоциированного профессора, кандидат технических наук, гуманитарная школа, международный университет туризма и гостеприимства, E-mail: <a href="mailto:t.turimbetov@iuth.edu.kz">t.turimbetov@iuth.edu.kz</a>
Жумадилаева Айнура Канадиловна	Доцент, кандидат технических наук, кафедра компьютерной и программной инженерии, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан, E-mail: <a href="mailto:zhumadillayeva_ak@enu.kz">zhumadillayeva_ak@enu.kz</a>

#### Information about the authors

Nazym Alimbekova	Department of Computer and Software Engineering, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: <a href="mailto:nazakhatovna@gmail.com">nazakhatovna@gmail.com</a>
Hari Mohan Rai	PhD, Associate Professor, School of Computing, Gachon University, Seong-nam-si, South Korea, E-mail: <a href="mailto:drhmrai@gachon.ac.kr">drhmrai@gachon.ac.kr</a>
Tursinbay Turymbetov	Acting Associate Professor, PhD in Engineering, Humanities school, International University of Tourism and Hospitality, Turkistan, Kazakhstan; E-mail: <a href="mailto:t.turimbetov@iuth.edu.kz">t.turimbetov@iuth.edu.kz</a>
Ainur Zhumadillayeva	Associate Professor, PhD in Engineering, Department of Computer and Software Engineering, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: <a href="mailto:zhumadillayeva_ak@enu.kz">zhumadillayeva_ak@enu.kz</a>

УДК 004.8  
МРНТИ 82.29.09  
[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_8](https://doi.org/10.53364/24138614_2025_36_1_8)

А.С. Құмарқанова<sup>1</sup>, З.Т. Хасенова<sup>1</sup>, Ю.А. Вайс<sup>1\*</sup>

<sup>1</sup>НАО «Восточно-Казахстанский технический университет имени Д. Серикбаева»,  
Усть-Каменогорск, Казахстан

E-mail: yuvais@edu.ektu.kz\*

## АНАЛИЗ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ СПРОСА НА БАЗЕ ИСТОРИЧЕСКИХ ДАННЫХ РАСХОДА ЛЕКАРСТВЕННЫХ СРЕДСТВ

***Аннотация.** Прогнозирование спроса на лекарственные средства играет ключевую роль в обеспечении устойчивого снабжения, эффективного управления запасами и своевременного доступа пациентов к жизненно важным лекарствам. В данной статье представлено исследование методов временных рядов ARIMA и экспоненциального сглаживания для прогнозирования спроса на антигипертензивное лекарственное средство от давления. Исследование направлено на разработку и выявление модели, обеспечивающей высокую точность и эффективность прогнозирования на основе выбранных наборов данных.*

*Анализ включает в себя исследование методов прогнозирования, проведение сбора и обработки данных, определение оптимальных параметров для каждого метода, разработку гибридной модели, оценку точности на основе заданных метрик и анализ результатов. В ходе проведенного исследования установлено, что наиболее эффективными методами прогнозирования являются подходы на основе временных рядов, включая модели ARIMA и методы экспоненциального сглаживания. А разработанная гибридная модель демонстрирует высокую точность прогнозов благодаря комбинированию преимуществ двух подходов. Результаты показывают, что гибридная модель превосходит ARIMA и экспоненциальное сглаживание по всем ключевым метрикам. На основании полученных выводов предложено внедрение гибридной модели для повышения точности прогнозирования спроса в фармацевтической отрасли.*

***Ключевые слова:** прогнозирование, методы прогнозирования, экспоненциальное сглаживание, ARIMA, временные ряды, лекарственное средство, гибридная модель.*

### **Введение.**

Прогнозирование спроса на лекарственные средства является ключевым элементом современной системы здравоохранения, способствуя эффективному управлению цепочками поставок, своевременному обеспечению медицинских учреждений и аптек необходимыми медикаментами, а также минимизации риска их дефицита. Несмотря на то, что в настоящее время лекарства в аптеках в большинстве случаев имеются в наличии, статистические данные, используемые для управления запасами, остаются недостаточно гибкими и оперативными. Это создает риск резкого всплеска спроса, при котором текущих запасов может не хватить. Поэтому применение прогнозных моделей приобретает особую значимость для обеспечения пациентов необходимыми лекарственными средствами, поскольку своевременное наличие препаратов в аптечных и медицинских учреждениях способствует повышению качества лечения, предотвращению обострений заболеваний и снижению риска летальных исходов. Особенно актуальной эта задача становится в отношении препаратов для лечения сердечно-сосудистых заболеваний и заболеваний системы кровообращения, так как смертность от этих болезней самая высокая во всем мире. По данным Министерства здравоохранения Республики Казахстан ежегодно в нашей

стране от сердечно-сосудистых заболеваний умирает порядка 40 тысяч человек [1]. При этом болезни системы кровообращения составили 22,7% от общего числа зарегистрированных случаев смертности в 2022 году, занимая лидирующую позицию среди причин летальных исходов [2]. Такой высокий показатель подчеркивает необходимость системного подхода к обеспечению доступности антигипертензивных препаратов.

Прогнозирование спроса на лекарственные средства осуществляется с использованием различных подходов, выбор которых определяется их категорией и влияющими факторами. Данный процесс представляет собой сложную задачу, требующую применения разнообразных аналитических подходов. Среди применяемых методов выделяются алгоритмы машинного обучения, включая нейронные сети, регрессионный анализ, методы интеллектуального анализа данных и модели временных рядов.

Алгоритмы машинного обучения применяются для прогнозирования временных рядов благодаря их способности моделировать сложные нелинейные зависимости. Например, в статье Мбониншуту, Нкурунзиза, Нийобухунгиро и Кайитаре [3] исследуется использование случайных лесов, линейной регрессии и метода искусственных нейронных сетей для прогнозирования спроса на жизненно важные лекарства на основе данных потребления в Руанде, где подчеркивается их высокая точность при наличии большого объема обучающих данных. Однако эти методы часто требуют значительных вычислительных ресурсов и большого объема данных для обучения, что делает их менее практичными для применения в условиях ограниченного времени и ресурсов.

Рекуррентные нейронные сети (RNN) и долгая краткосрочная память (LSTM) демонстрируют значительные успехи в задачах прогнозирования. В исследовании Элсворт и Гюттель [4] показано, что нейронные сети способны выявлять сложные нелинейные зависимости в данных, что делает их эффективными в задачах прогнозирования спроса. Тем не менее, их основной недостаток заключается в риске переобучения, особенно при небольшом объеме данных, а также в высокой вычислительной сложности.

Регрессионные методы, такие как линейная регрессия, традиционно используются для моделирования зависимостей в данных. Ученые Моуса и Аль-Хатиб [5] анализируют их применение для прогнозирования временных рядов в контексте фармацевтической логистики. Хотя регрессионные модели просты в интерпретации, они не всегда способны эффективно справляться с задачами, связанными с сезонностью или изменением трендов, что ограничивает их применение в сложных задачах прогнозирования.

Модели временных рядов, как ARIMA и экспоненциальное сглаживание, остаются стандартом в задачах прогнозирования спроса. За счет своей универсальности и эффективности они находят широкое применение в различных областях, включая медицину [6], транспорт [7], энергетику [8] и другие. В исследовании Захра и Путра [9] отмечается, что ARIMA обеспечивает высокий уровень точности при моделировании временных зависимостей, в то время как экспоненциальное сглаживание лучше справляется с задачами прогнозирования в условиях изменчивости данных. Преимущество этих методов заключается в их простоте, интерпретируемости и сравнительно низкой вычислительной сложности. Они хорошо адаптируются к задачам, связанным с сезонностью и трендами, что делает их наиболее подходящими для применения в фармацевтической отрасли.

Эффективное прогнозирование спроса лекарственных средств требует использования современных аналитических методов, отличающихся гибкостью, способностью учитывать сезонные колебания, изменения трендов и нерегулярные всплески спроса, а также обеспечивать оперативность и точность результатов. Сравнительный анализ различных методов прогнозирования показывает, что, несмотря на растущую популярность нейронных сетей и алгоритмов машинного обучения, традиционные модели временных рядов, такие как ARIMA и экспоненциальное сглаживание, остаются наиболее

эффективными для задач прогнозирования спроса на лекарственные средства. Эти методы обладают необходимой гибкостью, простотой настройки и способностью учитывать ключевые характеристики данных, включая тренды и сезонные колебания [10-13].

Цель исследования состоит в улучшении точности прогнозирования спроса на выбранную категорию лекарственных средств путем анализа различных моделей временных рядов. Новизна исследования заключается в разработке гибридной модели, интегрирующей подходы ARIMA и экспоненциального сглаживания. Используя сравнительный анализ определим точность прогнозирования спроса на выбранную категорию лекарственных средств для разработанной гибридной модели и традиционных моделей временных рядов.

### Материалы и методы.

Для проведения прогноза использовались данные по еженедельным расходам антигипертензивного лекарственного препарата физиотенз в медицинском центре (Рисунок 1). Данный набор охватывает временной интервал с 2003 года по конец 2023 года, что составляет 1096 записей. Данные включают информацию о расходах препарата, представленную в формате временных рядов, с указанием даты и соответствующего объема.

	A	B	C	D	E	F	G	H	I	J
1	Date	Weekly_ Consumption	Male_ Consumption	Female_ Consumption	Consumption_ Age_18_34	Consumption_ Age_35_59	Consumption_ Age_60_Plus	Price_ Euro	Percentage_with_ Cardiovascular_ Diseases	Percentage_Over_65_ Years_ Old
2	2003-01-05	117	50	67	9	42	66	7,12	12,6	17,48
3	2003-01-12	111	48	63	8	41	62	7,12	12,53	17,47
4	2003-01-19	113	47	66	9	41	63	7,12	12,54	17,52
5	2003-01-26	117	51	66	6	44	67	7,12	12,51	17,44
6	2003-02-02	120	52	68	8	43	69	7,12	12,53	17,48
7	2003-02-09	115	47	68	6	45	64	7,12	12,53	17,49
8	2003-02-16	116	48	68	10	41	65	7,12	12,5	17,45
9	2003-02-23	112	47	65	6	45	61	7,12	12,52	17,42
10	2003-03-02	116	48	68	7	43	66	7,12	12,61	17,47

Рисунок 1 – Набор данных по еженедельным расходам антигипертензивного лекарственного препарата

В ходе предварительного анализа было проведено исследование временных зависимостей в данных с использованием функций автокорреляции (ACF) и частичной автокорреляции (PACF), результаты которых отображены на графиках рисунка 2.

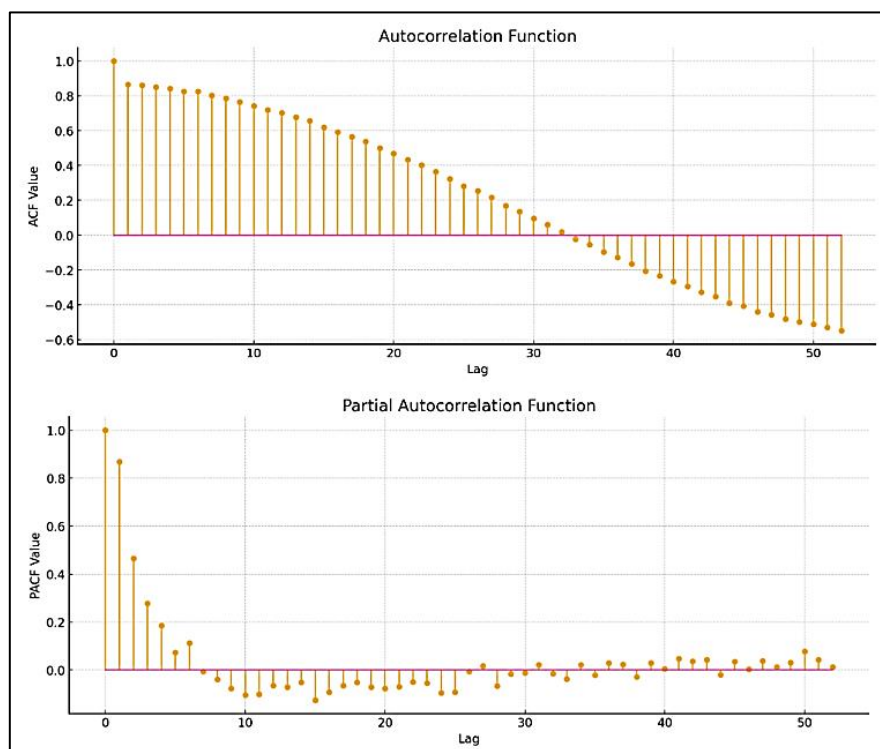


Рисунок 2 – ACF и PACF для анализа временного ряда потребления лекарств

Как показано на рисунке 2, ACF указывает на убывающий тренд зависимости, что свидетельствует о наличии автокорреляции в данных. График показывает сильную автокорреляцию при низких задержках, которая указывает на наличие сезонности в данных. PACF отражает наиболее значимые зависимости на первых задержках, что является ключевым фактором при выборе параметров модели ARIMA. Предварительный анализ данных выявил высокую регулярность и структурированность, подтверждающую их пригодность для методов временных рядов.

Модели обучаются на указанных выше исторических данных, что обеспечивает выявление основных закономерностей и тенденций. Обучающая выборка составляет 80% данных, в то время как оставшиеся 20% используются для тестирования.

Для анализа временных рядов и прогнозирования спроса на лекарственный препарат использовались два ключевых метода: ARIMA и экспоненциальное сглаживание.

Метод ARIMA (Autoregressive Integrated Moving Average) применяется для прогнозирования временных рядов за счет учета временной зависимости, преобразования нестационарных данных в стационарные и сглаживания ошибок прогноза [14,15]. Он включает три компонента: авторегрессию (AR), интеграцию (I) и скользящее среднее (MA). Компонент авторегрессии описывает связь переменной с ее предыдущими значениями и описывается формулой (1) ниже:

$$\text{AR}(p): X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t, \quad (1)$$

где,  $X_t$  - текущее значение временного ряда,  $c$  - постоянная,  $\phi_p$  - коэффициенты авторегрессии, характеризующие влияние предыдущих значений временного ряда,  $X_{t-p}$  - предыдущие значения временного ряда (с лагами от 1 до  $p$ ),  $\varepsilon_t$  - ошибка прогноза.

Интеграционный компонент использует дифференцирование временных рядов для достижения стационарности и определяется по следующей формуле (2):

$$\text{I}(d): (1 - L)^d X_t = X_t - X_{t-1}, \quad (2)$$

где,  $(1-L)^d$  - оператор дифференцирования порядка  $d$ , который преобразует нестационарный ряд в стационарный,  $L$  - лаговый оператор,  $d$  - порядок дифференцирования,  $X_{t-1}$  - предыдущее значение временного ряда.

Компонент скользящей средней описывает временной ряд с учетом прошлых ошибок прогнозирования и описывается формулой (3), приведенной ниже:

$$MA(q): X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (3)$$

где,  $\mu$  - среднее значение временного ряда,  $\varepsilon_t$  - текущая ошибка прогноза,  $\theta_q$  - коэффициенты скользящего среднего, характеризующие влияние ошибок предыдущих прогнозов,  $\varepsilon_{t-q}$  - предыдущие ошибки прогноза.

Перед применением ARIMA к данным была проведена проверка их стационарности, показавшая наличие тренда. Для его устранения применялось дифференцирование. Модель ARIMA была построена с параметрами  $p, d, q$ , а оптимальные значения были определены на основе анализа функций ACF и PACF для использованного набора данных.

Экспоненциальное сглаживание — это метод, который использует средневзвешенное значение исторических данных для прогнозирования будущих значений. Основная идея состоит в том, чтобы сгладить старые данные, которые со временем экспоненциально уменьшаются за счет придания веса [16]. Для старых данных вес уменьшается в геометрической прогрессии. Основное уравнение метода экспоненциального сглаживания выглядит следующим образом (4):

$$S_t = \alpha \cdot Y_t + (1 - \alpha) \cdot S_{t-1}, \quad (4)$$

где,  $S_t$  - прогнозируемое (сглаженное) значение во времени  $t$ ,  $Y_t$  - фактическое значение временного ряда в момент времени  $t$ ,  $S_{t-1}$  - прогнозируемое значение в предыдущий момент времени  $t-1$ ,  $\alpha$  - коэффициент сглаживания (от 0 до 1).

Для повышения точности прогноза была разработана гибридная модель, объединяющая преимущества ARIMA и экспоненциального сглаживания (Рисунок 3).



Рисунок 3 – Схема построения гибридной модели

Как представлено на рисунке 3, процесс построения гибридной модели включает пять этапов:

1. Декомпозиция временного ряда. Исходные данные временного ряда разделяются на трендовые и остаточные компоненты.

2. Анализа трендовой части методом ARIMA.

3. Моделирование остаточной части для учета сезонных изменений методом экспоненциального сглаживания.

4. Комбинирование результатов на основе объединения прогнозов, используя взвешенный подход, трендовой и остаточной компонент.

## 5. Формирование гибридной модели с наибольшей точностью прогноза.

### Результаты и обсуждение.

В рамках исследования были использованы три подхода для прогнозирования спроса на лекарственный препарат от давления: ARIMA, экспоненциальное сглаживание и гибридная модель, объединяющая оба метода. Процесс прогнозирования включает этапы выбора параметров, обучения модели, формирования прогноза и последующего анализа результатов.

В результате вычислительного эксперимента для модели ARIMA оптимальными параметрами на используемом наборе данных определены значения  $p=1$ ,  $d=0$ ,  $q=1$ . Для прогнозирования спроса на лекарства с использованием метода экспоненциального сглаживания был настроен коэффициент сглаживания ( $\alpha$ ). Параметр  $\alpha$  подбирается с учетом минимизации ошибки прогнозирования. Оптимальное значение было определено на основе минимального показателя средней абсолютной процентной ошибки (MAPE), представленное в таблице 1. Тестирование выявило, что наименьшее значение ошибки достигается при параметре  $\alpha=0,3$ .

Таблица 1 - Значения параметра сглаживания на основе значений MAPE

Alpha	MAPE
$\alpha=0$	5.66%
$\alpha=0,1$	4.79%
$\alpha=0,2$	4.39%
<b><math>\alpha=0,3</math></b>	<b>4.29%</b>
$\alpha=0,4$	4.44%
$\alpha=0,5$	4.78%
$\alpha=0,6$	5.32%
$\alpha=0,7$	5.98%
$\alpha=0,8$	6.69%
$\alpha=0,9$	7.45%
$\alpha=1$	8.23%

Результаты сопоставления моделей ARIMA, экспоненциального сглаживания и гибридной модели с исходными данными представлены на рисунках 4–6.

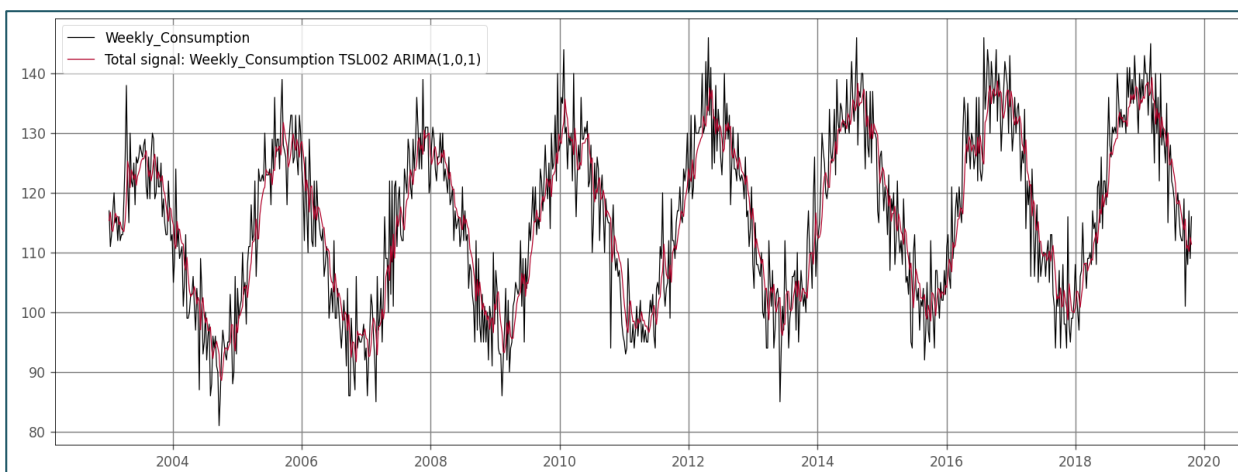


Рисунок 4 – Сопоставление модели ARIMA с историческими данными



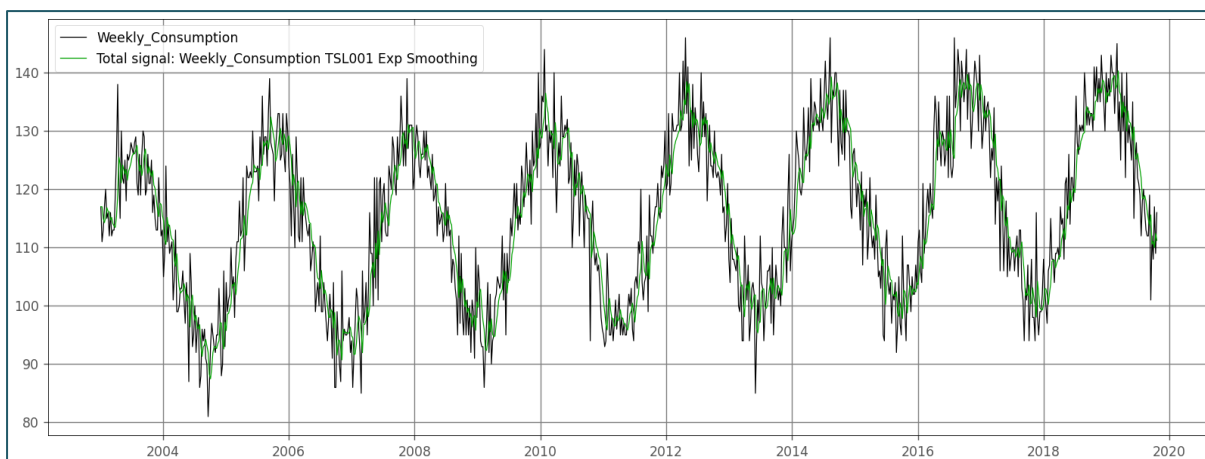


Рисунок 5 – Сопоставление модели экспоненциального сглаживания с историческими данными

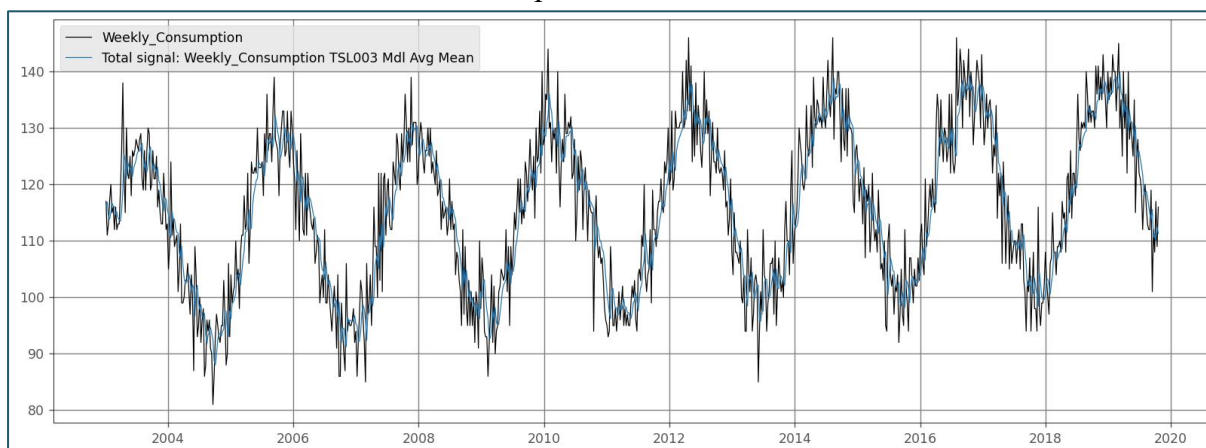


Рисунок 6 – Сопоставление гибридной модели с историческими данными

После этапа обучения модель применяется для прогнозирования. На рисунке 7 представлен результат прогноза спроса на лекарственное средство с использованием моделей ARIMA, экспоненциального сглаживания и гибридной модели. Фактические данные представлены черной линией, а затененная область в правой части графика соответствует периоду прогнозирования. Она отражает будущий прогноз спроса на основе трех моделей, что позволяет сравнить их точность и поведение.

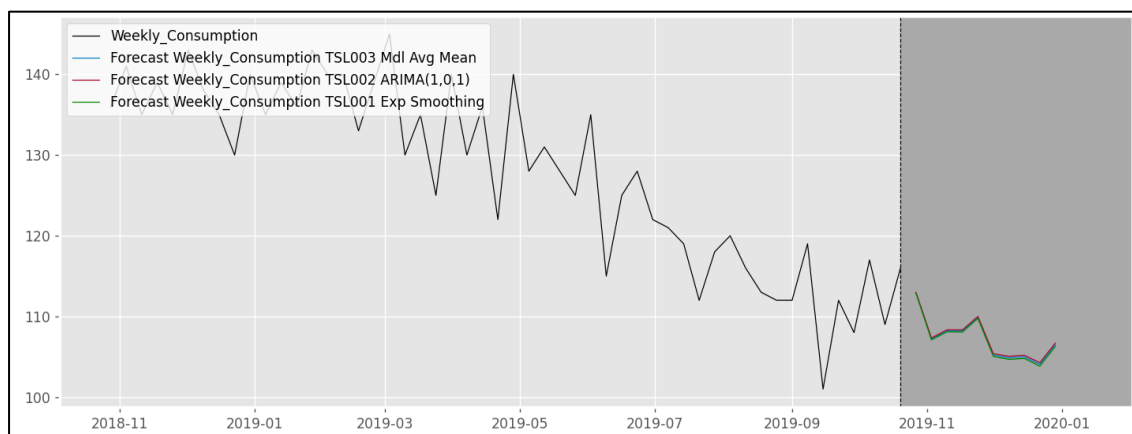


Рисунок 7 – Результаты прогнозирования спроса по трем моделям



Для оценки эффективности моделей прогнозирования были использованы метрики: средняя абсолютная ошибка (MAE), среднеквадратичная ошибка (RMSE) и средняя абсолютная процентная ошибка (MAPE). Таблица статистических показателей для всех трех моделей представлена на рисунке 8.

TRAINING SAMPLE MODEL FIT			
<b>Variable:</b> Weekly_Consumption			
<b>Model(s):</b>			
TSL001 Exp Smoothing			
TSL002 ARIMA(1,0,1)			
TSL003 Mdl Avg Mean			
	<b>TSL001</b>	<b>TSL002</b>	<b>TSL003</b>
in-sample MSE	38.153	37.906	37.910
... RMSE	6.177	6.157	6.157
... MAE	4.891	4.880	4.876
... MAPE	4.299	4.294	4.287
Sample size	877	877	877
Effective sample size	876	875	875
* based on one-step-ahead forecast errors			

Рисунок 8 – Показатели оценки эффективности

Сравнительный анализ показал, что гибридная модель (TSL003) продемонстрировала самые низкие значения метрик MAE, RMSE и MAPE по сравнению с моделями ARIMA (TSL002) и экспоненциального сглаживания (TSL001), что указывает на более высокое качество модели. Таким образом, гибридная модель демонстрирует лучшую точность в прогнозировании.

#### **Заключение.**

В рамках исследования было проведено сравнение методов прогнозирования временных рядов ARIMA, экспоненциального сглаживания и их гибридной модели в прогнозировании спроса на антигипертензивное лекарственное средство физиотенз. Исследование охватывало временной интервал с 2003 по 2023 год и включало анализ еженедельных расходов препарата.

Результаты сравнительного анализа показали, что гибридная модель превосходит отдельные подходы ARIMA и экспоненциального сглаживания. Гибридная модель обеспечила самые низкие значения ошибок, что подтверждает её высокую точность и надежность. Такой подход позволил объединить сильные стороны обеих моделей: способность ARIMA эффективно моделировать долгосрочные тренды и адаптивность экспоненциального сглаживания к сезонным изменениям.

Практическая значимость исследования заключается в применении предложенной гибридной модели для оптимизации управления запасами медикаментов, повышения эффективности планирования поставок и предотвращения дефицита препаратов. Это особенно важно для медицинских учреждений и аптек, где точное прогнозирование спроса является ключевым фактором обеспечения доступности лекарств для пациентов. Проведенное исследование разработки гибридной модели можно распространить на другие категории лекарственных средств с более сложной структурой спроса.

Таким образом, данное исследование подтверждает перспективность использования гибридных моделей временных рядов для решения задач прогнозирования в фармацевтической отрасли, обеспечивая их практическую применимость и высокую точность.

#### **Список литературы**

1. Национальный центр общественного здравоохранения Министерства здравоохранения Республики Казахстан. (2023). Сердечно-сосудистые заболевания [Электронный ресурс]. URL: <https://hls.kz/ru/archives/41346> (дата обращения: 15.01.2025)

2. Национальный центр общественного здравоохранения Министерства здравоохранения Республики Казахстан. (2023). Анализ здоровья казахстанцев за 2022 год [Электронный ресурс]. URL: <https://hls.kz/ru/archives/41153> (дата обращения: 15.01.2025)
3. Мбониншуту, Ф., Нкурунзиза, Дж., Нийобухунги, Дж., & Кайитаре, Э. (2022). Прогнозирование спроса на основные лекарственные средства: машинное обучение на основе данных о потреблении в Руанде. *Процессы*, 10(1). <https://doi.org/10.3390/pr10010026>.
4. Элсворт, С., & Гюттель, С. (2020). Прогнозирование временных рядов с использованием сетей LSTM: символичный подход. *Архив*, abs/2003.05672. <https://doi.org/10.48550/arXiv.2003.05672>
5. Моуса, Б., & Аль-Хатиб, Б. (2023). Прогнозирование спроса на лекарства с использованием методов глубокого обучения: обзор. *Журнал интеллектуальных систем*, 32, 20220297. <https://doi.org/10.1515/jisys-2022-0297>
6. Гонсалес, Р., Гонсалес, Л., & Санчес, А. (2021). Прогнозирование спроса на лекарства фармацевтической организации с использованием модели ARIMA. *Университет и общество*, 13, 119–130.
7. Daultbakov, V., Ruzieva, E., & Balkhybekova, K. (2024). Metody analiza vremennykh ryadov i prognozirovaniya dlya otsenki sezonnoi sostavlyayushchei zheleznodorozhnogo transporta. *Vestnik KazATK*, 131(2), 260–277. <https://doi.org/10.52167/1609-1817-2024-131-2-260-277>.
8. Dzhetsmekova, S., & Shukirova, A. (2024). Determinirovannye i stokhasticheskie komponenty vremennykh ryadov: modelirovanie elektropotrebleniya. *Vestnik KazATK*, 130(1), 163–170. <https://doi.org/10.52167/1609-1817-2024-130-1-163-170>
9. Захра, А., & Путра, Ю. Х. (2019). Сравнение методов прогнозирования на основе сезонных паттернов для предсказания потребности в лекарствах с использованием метода ARIMA и экспоненциального сглаживания. *Серия конференций IOP: Материаловедение и инженерия*, 662, 022030. <https://doi.org/10.1088/1757-899X/662/2/022030>
10. Буринскиене, А. (2022). Модель прогнозирования: случай фармацевтической розничной торговли. *Медицина будущего*, 9, 582186. <https://doi.org/10.3389/fmed.2022.582186>
11. Ван Ахмад, В. К. А., & Ахмад, С. (2013). Модель ARIMA и метод экспоненциального сглаживания: сравнение. *Материалы конференции AIP*, 1522, 1312–1321. <https://doi.org/10.1063/1.4801282>
12. Тас, И., & Сатоглу, С. И. (2023). Прогнозирование спроса в фармацевтической отрасли в условиях пандемии COVID-19 с использованием машинного обучения и анализа временных рядов. *Интеллектуальные и нечеткие системы*, 759, 157–165. [https://doi.org/10.1007/978-3-031-39777-6\\_19](https://doi.org/10.1007/978-3-031-39777-6_19)
13. Хайндман, Р. Дж., & Афанасопулос, Г. (2018). Прогнозирование: Принципы и практика (2-е изд.). Австралия: OText
14. Фаттах, Дж., Эзин, Л., Аман, З., Эль Муссами, Х., & Лачхаб, А. (2018). Прогнозирование спроса с использованием модели ARIMA. *Международный журнал инженерного бизнеса и управления*, 10, 1–9. <https://doi.org/10.1177/1847979018808673>
15. Дай, Дж., & Чен, С. (2019). Применение модели ARIMA для прогнозирования данных о численности населения. *Журнал физики: Серия конференций*, 1324, 012100. <https://doi.org/10.1088/1742-6596/1324/1/012100>
16. Нирмала, В. В., Харджади, Д., & Авалуддин, Р. (2021). Прогнозирование продаж с использованием метода экспоненциального сглаживания и метода тренда для оптимизации продаж продукции в РТ. *Zamrud Bumi Indonesia в условиях пандемии COVID-19. Международный журнал инженерии, науки и информационных технологий*, 1(4), 59–64. <https://doi.org/10.52088/ijesty.v1i4.16>

## References

1. Natsional'nyi tsentr obshchestvennogo zdavookhraneniya Ministerstva zdavookhraneniya Respubliki Kazakhstan. (2023). Serdechno-sosudistye zabolvaniya [Elektronnyi resurs]. URL: <https://hls.kz/ru/archives/41346> (data obrashcheniya: 15.01.2025)
2. Natsional'nyi tsentr obshchestvennogo zdavookhraneniya Ministerstva zdavookhraneniya Respubliki Kazakhstan. (2023). Analiz zdorov'ya kazakhstantsev za 2022 god [Elektronnyi resurs]. URL: <https://hls.kz/ru/archives/41153> (data obrashcheniya: 15.01.2025)
3. Mboninshuti, F., Nkurunziza, Dzh., Niiobukhungiro, Dzh., & Kaiitare, E. (2022). Prognozirovaniye sprosa na osnovnye lekarstvennyye sredstva: mashinnoe obuchenie na osnove dannykh o potreblenii v Ruande. *Protsessy*, 10(1). <https://doi.org/10.3390/pr10010026>
4. Elsvort, S., & Gyuttel', S. (2020). Prognozirovaniye vremennykh ryadov s ispol'zovaniem setei LSTM: simvol'nyi podkhod. *Arkhiv*, abs/2003.05672. <https://doi.org/10.48550/arXiv.2003.05672>
5. Mousa, B., & Al'-Khatib, B. (2023). Prognozirovaniye sprosa na lekarstva s ispol'zovaniem metodov glubokogo obucheniya: obzor. *Zhurnal intellektual'nykh sistem*, 32, 20220297. <https://doi.org/10.1515/jisys-2022-0297>
6. Gonsales, R., Gonsales, L., & Sanches, A. (2021). Prognozirovaniye sprosa na lekarstva farmatsevticheskoi organizatsii s ispol'zovaniem modeli ARIMA. *Universitet i obshchestvo*, 13, 119–130.
7. Dauletbakov, B., Ruzieva, E., & Balkhybekova, K. (2024). Metody analiza vremennykh ryadov i prognozirovaniya dlya otsenki sezonnoi sostavlyayushchei zhelezodorozhnogo transporta. *Vestnik KazATK*, 131(2), 260–277. <https://doi.org/10.52167/1609-1817-2024-131-2-260-277>
8. Dzhetsmekova, S., & Shukirova, A. (2024). Determinirovannyye i stokhasticheskie komponenty vremennykh ryadov: modelirovaniye elektropotrebleniya. *Vestnik KazATK*, 130(1), 163–170. <https://doi.org/10.52167/1609-1817-2024-130-1-163-170>
9. Zakhra, A., & Putra, Yu. Kh. (2019). Sravneniye metodov prognozirovaniya na osnove sezonnykh patternov dlya predskazaniya potrebnosti v lekarstvakh s ispol'zovaniem metoda ARIMA i eksponentsial'nogo sglazhivaniya. *Seriya konferentsii IOP: Materialovedeniye i inzheneriya*, 662, 022030. <https://doi.org/10.1088/1757-899X/662/2/022030>
10. Burinskiene, A. (2022). Model' prognozirovaniya: sluchai farmatsevticheskoi roznichnoi trgovli. *Medsitina budushchego*, 9, 582186. <https://doi.org/10.3389/fmed.2022.582186>
11. Van Akhmad, V. K. A., & Akhmad, S. (2013). Model' ARIMA i metod eksponentsial'nogo sglazhivaniya: sravneniye. *Materialy konferentsii AIP*, 1522, 1312–1321. <https://doi.org/10.1063/1.480128>
12. Tas, I., & Satoglu, S. I. (2023). Prognozirovaniye sprosa v farmatsevticheskoi otrasli v usloviyakh pandemii COVID-19 s ispol'zovaniem mashinnogo obucheniya i analiza vremennykh ryadov. *Intellektual'nye i nechetkie sistemy*, 759, 157–165. [https://doi.org/10.1007/978-3-031-39777-6\\_19](https://doi.org/10.1007/978-3-031-39777-6_19)
13. Khaindman, R. Dzh., & Afanasopulos, G. (2018). Prognozirovaniye: *Printsipy i praktika* (2-e izd.). Avstraliya: OTexts.
14. Fattakh, Dzh., Ezzin, L., Aman, Z., El' Mussami, Kh., & Lachkhab, A. (2018). Prognozirovaniye sprosa s ispol'zovaniem modeli ARIMA. *Mezhdunarodnyi zhurnal inzhenernogo biznesa i upravleniya*, 10, 1–9. <https://doi.org/10.1177/1847979018808673>
15. Dai, Dzh., & Chen, S. (2019). Primeneniye modeli ARIMA dlya prognozirovaniya dannykh o chislennosti naseleniya. *Zhurnal fiziki: Seriya konferentsii*, 1324, 012100. <https://doi.org/10.1088/1742-6596/1324/1/012100>
16. Nirmala, V. V., Khardzhadi, D., & Avaluddin, R. (2021). Prognozirovaniye prodazh s ispol'zovaniem metoda eksponentsial'nogo sglazhivaniya i metoda trenda dlya optimizatsii prodazh produktsii v PT. *Zamrud Bumi Indonesia v usloviyakh pandemii COVID-19*.

*Mezhdunarodnyi zhurnal inzhenerii, nauki i informatsionnykh tekhnologii*, 1(4), 59–64.  
<https://doi.org/10.52088/ijesty.v1i4.169>

## ДӘРІЛІК ЗАТТАРДЫ ТҰТЫНУДЫҢ ТАРИХИ ДЕРЕКТЕРІ НЕГІЗІНДЕ СҰРАНЫСТЫ БОЛЖАУ МОДЕЛЬДЕРІН ТАЛДАУ

**Аңдатпа.** Дәрілік заттарға сұранысты болжау пациенттердің тұрақты жеткізілімін, қорларды тиімді басқаруын және өмірлік маңызды дәрілерге уақтылы қол жеткізуін қамтамасыз етуде шешуші рөл атқарады. Бұл мақалада антигипертензиялық қысымға қарсы дәрілік затқа сұранысты болжау мақсатында ARIMA мен экспоненциалды тегістеу уақыт қатарлары әдістерін зерттеу ұсынылған. Зерттеу таңдалған деректер жиынтығына негізделген болжаудың жоғары дәлдігі мен тиімділігін қамтамасыз ететін модельді әзірлеуге және анықтауға бағытталған.

Талдау болжау әдістерін зерттеуді, деректерді жинауды және өңдеуді, әрбір әдіс үшін оңтайлы параметрлерді анықтауды, гибридті модельді әзірлеуді, берілген көрсеткіштерге негізделген дәлдікті бағалауды және нәтижелерді талдауды қамтиды. Зерттеу барысында болжаудың ең тиімді әдістері уақыт қатарына негізделген тәсілдер, соның ішінде ARIMA модельдері мен экспоненциалды тегістеу әдістері екендігі анықталды. Ал әзірленген гибридті модель екі тәсілдің артықшылықтарын біріктіру арқылы болжамдардың жоғары дәлдігін береді. Нәтижелер гибридті модель ARIMA мен экспоненциалды тегістеуден барлық негізгі көрсеткіштер бойынша басым екенін көрсетеді. Алынған қорытындылар негізінде фармацевтика саласында сұранысты болжау дәлдігін арттыру үшін гибридті модельді енгізу ұсынылды.

**Түйін сөздер:** болжау, болжау әдістері, экспоненциалды тегістеу, ARIMA, уақыт қатарлары, дәрілік зат, гибридті модель.

## ANALYSING DEMAND FORECASTING MODELS BASED ON HISTORICAL DRUG CONSUMPTION DATA

**Abstract.** Forecasting drug demand plays a key role in ensuring sustainable supply, effective inventory management, and timely patient access to life-saving medicines. This article presents a study of ARIMA time series and exponential smoothing methods for predicting demand for an antihypertensive blood pressure drug. The research is aimed at developing and identifying a model that ensures high accuracy and efficiency of forecasting based on selected data sets.

The analysis includes the study of forecasting methods, data collection and processing, determining the optimal parameters for each method, developing a hybrid model, evaluating accuracy based on specified metrics, and analyzing the results. In the course of the conducted research, it was found that the most effective forecasting methods are time series-based approaches, including ARIMA models and exponential smoothing methods. And the developed hybrid model demonstrates high forecast accuracy by combining the advantages of the two approaches. The results show that the hybrid model is superior to ARIMA and exponential smoothing in all key metrics. Based on the findings, the introduction of hybrid models is proposed to improve the accuracy of forecasting demand in the pharmaceutical industry.

**Keywords:** forecasting, forecasting methods, exponential smoothing, ARIMA, time series, drug, hybrid model.

**Сведения об авторах**

Құмарқанова Ақбота Серікқызы	Магистр, преподаватель «Школа цифровых технологий и искусственного интеллекта», НАО «Восточно-Казахстанский технический университет имени Д. Серикбаева», г. Усть-Каменогорск, Казахстан, тел: 87474371007; e-mail: <a href="mailto:akbota.vko@gmail.com">akbota.vko@gmail.com</a>
Хасенова Зарина Толлеубековна	PhD, декан «Школа цифровых технологий и искусственного интеллекта», НАО «Восточно-Казахстанский технический университет имени Д. Серикбаева», г. Усть-Каменогорск, Казахстан, тел: 87715492087; E-mail: <a href="mailto:zhasenova@mail.ru">zhasenova@mail.ru</a>
Вайс Юрий Андреевич	Кандидат технических наук, ассоциированный профессор «Школа цифровых технологий и искусственного интеллекта», НАО «Восточно-Казахстанский технический университет имени Д. Серикбаева», г. Усть-Каменогорск, Казахстан, тел: 87052502872; E-mail: <a href="mailto:yuvais@edu.ektu.kz">yuvais@edu.ektu.kz</a>

**Авторлар туралы мәлімет**

Құмарқанова Ақбота Серікқызы	Магистр, «Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті» КеАҚ-нің «Цифрлық технологиялар және жасанды интеллект мектебі» оқытушысы, Өскемен қ., Қазақстан, тел: 87474371007; E-mail: <a href="mailto:akbota.vko@gmail.com">akbota.vko@gmail.com</a>
Хасенова Зарина Толлеубековна	PhD, «Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті» КеАҚ-нің «Цифрлық технологиялар және жасанды интеллект мектебі» деканы, Өскемен қ., Қазақстан, тел: 87715492087; E-mail: <a href="mailto:zhasenova@mail.ru">zhasenova@mail.ru</a>
Вайс Юрий Андреевич	Техника ғылымдарының кандидаты, «Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті» КеАҚ-нің «Цифрлық технологиялар және жасанды интеллект мектебі» қауымдастырылған профессоры, Өскемен қ., Қазақстан, тел: 87052502872; E-mail: <a href="mailto:yuvais@edu.ektu.kz">yuvais@edu.ektu.kz</a>

**Information about the authors**

Kumarkanova Akbot	Master, Lecturer, School of Digital Technology and Artificial Intelligence, NCJSC «D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan, Phone: 87474371007; E-mail: <a href="mailto:akbota.vko@gmail.com">akbota.vko@gmail.com</a>
Khasenova Zarina	PhD, Dean, School of Digital Technology and Artificial Intelligence, NCJSC «D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan, Phone: 87715492087; E-mail: <a href="mailto:zhasenova@mail.ru">zhasenova@mail.ru</a>
Weiss Yuri	Candidate of Technical Sciences, Associate Professor, School of Digital Technology and Artificial Intelligence, NCJSC «D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan. Phone: 87052502872; E-mail: <a href="mailto:yuvais@edu.ektu.kz">yuvais@edu.ektu.kz</a>

УДК: 519.6

МРНТИ 50.05, 50.41

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_9](https://doi.org/10.53364/24138614_2025_36_1_9)

Д.Е.Койшиева<sup>1</sup>, М.А.Сыдыбаева<sup>2</sup>, С.А.Бельгинова<sup>3\*</sup>,  
А.М. Жаксыбаев<sup>4</sup>, Ж.Е.Ерсаинова<sup>5</sup>

<sup>1,2</sup>Алматинский университет энергетики и связи имени Г. Даукеева,  
Алматы, Казахстан

<sup>3</sup>Университет «Туран» Алматы, Казахстан

<sup>4</sup>Шакарим университет, Семей, Казахстан

<sup>5</sup>Восточно-Казахстанский технический университет им. Д. Серикбаева,  
Усть-Каменогорск, Казахстан

E-mail: [sbelginova@gmail.com](mailto:sbelginova@gmail.com)\*

## ИССЛЕДОВАНИЕ ВЛИЯНИЯ МЕТОДОВ РАСШИРЕНИЯ ДАННЫХ И ШУМОВ НА ТОЧНОСТЬ СЕГМЕНТАЦИИ ЯДЕР В ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОЙ НЕЙРОННОЙ СЕТИ

**Аннотация.** *Сегментация медицинских изображений является одной из ключевых задач биомедицинской информатики, особенно в контексте диагностики онкологических заболеваний. Особенно актуальной является задача точной сегментации ядер клеток, что необходимо для идентификации морфологических изменений в тканях. Однако, основным вызовом и проблематикой в данной области остаются изменчивость качества изображений, ограниченные объемы доступных данных и необходимость обеспечения высокой точности и устойчивости моделей глубокого обучения. В данном исследовании был проведен анализ влияния десяти техник аугментации данных, включая добавление случайного шума различной интенсивности, на производительность модели сегментации ядер клеток. Также в эксперименте особое внимание уделено техникам, использующим искусственное добавление шума, моделирующее реальные условия, такие как вариации освещения, артефакты и дефекты при подготовке образцов. Для добавления шума использовались два вида: аддитивный гауссов шум и равномерный случайный шум с интенсивностью от минус 0.05 до 0.05 продемонстрировали значительное влияние на обобщающую способность модели, улучшая её устойчивость к гетерогенным данным. Используемая архитектура была основана на модифицированной модели UNet с внедрением модуля СВМ, который акцентирует внимание модели на значимых областях изображения. Модуль был добавлен в декодирующую часть архитектуры модели. Для экспериментов были объединены два эталонных набора данных CryoNuSeg и MoNuSeg с предварительной обработкой, включающей разбиение на патчи и преобразование формата изображений. Примененные техники аугментации включали горизонтальное и вертикальное отражение, случайное вращение, изменение контрастности, эластичную деформацию и добавление шума. Кроме того, в ходе экспериментов были исследованы комбинации аугментации. Результаты исследования показали, что техники аугментации горизонтального и вертикального отражения и случайного вращения достигли минимальных потерь, при этом точность моделей превышала девяносто процентов. Техника изменения контрастности продемонстрировала наиболее сбалансированную производительность, обеспечив средний коэффициент пересечения над объединением 0.928 и точность 92.2 процента. Данные по результатам исследования подчеркивают важность использования искусственного добавления шума как метода повышения*

*устойчивости модели к артефактам и вариативности данных и также использование комбинированных техник аугментации. В перспективе комбинированные техники аугментации могут стать основой для разработки адаптивных алгоритмов, способных эффективно работать с гетерогенными биомедицинскими изображениями.*

**Ключевые слова.** *Гистологические изображения, рак, сегментация ядер, сверточные нейронные сети, модуль внимания, расширение данных.*

### **Введение.**

Сегментация медицинских изображений представляет собой одну из ключевых задач биомедицинской информатики и компьютерного зрения [1]. Способность автоматизированных систем точно выделять ядра клеток или границы органов на сложных гистологических изображениях открывает перспективы для ускорения диагностики. В Казахстане рак остается серьезной проблемой, с высокими показателями заболеваемости и смертности [2]. По данным Global Cancer Statistics, в 2022 году было зарегистрировано почти 20 миллионов новых случаев рака и 9,7 миллиона смертей от него, что подчеркивает необходимость в эффективных методах диагностики [3]. Способность автоматизированных систем глубокого обучения выделять ядра клеток или границы органов на сложных гистологических изображениях открывает перспективы для ускорения диагностики рака [4]. Тем не менее, процесс сегментации ядер и клеток в гистологических изображениях сталкивается с рядом вызовов. Во-первых, качество изображений может варьироваться в зависимости от условий их получения, таких как различия в окрашивании, микроскопии и методах подготовки образцов, что затрудняет обобщение моделей. Во-вторых, доступные наборы данных обычно ограничены, что усложняет обучение глубоких нейронных сетей, которым для достижения высокой точности и устойчивости к вариативности данных требуется большой объем информации [5]. Одним из подходов к улучшению способности нейросетей к обобщению является расширение данных — техника, позволяющая искусственно расширить обучающий набор посредством различных преобразований оригинальных изображений [6]. На ограниченных наборах данных расширение способствует повышению качества сегментации, однако влияние отдельных методов расширения данных на модели сегментации изучено недостаточно. Важным аспектом исследования становится добавление искусственного шума, который моделирует реальные условия, такие как неидеальное освещение или дефекты при подготовке образцов [7].

В данном исследовании был проведен анализ влияния десяти различных техник аугментации и различных уровней шума на задачу сегментации гистологических клеток с использованием модифицированной архитектуры Глубокой Нейронной Сети (ГНС) [8]. В данной работе рассматривается применение данной модели, специально адаптированной для задачи сегментации, с целью оценки её производительности и стабильности в различных условиях данных.

### **Материалы и методы исследования.**

Сегментация ядер с использованием модели ГНС представляет собой важный метод в области биомедицинской визуализации. Среди основных подходов выделяются архитектуры UNet и ее усовершенствованные варианты, такие как UNet++ и Attention UNet, которые широко используются для сегментации на уровне пикселей [9]. В данной работе представлена модифицированная архитектура UNet с внедренным блоком внимания СВМ (Convolutional Block Attention Module) [10]. Блок СВМ включает канальное и пространственное внимание, усиливающее фокус на значимых областях изображения. Модель имеет общий объем параметров, равный 34 573 825, что соответствует 131.89 МВ и обеспечивает достаточную емкость для обучения.



В настоящем исследовании были использованы два набора данных: MoNuSeg Challenge и CryoNuSeg [11, 12]. Набор данных MoNuSeg Challenge включает 30 обучающих изображений с аннотированными границами ядер, что в общей сложности составляет 21 623 ядра. Эти изображения представлены в формате TIFF с разрешением 1000x1000 пикселей, полученные при увеличении 40x, и охватывают ткани семи различных органов. Набор данных *CryoNuSeg* представляет собой первый полностью аннотированный набор данных для сегментации ядер в криосрезах замороженных тканей, окрашенных гематоксилином и эозином. В этот набор входят изображения из 10 различных органов человека, таких как надпочечники, гортань, лимфатические узлы, средостение, поджелудочная железа, плевра, кожа, яички, тимус и щитовидная железа. Набор включает 30 изображений размером 512×512 пикселей в формате TIFF. Для целей исследования оба набора данных были объединены, переведены в формат PNG, а также разбиты на патчи размером до 256x256 пикселей, с метками 0 и 1 для дальнейшего использования в модели. Ниже на рисунке 1 представлены примеры с каждого набора.

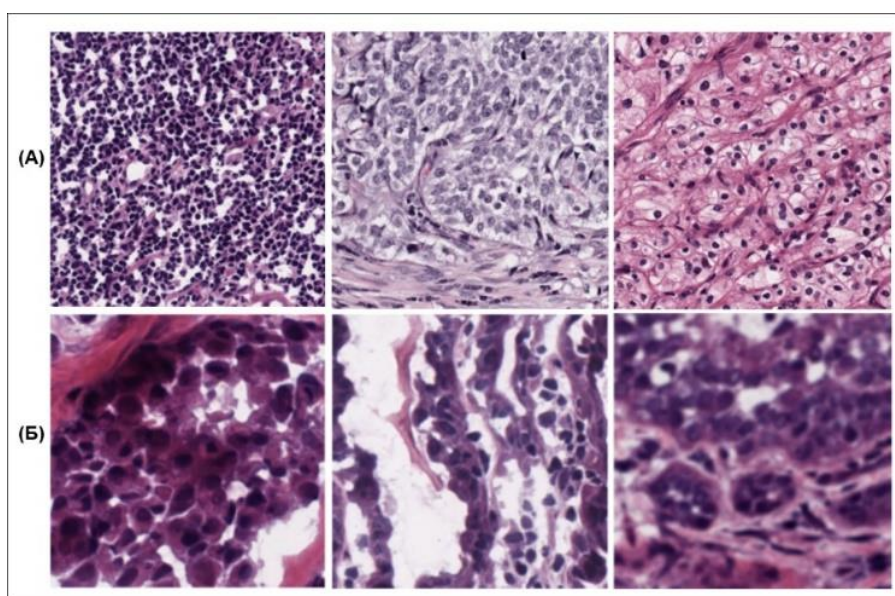


Рисунок 1 – Примеры с набора MoNuSeg Challenge (А) и CryoNuSeg (Б)

Для улучшения сегментации гистологических изображений применяются десять различных техник расширения данных, которые далее будут обозначены как *Техника-1* — *Техника-10*.

*Техника-1*: случайное масштабирование. Операция масштабирования изменяет размер объектов на изображении на коэффициент  $s$  в определенном диапазоне  $s \in [0.8, 1.2]$ .

*Техника-2*: горизонтальное и вертикальное отражение. Этот метод предполагает переворот изображения по горизонтальной и вертикальной осям.

*Техника-3*: регулировка яркости. Регулировка яркости изменяет интенсивность пикселей на коэффициент, произвольно выбранный из  $[0, 0.8, 1, 2]$ .

*Техника-4*: эластичная деформация. Эластичная деформация моделируется путем смещения пикселей в соответствии со сглаженным случайным полем смещений.

*Техника-5*: случайное размытие. Этот метод применяет к изображению гауссово размытие, используя фильтр Гаусса со стандартным отклонением  $=1.3$  [8].

*Техника-6*: случайное вращение. Для данного изображения случайным образом вращаются на угол, выбранный из диапазона  $[-30^\circ, 30^\circ]$ . Это помогает модели стать инвариантной к вращательным изменениям входных данных [7].



Техника-7: добавление случайного шума. Случайный шум добавляется к каждому пикселю изображения с параметром интенсивности шума  $p$ , выбранным из диапазона  $p \in [-0.05, 0.05]$ .

Техника-8: добавление случайного шума. Случайный шум добавляется к каждому пикселю изображения с параметром интенсивности шума  $p$ , выбранным из диапазона  $p \in [-0.05, 0.05]$ .

Техника-9: случайное изменение контрастности. Контраст изображения изменяется на коэффициент  $k$ , выбранный из диапазона  $k \in [0.5, 1.5]$ . Это улучшает способность модели распознавать изображения с разной контрастностью.

Техника-10: обрезка и изменение положения. Операция случайного обрезания изменяет размер изображения на заданный коэффициент  $s$ , выбранный из диапазона  $s \in [0.75, 1.0]$ , а затем сдвигает его на случайное смещение  $\Delta x$  и  $\Delta y$ , чтобы изменить положение объекта в тензоре. Данная техника увеличивает вариативность положения объектов и помогает модели быть более устойчивой к изменению положения объектов на изображении.

В данном исследовании мы оценивали эффективность модели по четырем ключевым метрикам: Точность, Чувствительность, Потери и Среднее пересечение над объединением (Средняя IoU). Чтобы оценить стабильность каждой метрики, мы также рассчитали стандартное отклонение, которое показывает разброс значений метрик для различных техник расширения данных. Чем ниже стандартное отклонение, тем стабильнее результаты модели для конкретной метрики [13].

Точность — это отношение правильно предсказанных образцов к общему числу образцов, формула представлена ниже:

$$\text{Точность} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

где TP - истинно положительные результаты, TN - истинно отрицательные результаты, FP - ложноположительные результаты, FN - ложноотрицательные результаты.

Средний IoU, также называемый индексом Жаккара, измеряет степень совпадения между предсказанной сегментацией и истинной. Средний IoU усредняется по всем классам. Формула для IoU выглядит следующим образом:

$$\text{Средняя IOU} = \frac{TP}{TP + FP + FN} \quad (2)$$

Функция потерь оценивает разницу между предсказанной сегментацией и истинной. В данном исследовании мы использовали бинарную кросс-энтропийную потерю, которая определяется как:

$$\text{Потери} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (3)$$

где  $y_i$  - истинная метка,  $p_i$  - предсказанная вероятность, а  $N$  - общее количество пикселей в изображении [14].

Чувствительность, измеряет долю правильно идентифицированных положительных результатов:

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

Эксперименты выполнялись с использованием на языке Python 3.7 и фреймворков машинного обучения TensorFlow 2.4.1 и Keras 2.4.1 для построения и обучения модели. Техники расширения изображений осуществлялась библиотекой Albumentations 0.5.2, предварительная обработка изображений проводилась с помощью OpenCV 4.2.0 и Pillow 7.0.0. Для численных вычислений использовались библиотеки NumPy версии 1.18.1 и SciPy 1.4.1, а для обработки данных — Pandas 1.2.0. Визуализация данных выполнялась с

использованием Matplotlib 3.1.3 и Seaborn 0.10.0. Обучение модели проводилось на графическом процессоре NVIDIA Tesla V100, что позволило значительно ускорить вычисления.

**Результаты и их обсуждение.** Для данного эксперимента модель обучалась на 160 эпохах. В ходе анализа было установлено, что после 160 эпох модель демонстрировала стабильные результаты, а увеличение количества эпох не приводило к значительному улучшению индикаторов. Начальная скорость обучения была установлена на уровне 0.001, с течением времени использовался коэффициент затухания для уменьшения скорости обучения. Размер батча был выбран равным 16 для всех экспериментов. Изначальный набор данных включал 60 изображений, объединенных из наборов данных MoNuSeg Challenge и CryoNuSeg. После разбиения на патчи размером 256x256 пикселей и применения аугментации, объем данных был увеличен до 132 патчей, обеспечив разнообразие данных для более устойчивого обучения модели. Набор данных был разделен на обучающий, проверочный и тестовый наборы в соотношении 80% для обучения, 10% для проверки и 10% для тестирования. В таблице 1 показаны результаты.

Таблица 1 – Результаты производительности модели ГНС на валидационных данных для различных техник аугментации

<i>Техники расширения данных</i>	<i>Потери</i>	<i>Точность</i>	<i>Средняя IOU</i>	<i>Чувствительность</i>
<i>Техника-1</i>	0.0840	0.9282	0.8986	0.8382
<i>Техника-2</i>	0.0631	0.9272	0.9223	0.8546
<i>Техника-3</i>	0.0796	0.9260	0.8980	0.8514
<i>Техника-4</i>	0.0847	0.8757	0.9106	0.8475
<i>Техника-5</i>	0.2341	0.9289	0.7590	0.7955
<i>Техника-6</i>	0.0655	0.9020	0.8166	0.8475
<i>Техника-7</i>	0.1498	0.9239	0.9044	0.8518
<i>Техника-8</i>	0.0803	0.9193	0.9106	0.8188
<i>Техника-9</i>	0.0655	0.9220	0.9280	0.8395
<i>Техника-10</i>	0.1597	0.9270	0.9170	0.8321

На основании данных, представленных в Таблице 1, проведен анализ производительности модели ГНС на валидационном наборе при использовании различных техник аугментации.

Точность и Средняя IoU продемонстрировали различное влияние техник аугментации на качество сегментации. Наибольшая точность была достигнута при использовании техники 5 (0.9289), однако данная техника показала относительно низкое значение Средней IoU (0.7590), что указывает на недостаточную устойчивость модели. В то же время, техника 9 обеспечила высокое значение Средней IoU (0.9280) при высокой точности (0.9220), что свидетельствует о хорошей способности модели к обобщению на задачах сегментации.

Анализ Потерь выявил, что наименьшие значения этой метрики достигнуты при использовании техник 2 (0.0631) и 6 (0.0655), что указывает на меньшие ошибки модели в процессе валидации и более высокую стабильность данных подходов. Эти техники также продемонстрировали высокие значения точности и средней IoU, подтверждая их эффективность для решения поставленной задачи.

Что касается Чувствительности, наивысшее значение было достигнуто с помощью техники 2 (0.8546), что указывает на превосходную способность модели точно идентифицировать положительные сегменты. Техники 3, 4 и 7 также продемонстрировали

высокие показатели чувствительности (более 0.85), что подчеркивает их значимость для задач, где требуется точное обнаружение истинных положительных сегментов. В целом, результаты показывают, что техника 9 является наиболее сбалансированной, так как обеспечивает высокие значения по всем основным метрикам. На Рисунке 2 представлена диаграмма, отображающая стандартное отклонение по основным метрикам.

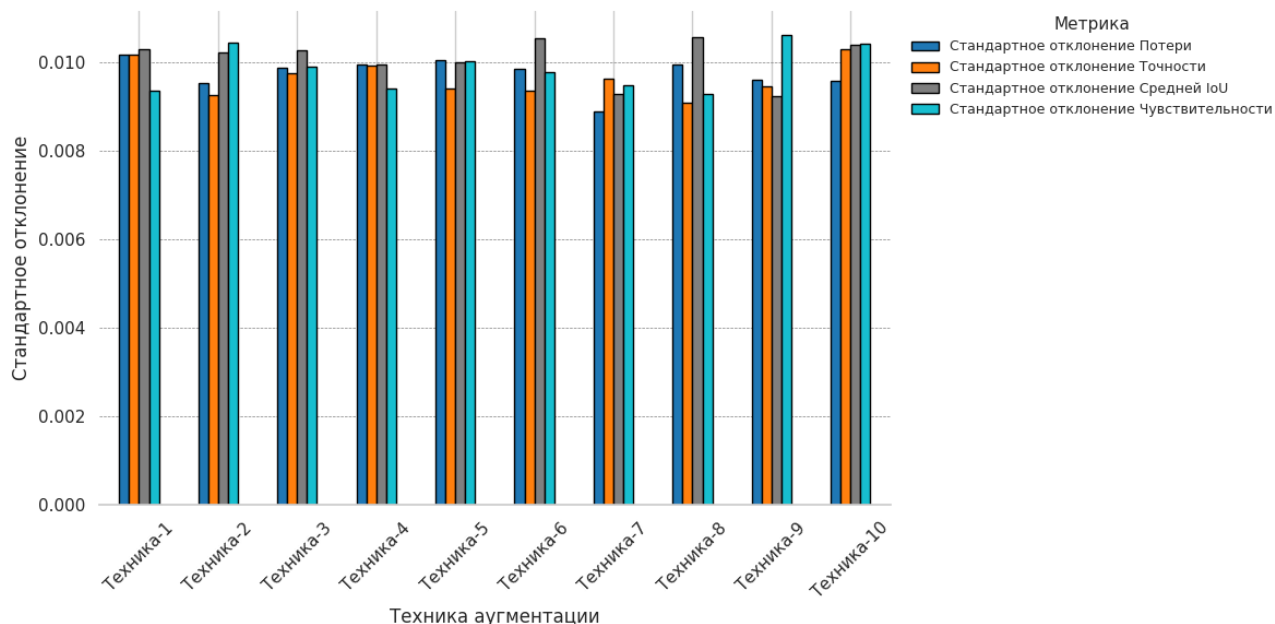


Рисунок 2 – Сравнение стандартного отклонения метрик для разных техник расширения данных

Техники 2 и 6 демонстрируют минимальные значения стандартного отклонения для всех метрик, что говорит об их стабильности и способности модели показывать более согласованные результаты. В то же время техники 3 и 7 показывают более высокие значения стандартного отклонения по ряду метрик, указывая на большую изменчивость и менее устойчивую производительность модели при их применении. Напротив, техники 3 и 7 характеризуются более выраженной вариативностью значений, особенно на ранних этапах обучения. Это может быть связано с тем, что данные техники приводят к увеличению внутриклассового разброса, что усложняет процесс оптимизации параметров модели. Несмотря на то, что данные техники обеспечивают высокий уровень чувствительности, их повышенное стандартное отклонение указывает на возможность появления нестабильных предсказаний на тестовой выборке. Дополнительно был проведен сравнительный анализ времени обучения моделей при использовании различных техник расширения данных. Результаты показывают, что техники 2 и 6 не только способствуют улучшению сегментации, но и позволяют достичь высокой точности за меньшее число эпох. В отличие от них, техники 3 и 7 потребовали большего количества эпох для достижения аналогичных показателей, что может подтвердить сложность их оптимизации. Ниже на Рисунке 3 представлена динамика изменения индикаторов по эпохам.

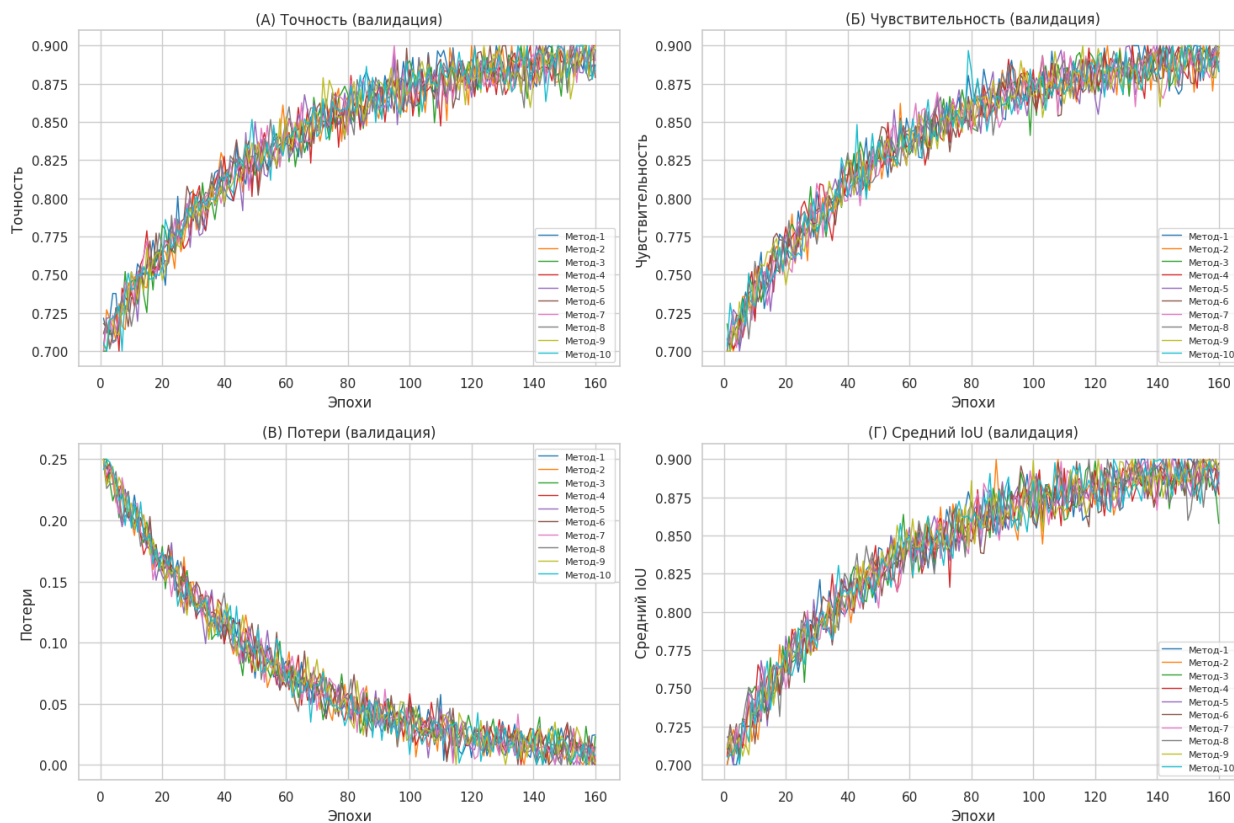


Рисунок 3 – Анализ изменений метрик сегментации по эпохам для различных техник расширения данных

В данном исследовании установлено, что различные техники аугментации расширения данных оказывают существенное влияние на производительность и устойчивость модели в задаче сегментации биомедицинских изображений. Наиболее стабильные и воспроизводимые результаты были достигнуты при использовании техник 2 и 6, которые продемонстрировали минимальные значения стандартного отклонения. В то же время техника 9 показала сбалансированную эффективность по ключевым метрикам, что свидетельствует о её высокой способности к обобщению. Техники с более высоким стандартным отклонением, такие как 3 и 7, характеризуются значительной вариативностью, что может быть полезно для задач, требующих значительного увеличения разнообразия данных.

В перспективе представляется целесообразным проведение исследований, направленных на тестирование комбинированных техник аугментации для достижения оптимального баланса между разнообразием данных и стабильностью модели. Ожидается, что такие комбинации могут существенно повысить способность модели к обобщению, особенно при обработке гетерогенных биомедицинских данных, требующих адаптивного подхода.

Кроме того, будущие исследования могут быть сосредоточены на адаптивной настройке параметров расширения данных для каждой техники в рамках комбинаций.

### Заключение.

В данном исследовании установлено, что техники расширения данных существенно влияют на производительность модели сегментации ядер клеток на гистологических изображениях. Техники 2 и 6 продемонстрировали минимальные значения стандартного отклонения и высокую стабильность, с потерями 0.0631 и 0.0655 соответственно,

обеспечивая точность более 90% и среднюю IoU свыше 0.81. Техника 9 показала сбалансированные результаты с точностью 92.2% и средней IoU 0.9280. Эти данные свидетельствуют о важности выбора оптимальных методик аугментации для повышения точности и воспроизводимости сегментации, что может стать основой для разработки более надежных инструментов ранней диагностики и количественного анализа онкологических заболеваний.

### Список литературы

1. Conze, P. -H., Andrade-Miranda, G., Singh, V. K., Jaouen, V. & Visvikis, D. (2023). Current and Emerging Trends in Medical Image Segmentation with Deep Learning. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 7(6), 545-569, <https://doi.org/10.1109/TRPMS.2023.3265863>.
2. Akhmedullin, R., Aimyshev, T., Zhakhina, G., Yerdessov, S., Beyembetova, A., Ablayeva, A., Biniyazova, A., Seyil, T & Gaipov, A. (2024). In-depth Analysis and Trends of Cancer Mortality in Kazakhstan: a Joinpoint Analysis of Nationwide Healthcare Data 2014–2022. *BMC Cancer* 24, 1340 <https://doi.org/10.1186/s12885-024-13128-2>
3. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R., Soerjomataram, I. & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 74(3), 229-263. <https://doi.org/10.3322/caac.21834>
4. Rai, H. M., Yoo, J. & Razaque, A. (2024). Comparative Analysis of Machine Learning and Deep Learning Models for Improved Cancer Detection: A Comprehensive Review of Recent Advancements in Diagnostic Techniques. *Expert Systems with Applications*, 255, 124838. <https://doi.org/10.1016/j.eswa.2024.124838>
5. Gabdullin, M.T., Mukasheva, A., Koishiyeva, D. Umarov, T., Bissembayev, A., Kim, K-S. & Won, K., (2024). Automatic Cancer Nuclei Segmentation on Histological Images: Comparison Study of Deep Learning Methods. *Biotechnology and Bioprocess Engineering*, 29, 1034–1047 (2024). <https://doi.org/10.1007/s12257-024-00130-5>
6. Xu, M., Yoon, S., Fuentes, A. & Park, D. A Comprehensive Survey of Image Augmentation Techniques for Deep Learning. *Pattern Recognition*, 137(1), 109347. <https://doi.org/10.1016/j.patcog.2023.109347>
7. Mukasheva, A., Koishiyeva, D., Suimenbayeva, Z., Rakhmetulayeva, S., Bolshibayeva, A., & Sadikova, G. (2023). Comparison evaluation of unet-based models with noise augmentation for breast cancer segmentation on ultrasound image. *Eastern-European Journal of Enterprise Technologies*, 5(9 (125), 85–97. <https://doi.org/10.15587/1729-4061.2023.289044>
8. Zhao, X., Wang, L., Zhang, Y. Han, X., Devenci, M. & Parmar, M. (2024). A Review of Convolutional Neural Networks in Computer Vision. *Artificial Intelligence Review*, 57(4), 57-99. <https://doi.org/10.1007/s10462-024-10721-6>
9. Mukasheva, A., Koishiyeva, D., Sergazin, G., Sydybayeva, M., Mukhammejanova, D., & Seidazimov, S. (2024). Modification of U-Net with Pre-Trained ResNet-50 and Atrous Block for Polyp Segmentation: Model TASPP-UNet. *Engineering Proceedings*, 70(1), 16. <https://doi.org/10.3390/engproc2024070016>
10. Zhu, C., Cheng, K., & Hua, X. (2024). A Medical Image Segmentation Network with Multi-Scale and Dual-Branch Attention. *Applied Sciences*, 14(14), 6299. <https://doi.org/10.3390/app14146299>
11. Mahbod, A., Schaefer, G., Bancher, B., Löw, C., Dorffner, G., Ecker, R. & Ellinger, I. (2023). CryoNuSeg: A dataset for nuclei instance segmentation of cryosectioned H&E-stained histological images. *Computers In Biology and Medicine*, 132, 104349. <https://doi.org/10.1016/j.compbiomed.2021.104349>

12. Kumar, N., Verma, R., Anand, D., Zhou, Y., Onder, O. & Tsougenis, E. (2019). A Multi-Organ Nucleus Segmentation Challenge. *IEEE Transaction on Medical Imaging*, 39(5), 1380-1391. <https://doi.org/10.1109/TMI.2019.2947628>

13. Jadon S. A Survey of Loss Functions for Semantic Segmentation. (2020). *IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. – IEEE, 1-7. <https://doi.org/10.1109/CIBCB48159.2020.9277638>

14. Nguyen, T-H., Pham, T-L., Tran, Q-V., Le, T-L., Tran, T-H. & Vu, H. (2023). Systematic Evaluation of Loss Functions for Ovarian Tumors Segmentation from Ultrasound Images. *1st International Conference on Health Science and Technology (ICHST)*. 1-6. <https://doi.org/10.1109/ICHST59286.2023.10565336>

## ГИСТОЛОГИЯЛЫҚ БЕЙНЕЛЕРДЕГІ ЯДРОЛАРДЫ СЕГМЕНТАЦИЯЛАУ ДӘЛДІГІНЕ ДЕРЕКТЕРДІ КЕҢЕЙТУ ӘДІСТЕРІ МЕН ШУДЫҢ ӘСЕРІН ТЕРЕҢ НЕЙРОНДЫҚ ЖЕЛІ ҚОЛДАНУ АРҚЫЛЫ ЗЕРТТЕУ

**Аңдатпа.** Медициналық кескіндерді сегменттеу биомедициналық информатиканың негізгі міндеттерінің бірі болып табылады, әсіресе онкологиялық ауруларды диагностикалау контекстінде. Тіндердегі морфологиялық өзгерістерді анықтау үшін қажет жасуша ядроларын дәл сегментациялау міндеті ерекше өзекті болып табылады. Алайда, бұл саладағы негізгі сын-қатер мен проблемалар кескін сапасының өзгергіштігі, қол жетімді деректердің шектеулі көлемі және терең оқыту модельдерінің жоғары дәлдігі мен тұрақтылығын қамтамасыз ету қажеттілігі болып қала береді. Бұл зерттеу жасуша ядросының сегменттеу моделінің өнімділігіне әртүрлі қарқындылықтағы кездейсоқ шуды қосуды қоса алғанда, деректерді күшейтудің он әдісінің әсерін талдады. Сондай-ақ, экспериментте жарықтандырудың өзгеруі, артефактілер және үлгілерді дайындаудағы ақаулар сияқты нақты жағдайларды модельдейтін жасанды шу қосуды қолданатын әдістерге ерекше назар аударылады. Шуды қосу үшін екі түрі қолданылды: аддитивті Гаусс шу және минус 0.05-тен 0.05-ке дейінгі қарқындылықтағы біркелкі кездейсоқ шу модельдің жалпылау қабілетіне айтарлықтай әсер етіп, оның гетерогенді деректерге төзімділігін жақсартты. Пайдаланылған архитектура Svat модулін енгізе отырып, модификацияланған UNet моделіне негізделген, ол модельдің кескінің маңызды аймақтарына назар аударады. Модуль модель архитектурасының декодтау бөлігіне қосылды. Эксперименттер үшін *suposeg* және *topuseg* екі анықтамалық деректер жиынтығы патчтарға бөлуді және кескін пішімін түрлендіруді қамтитын алдын ала өңдеумен біріктірілді. Қолданылған күшейту әдістеріне көлденең және тік шағылысу, кездейсоқ айналу, контрастты өзгерту, серпінді деформация және Шу қосу кірді. Сонымен қатар, эксперименттер күшейту комбинацияларын зерттеді. Зерттеу нәтижелері көлденең және тік шағылысу мен кездейсоқ айналуды күшейту әдістері модельдердің дәлдігі тоқсан пайыздан асатын минималды шығындарға жеткенін көрсетті. Контрастты өзгерту әдісі ең теңдестірілген өнімділікті көрсетті, бұл 0.928 бірлестігінің орташа қиылысу коэффициентін және 92.2 пайыздық дәлдікті қамтамасыз етті. Зерттеу нәтижелері бойынша мәліметтер модельдің артефактілерге төзімділігі мен деректердің өзгергіштігін арттыру әдісі ретінде жасанды шу қосуды қолданудың, сондай-ақ біріктірілген күшейту әдістерін қолданудың маңыздылығын көрсетеді. Болашақта біріктірілген күшейту әдістері гетерогенді биомедициналық кескіндермен тиімді жұмыс істей алатын адаптивті алгоритмдерді әзірлеуге негіз бола алады.

**Түйін сөздер.** Гистологиялық кескіндер, қатерлі ісік, ядро сегментациясы, конволюциялық нейрондық желілер, зейін модулі, деректерді кеңейту.

**INVESTIGATION OF THE IMPACT OF DATA AUGMENTATION METHODS AND NOISE ON THE ACCURACY OF NUCLEI SEGMENTATION IN HISTOLOGICAL IMAGES USING A DEEP NEURAL NETWORK**

**Abstract.** Segmentation of medical images is one of the key tasks of biomedical informatics, especially in the context of cancer diagnosis. The task of precise segmentation of cell nuclei is particularly relevant and necessary to identify morphological changes in tissues. However, the main challenges and problems in this area remain the variability of image quality, limited amounts of available data, and the need to ensure the high accuracy and stability of deep learning models. This study analysed the impact of ten data augmentation techniques, including the addition of random noise of varying intensity, on the cell nucleus segmentation model's performance. The experiment also specialises in techniques using artificial noise addition, simulating real-world conditions such as lighting variations, artefacts, and defects during sample preparation. Two types of noise were used to add: additive Gaussian noise and uniform random noise with an intensity from minus 0.05 to 0.05, which demonstrated a significant effect on the generalizing ability of the model, improving its resistance to heterogeneous data. The architecture used was based on a modified UNet model with the introduction of the CBAM module, which focuses the model's attention on significant areas of the image. The module has been added to the decoding part of the model architecture. For the experiments, two reference datasets CryoNuSeg and MoNuSeg were combined with preprocessing, including patching and image format conversion. Applied augmentation techniques included horizontal and vertical reflection, random rotation, contrast change, elastic deformation, and noise addition. In addition, augmentation combinations were investigated during the experiments. The study results showed that the techniques of horizontal and vertical reflection augmentation and random rotation achieved minimal losses, while the accuracy of the models exceeded ninety per cent. The contrast change technique demonstrated the most balanced performance, providing an average intersection ratio over the union of 0.928 and an accuracy of 92.2 per cent. The data from the study results emphasize the importance of using artificial noise addition to increase the model's resistance to artefacts and data variability, as well as the use of combined augmentation techniques. In the future, combined augmentation techniques may become the basis for the development of adaptive algorithms capable of effectively working with heterogeneous biomedical images

**Keywords.** Histological images, cancer, nuclei segmentation, convolutional neural networks, attention module, augmentation.

**Авторлар туралы мәлімет**

Койшиева Дина Еркиновна	Ғ.Даукеев атындағы Алматы энергетика және коммуникация университеті, М094-Ақпараттық технологиялар мамандығының магистранты, Алматы, Қазақстан, E-mail: <a href="mailto:d.koishiyeva@aes.kz">d.koishiyeva@aes.kz</a>
Сыдыбаева Мадина Алданышовна	Ғ.Дәукеев атындағы Алматы энергетика және коммуникация университетінің ақпараттық жүйелер және киберқауіпсіздік кафедрасының аға оқытушысы, Алматы, Қазақстан, E-mail: <a href="mailto:m.sydybaeva@alt.edu.kz">m.sydybaeva@alt.edu.kz</a>
Бельгинова Сауле	PhD, Тұран Университетінің қауымдастырылған профессоры, Алматы қ., Қазақстан, E-mail: <a href="mailto:sbelginova@gmail.com">sbelginova@gmail.com</a>
Жаксыбаев Айдын Мадиярович	«Технологиялық машиналар мен жабдықтар» мамандығы бойынша магистр, Шәкәрім университеті, «ОлжаПРОЕКТ» ЖШС басшысы, Өскемен, Қазақстан, E-mail: <a href="mailto:olzha2010@mail.ru">olzha2010@mail.ru</a>
Ерсаинова Жансая	«АТЖЗЖМ» оқытушысы, Д.Серікбаев атындағы ШҚТУ, Өскемен қ., Қазақстан, E-mail: <a href="mailto:Zhansaya.ersayynova@mail.ru">Zhansaya.ersayynova@mail.ru</a>

**Сведение об авторах**

Койшиева Дина Еркиновна	Магистрант по специальности М094-Информационные технологии, Алматинский университет энергетики и связи имени Г.Даукеева, Алматы, Казахстан, E-mail: <a href="mailto:d.koishiyeva@aes.kz">d.koishiyeva@aes.kz</a>
Сыдыбаева Мадина Алданышовна	Старший преподаватель кафедры «Информационных систем и кибербезопасности», Алматинский университет энергетики и связи имени Г.Даукеева, Алматы, Казахстан, E-mail: <a href="mailto:m.sydybaeva@alt.edu.kz">m.sydybaeva@alt.edu.kz</a>
Бельгинова Сауле	PhD, ассоц. профессор, Университета Туран, к.т.н. г. Алматы, Казахстан, E-mail: <a href="mailto:sbelginova@gmail.com">sbelginova@gmail.com</a>
Жаксыбаев Айдын Мадиярович	Магистр по специальности «Технологические машины и оборудование», Шакарим университет, Руководитель ТОО «ОлжаПРОЕКТ» г. Усть-Каменогорск, Казахстан, E-mail: <a href="mailto:olzha2010@mail.ru">olzha2010@mail.ru</a>
Ерсаинова Жансая	Преподаватель «ШИТиИС», ВКТУ им.Д. Серикбаева, г. Усть-Каменогорск, Казахстан, E-mail: <a href="mailto:Zhansaya.ersayynova@mail.ru">Zhansaya.ersayynova@mail.ru</a>

**Information about the authors**

Koyshieva Dina	Master's student in the specialty M094-Information Technology, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, E-mail: <a href="mailto:d.koishiyeva@aes.kz">d.koishiyeva@aes.kz</a>
Sydybaeva Madina	Senior Lecturer, Department of Information Systems and Cybersecurity, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, E-mail: <a href="mailto:m.sydybaeva@alt.edu.kz">m.sydybaeva@alt.edu.kz</a>
Belginova Saule	PhD, Associate Professor of Turan University, Almaty, Kazakhstan, E-mail: <a href="mailto:sbelginova@gmail.com">sbelginova@gmail.com</a>
Zhaksybaev Aidyn	Master's degree in Technological Machines and Equipment, Shakarim University, Head of «OlzhaПРОЕКТ» LLP, Ust-Kamenogorsk, Kazakhstan, E-mail: <a href="mailto:olzha2010@mail.ru">olzha2010@mail.ru</a>
Ersainova Zhansaya	Teacher of " SITaIS ", D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: <a href="mailto:Zhansaya.ersayynova@mail.ru">Zhansaya.ersayynova@mail.ru</a>



УДК 004.912

МРНТИ 81.93.29

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_10](https://doi.org/10.53364/24138614_2025_36_1_10)М. Мехдиев<sup>1</sup>, А.К. Шайханова<sup>1\*</sup>, Г.Б. Бекешова<sup>1</sup>, И.Е. Икласова<sup>2</sup>, К.С. Бакенова<sup>1</sup><sup>1</sup>Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан<sup>2</sup>Северо-Казахстанский университет им. Манаша Козыбаева, Петропавловск, Казахстан

E-mail: shaikhanova\_ak@enu.kz\*

## МЕТОДЫ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ТЕКСТОВ (TEXT MINING)

**Аннотация.** В статье рассматриваются методы интеллектуальной обработки текстов (Text Mining), которые позволяют преобразовывать слабоструктурированные текстовые данные в структурированную и легко анализируемую информацию. С ростом объемов данных в цифровую эпоху Text Mining становится незаменимым инструментом анализа текстов в различных сферах. Эти технологии находят широкое применение в информационной безопасности, где анализ текстов помогает выявлять угрозы и аномалии, в здравоохранении — для обработки медицинских записей и извлечения диагностической информации, в маркетинге — для анализа потребительских предпочтений, а также в юридической практике, где автоматизация анализа документов повышает точность и снижает временные затраты.

В статье подробно рассматриваются как традиционные статистические методы, такие как TF-IDF, Word2Vec, Latent Dirichlet Allocation (LDA), так и современные подходы, включая модели глубокого обучения на основе архитектуры трансформеров, например BERT, GPT и их производные. Современные методы демонстрируют значительные успехи в учёте контекста, анализе семантики и извлечении скрытых смыслов из текстов, что делает их незаменимыми для решения сложных задач.

Особое внимание уделено сравнению эффективности различных методов и их применимости в задачах автоматизации. Описаны возможности интеграции Text Mining для анализа больших объемов данных, выявления закономерностей и автоматизации процессов извлечения знаний. Представленные результаты исследования подчеркивают актуальность использования этих технологий для повышения эффективности работы специалистов, ускорения процессов анализа информации и решения задач в ключевых отраслях, что открывает новые перспективы для внедрения интеллектуальных систем обработки данных.

**Ключевые слова:** Text Mining, интеллектуальная обработка текстов, машинное обучение, обработка естественного языка, TF-IDF, Word2Vec, BERT, GPT, автоматизация анализа текстов.

### Введение.

Современная эпоха цифровых технологий характеризуется быстрым ростом объема текстовой информации, создаваемой и распространяемой в различных областях деятельности, начиная от научных исследований и заканчивая социальными сетями и корпоративными документами. Этот огромный объем данных содержит ценную информацию, которую можно извлекать, анализировать и применять для решения разнообразных задач. Однако традиционные методы обработки данных оказываются недостаточно эффективными для работы с текстами, так как они не способны адекватно учитывать лингвистическую структуру и семантические особенности текста.

Интеллектуальная обработка текстов, также известная как Text Mining, представляет собой междисциплинарное направление, включающее методы обработки естественного языка, машинного обучения и статистического анализа. Основная цель Text Mining — преобразование слабоструктурированных текстовых данных в структурированную информацию, которая может быть использована для анализа и принятия решений. Применение технологий Text Mining позволяет решать такие задачи, как автоматическая классификация и аннотирование текстов, выделение ключевых тематических направлений, поиск скрытых закономерностей и построение предсказательных моделей на основе текстовых данных.

Актуальность исследования методов Text Mining обусловлена необходимостью автоматизации анализа текстовой информации, что особенно важно в таких областях, как информационная безопасность, медицина, маркетинг и право. Автоматизация позволяет существенно снизить трудозатраты и повысить точность анализа, что в условиях роста объемов информации становится ключевым фактором успеха. Настоящая статья направлена на обзор и анализ современных методов Text Mining, включая традиционные статистические методы, такие как TF-IDF и Word2Vec, и новейшие контекстно-зависимые модели, такие как BERT и GPT, для глубокого анализа текстов.

### **Материалы и методы исследования.**

Для выполнения исследования использовались различные методы интеллектуальной обработки текстов. На этапе подготовки данных проводился сбор слабоструктурированных текстов из открытых источников, таких как базы данных угроз, классификаторы и реестры, с последующей их предобработкой: удаление стоп-слов, лемматизация и нормализация текстов. Для представления текстовых данных применялись как традиционные подходы, такие как TF-IDF для выделения значимых слов и Word2Vec для создания плотных векторных представлений слов, так и современные модели глубокого обучения, включая трансформеры BERT и GPT, способные учитывать контекст и семантическую многозначность текста.

В процессе исследования использовались задачи тематической классификации текстов, автоматического реферирования, кластеризации и создания текстов с заданным содержанием, а также поиск информации по ключевым словам. Для моделирования применялись искусственные нейронные сети, обученные на количественных показателях актуальности угроз, что позволило интегрировать их с большим объемом текстовых данных.

Материалами исследования выступали текстовые корпуса, содержащие описания тактик и техник злоумышленников, взятые из специализированных баз данных, таких как MITRE ATT&CK или CVE. Для работы с текстами применялись программные библиотеки для обработки естественного языка, включая NLTK, spaCy, TensorFlow и PyTorch. Обработка больших массивов данных осуществлялась с использованием вычислительных мощностей, таких как GPU и облачные платформы, обеспечивающие эффективное выполнение задач глубокого обучения. Экспериментальная часть исследования опиралась на данные, представленные в ряде научных и прикладных работ по теме Text Mining.

### **Результаты и их обсуждение.**

Традиционно применяемые базы данных (классификаторы, реестры) содержат обширные текстовые описания угроз, уязвимостей и тактик (техник) злоумышленников. Несмотря на явные преимущества таких баз данных, работа по оценке и анализу угроз и уязвимостей на конкретном предприятии остаётся экспертно ориентированной, что требует значительных временных затрат, высокой когнитивной нагрузки и специализированных знаний. В исследовании [1] предлагается использовать количественные показатели

актуальности угроз информационной безопасности для создания искусственной нейронной сети (НС) и подачи этих факторов в качестве входных данных. В то же время современные технологии искусственного интеллекта позволяют обрабатывать большие массивы текстовых данных на естественном языке (ЕЯ), решая следующие задачи:

- поиск информации, по ключевым словам;
- тематическая классификация текстов;
- автоматическое реферирование;
- кластеризация текстов по содержанию;
- автоматическое создание текстов с заданным содержанием и т.д.

Этим вопросам посвящено множество исследований, таких как [2, 3, 4]. Направление анализа слабоструктурированных текстовых данных на ЕЯ, известное как интеллектуальный анализ текстов (Text Mining), продемонстрировало значимые результаты в работах [5, 6]. Переход к обработке текстов на естественном языке требует выполнения ряда этапов: подготовки корпуса текстовых данных, их векторного представления в многомерном семантическом пространстве и, наконец, решения задач семантического анализа.

Использование методов для создания векторных представлений слов и оценки их семантической близости является ключевым в анализе слабоструктурированных текстов на ЕЯ [7]. В современных подходах машинного обучения для таких задач применяются как классические методы (линейные методы классификации, гауссовские модели, деревья решений), так и методы обработки последовательностей (скрытые марковские модели, модели условных случайных полей). В последние годы большую популярность получили нейронные сети: многослойные перцептроны, сверточные и рекуррентные НС, для которых векторные представления слов создаются с помощью таких инструментов, как TF-IDF, Word2Vec, Doc2Vec, GloVe, FastText и других.

Прежде чем перейти к обсуждению методов машинного обучения, важно рассмотреть основные этапы предварительной обработки текстовых данных. Эти этапы включают:

1. Нормализация — упрощение текста, удаление пунктуации, аббревиатур, служебных слов (например, союзов и предлогов).
2. Токенизация — разбиение текста на слова и предложения.
3. Стеммизация — приведение слов к их корню, удаление суффиксов и окончаний.
4. Лемматизация — приведение слов к их канонической форме (например, инфинитиву или именительному падежу).
5. Фильтрация — удаление нерелевантных символов и слов.

При этом следует учитывать, что данный алгоритм не сохраняет порядок слов в предложении.

Метрика TF-IDF часто применяется для анализа текстов и информационного поиска, например, для определения релевантности документа запросу, а также при кластеризации документов.

Word2Vec [8] — инструмент, разработанный группой исследователей Google под руководством Т. Миколова в 2013 году, представляет собой набор алгоритмов для создания векторных представлений слов. На вход подаётся текстовый корпус, а на выходе формируются вектора слов. Инструмент обучается на большом объеме текстов, запоминая контексты, в которых встречается каждое слово. По завершении обучения каждому слову соответствует вектор в 300-мерном пространстве признаков (семантическом пространстве), где слова, близкие по смыслу, располагаются рядом. В Word2Vec реализованы две основные архитектуры: Continuous Bag of Words (CBOW) и Skip-gram.

CBOW и Skip-gram — нейросетевые модели Word2Vec, описывающие процесс обучения и формирования векторных представлений слов. В CBOW модель предсказывает слово по контексту, в то время как Skip-gram предсказывает контекст по заданному слову.

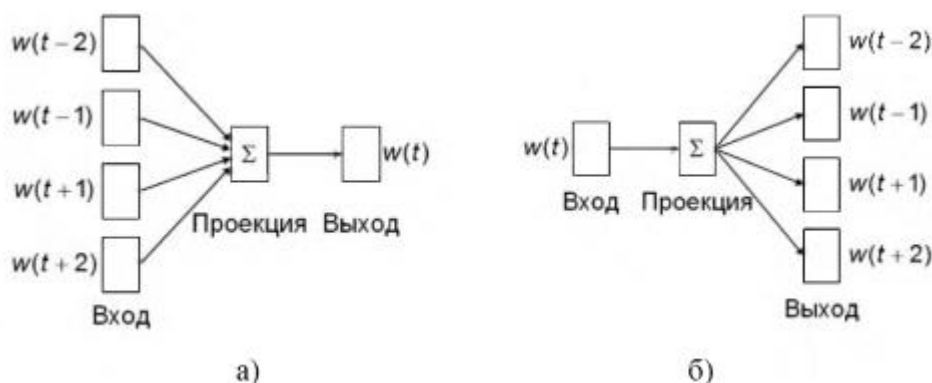


Рисунок 1 – Архитектуры алгоритмов обучения Continuous Bag of Words (а) и Skip-gram (б)

Хотя Word2Vec эффективно создает векторные представления для слов, у него имеются существенные ограничения: 1) алгоритм не учитывает порядок слов, что влияет на смысл фраз; 2) не принимает во внимание семантические значения; 3) не поддерживает формирование вектора для всего документа. Для решения этих проблем был создан инструмент Doc2Vec.

Doc2Vec (paragraph2vec) [9] — это инструмент, разработанный в 2014 году на основе Word2Vec, который позволяет формировать пространство признаков фиксированной длины для документов. Такой вектор можно применять для анализа семантической близости текстов (абзацев, документов) и использования в задачах кластеризации, классификации и прогнозирования.

Перечисленные модели относятся к первой волне обработки естественного языка (NLP). Вторая революция в NLP связана с развитием нейросетевых языковых моделей, таких как GPT и BERT, основанных на механизме внимания (Self-Attention), которые обозначаются как «трансформеры».

Первая версия GPT (Generative Pre-trained Transformer) [10], выпущенная OpenAI в 2018 году (создатели — Илон Маск и Сэм Альтман), стала прорывом в генеративных языковых моделях на архитектуре трансформера. Третья версия, GPT-3, выпущенная в 2020 году, была обучена на более чем 500 ГБ текстовых данных. Для русского языка существует аналогичная модель ruGPT-3 13B, разработанная компаниями SberDevices и SberCloud, обученная на 600 ГБ данных. Однако полное использование GPT требует мощных вычислительных ресурсов, что недоступно большинству исследователей.

Модель BERT (Bidirectional Encoder Representations from Transformers) [11] — двунаправленная трансформерная модель, предназначенная для предобучения на крупных текстовых корпусах для последующего применения в широком спектре NLP-задач. Мультиязычные версии BERT предобучены на больших наборах данных и могут быть легко интегрированы в проекты, позволяя избегать трудоёмкого обучения модели с нуля. BERT также может быть запущен как на локальном компьютере, так и на бесплатных облачных платформах, таких как Google Colab. В отличие от GPT, BERT имеет более лёгкую структуру: например, версия BERT-Large включает 24 слоя и 240 млн параметров. Модель BERT поддерживает тонкую настройку для конкретных задач, что делает её подходящей и для анализа данных в области ИБ. В отличие от Word2Vec и других традиционных языковых моделей, BERT генерирует контекстозависимые представления, что позволяет учитывать контекст и генерировать уникальные векторы для омонимов в зависимости от контекста.

Архитектура трансформера, представленного на рисунках, состоит из кодировщика (encoder) и декодировщика (decoder), каждый из которых включает слои с механизмом внутреннего внимания и НС прямого распространения (Feed-Forward NN).



Рисунок 2 – Общая схема архитектуры трансформера

Кодировщик преобразует входные слова в векторы (эмбединги) в семантическом пространстве, а декодировщик генерирует последовательность выходных слов для выполнения конкретных задач, таких как классификация, поиск или перевод. Механизмы внимания выделяют ключевые слова в тексте, формируя содержательное представление текста и оптимизируя обработку.

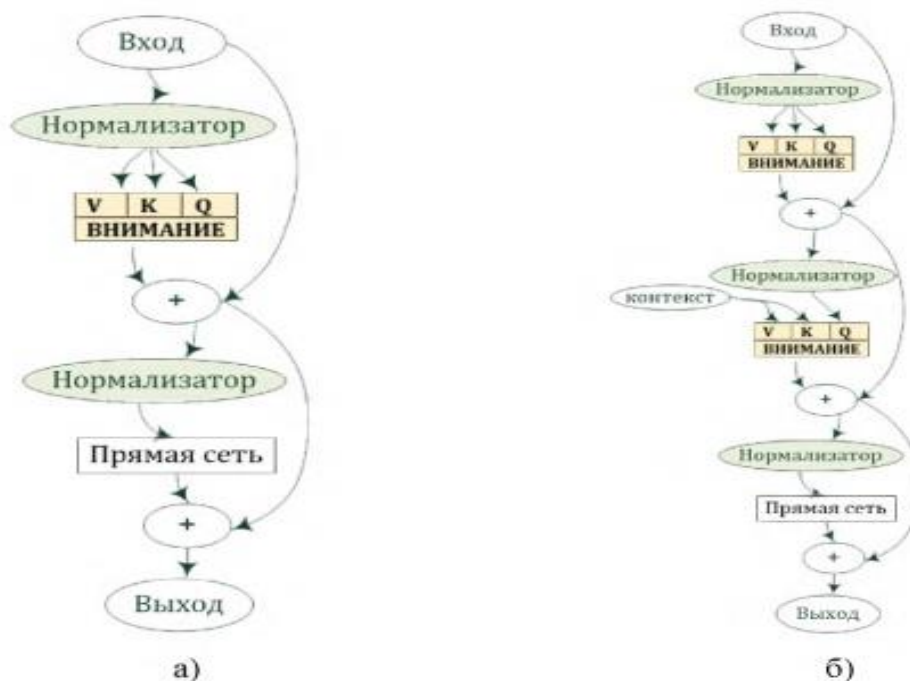


Рисунок 3 – Схема архитектуры трансформера: а) кодирующий слой; б) декодирующий слой

Методы векторизации, такие как TF-IDF, Word2Vec и Doc2Vec, относятся к статистическим. Эти методы часто используются в задачах NLP, однако они имеют определённые недостатки:

– не учитывают многозначность и контекст использования слов, что приводит к созданию одного усреднённого вектора для омонимов и может исказить результаты анализа;

– недостаточно эффективны при обработке редких или новых слов.

Контекстуализированные модели, такие как GPT и BERT, избавлены от этих недостатков: они вычисляют векторное представление слова с учётом его контекста. В исследовании [12] проводился сравнительный анализ векторизации текстов с использованием статистических моделей и одной из трансформерных моделей для анализа диалоговых чат-ботов. Результаты показали высокую эффективность как комбинации TF-IDF + Word2Vec, так и трансформерных моделей.

Алгоритмы кластеризации широко применяются для анализа текстовых описаний и определения сходства между документами. Основная их задача заключается в выявлении скрытой структуры текстовых корпусов на основе признаков, позволяющих группировать текстовые данные по смыслу. Центральное место в решении этой задачи занимает оценка семантической близости между векторными представлениями текстов, визуализируемыми как точки в пространстве признаков. Для этого используются метрики семантической близости, или метрики расстояний [13]. Суть этих метрик заключается в вычислении расстояния между точками (векторами) в многомерном пространстве, причём более близкие точки обозначают более схожие текстовые описания.

Кластеризация представляет собой ключевой этап анализа текстовой информации, когда текстовые векторы, полученные после предварительной обработки, делятся на группы по общим признакам [14].

Существует несколько подходов к реализации алгоритмов кластеризации:

1. Иерархическая кластеризация. Эти алгоритмы создают дендрограмму вложенных кластеров, где каждый новый кластер строится из существующих, образуя древовидную структуру. На каждом шаге вычисляется расстояние между кластерами и обновляется для вновь образованных кластеров. Расстояние  $R(U, V)$  между кластерами  $U$  и  $V$ , содержащими несколько элементов, может определяться различными функциями, выбор которых зависит от задачи.

2. Плоская кластеризация предполагает разбиение объектов сразу на несколько кластеров, к одному из которых будет принадлежать каждый объект. Самый распространённый алгоритм данного типа —  $k$ -means ( $k$ -средних). Он случайным образом определяет центры кластеров и находит ближайшие к ним векторы данных. Затем центры кластеров смещаются, и процедура повторяется, пока кластеры не стабилизируются. Недостатки  $k$ -means: 1) количество кластеров задаётся вручную; 2) алгоритм чувствителен к выбору начальных центров; 3) не позволяет элементу принадлежать к нескольким кластерам.

3. Чёткая кластеризация предполагает, что каждый объект принадлежит только одному кластеру, тогда как нечёткая кластеризация допускает принадлежность объекта к нескольким кластерам с определённой вероятностью. К нечетким алгоритмам относится  $c$ -means ( $c$ -средних). Однако этот метод не всегда корректен при разном разбросе значений по осям элементов в кластерах.

При кластеризации текстов одной из ключевых проблем является высокая размерность данных [15], что усложняет процесс машинного обучения. Чтобы справиться с этой проблемой, используются методы снижения размерности, описанные в ряде отечественных и зарубежных исследований [16, 17]. Процедура основывается на выборе наиболее значимых признаков, что позволяет:

- улучшить визуализацию результатов;
- повысить точность классификации;
- сократить ресурсы на вычисления;

– ускорить обучение моделей.

Снижение размерности векторного пространства является одним из эффективных методов для экономии вычислительных ресурсов при сохранении информативности результатов [18].

Популярные алгоритмы для снижения размерности:

– t-SNE (t-distributed stochastic neighbor embedding) [19] — стохастическое вложение соседей с использованием t-распределения Стьюдента. Этот итерационный процесс визуализирует данные в низкоразмерных пространствах, сохраняя информацию об относительном расположении точек, что способствует улучшению поиска глобальных минимумов и визуализации данных. Несмотря на низкую скорость работы, t-SNE хорошо сохраняет базовую структуру данных.

– UMAP (Uniform Approximation and Projection) [20] — высокоэффективный метод для снижения размерности, который строит взвешенный граф, соединяя ближайших соседей (объекты) ребрами. Граф в низкоразмерном пространстве оптимизируется так, чтобы приблизиться к исходному графу путем минимизации дивергенции Кульбака-Лейблера. UMAP отличается высокой скоростью и отсутствием ограничений на исходную размерность пространства, превосходя t-SNE по быстродействию.

– PCA (Principal Component Analysis) [21] — метод анализа главных компонент, выполняющий линейное преобразование данных в новую систему координат, где можно описать вариации данных меньшим числом измерений. PCA характеризуется высокой скоростью вычислений, однако это достигается за счёт некоторой потери информации.

Методы, применяемые в тематической классификации текстов:

– LSA (Latent Semantic Analysis) [22] — латентный семантический анализ, который позволяет выявлять скрытые взаимосвязи между документами и терминами. Он предполагает, что слова с близким значением чаще встречаются в схожих текстах. Для анализа создается матрица, где строки представляют слова, а столбцы — текстовые описания. Путем сингулярного разложения (SVD) уменьшается количество строк, сохраняя при этом структурные связи между столбцами.

– LDA (Latent Dirichlet Allocation) [23] — метод латентного распределения Дирихле, также применяемый в тематическом моделировании. В отличие от LSA, LDA предполагает, что каждый документ представляет собой вероятностный набор тем, а каждое слово в документе ассоциировано с одной из этих тем.

На основании анализа представленных исследований можно заключить, что актуальность автоматизации обработки текстовых данных обусловлена потребностью в снижении трудоемкости и повышении эффективности работы специалистов по ИБ. Это особенно важно для задач анализа актуальных угроз ИБ и уязвимостей программного обеспечения объектов критической информационной инфраструктуры.

Технологии интеллектуального анализа текстов (Text Mining) представляют собой набор методов и алгоритмов для анализа и обработки текстовых данных. Text Mining охватывает широкий спектр задач, таких как классификация, кластеризация, аннотирование, выявление скрытых связей и прогнозирование на основе текстовой информации. Данные технологии становятся особенно востребованными в условиях растущих объемов текстовых данных, где требуется автоматизация анализа для эффективного извлечения знаний.

Технологии Text Mining открывают новые возможности для обработки текстовых данных и извлечения значимой информации, что позволяет автоматизировать анализ и значительно повысить эффективность работы специалистов в различных областях.

### Заключение.

Интеллектуальная обработка текстов (Text Mining) занимает важное место в современных подходах к анализу текстовой информации, представляя собой набор инструментов и методов, позволяющих автоматизировать процессы извлечения знаний из слабоструктурированных текстовых данных.

В условиях роста объемов информации технологии Text Mining становятся незаменимыми, обеспечивая эффективный анализ и структурирование данных в различных областях — от информационной безопасности и здравоохранения до права и маркетинга.

В данной статье были рассмотрены ключевые методы Text Mining, начиная с традиционных статистических моделей, таких как TF-IDF и Word2Vec, и заканчивая современными трансформерными моделями, такими как BERT и GPT. Каждый из методов имеет свои особенности и ограничения, которые важно учитывать при выборе подхода к решению конкретной задачи. Традиционные методы эффективны при обработке больших объемов данных с минимальными требованиями к вычислительным ресурсам, в то время как более современные модели обеспечивают высокую точность анализа за счет учета контекста, но требуют значительных вычислительных мощностей.

Преимущества использования Text Mining очевидны: автоматизация обработки текстов снижает когнитивную нагрузку на специалистов, повышает точность анализа и открывает возможности для глубокого понимания и прогнозирования на основе текстовой информации. Развитие технологий Text Mining будет способствовать дальнейшему росту эффективности и расширению применения этих методов в различных отраслях.

### Список литературы

1. Жук, Р. В., Дзьобан, П. И., & Власенко, А. В. (2020). Определение актуальности угроз информационной безопасности в информационных системах обработки персональных данных с использованием математического аппарата нейронных сетей. *Прикаспийский журнал: управление и высокие технологии*, 1(49), 169-178. <https://doi.org/10.21672/2074-1707.2020.49.4.169-178>
2. Гузаиров, М. Б., & Машкина, И. В. (2013). Управление защитой информации на основе интеллектуальных технологий. Москва: Машиностроение.
3. Коноваленко, С. А., & Королев, И. Д. (2016). Выявление уязвимостей информационных систем. *Инновации в науке*, 9(58), 12-20.
4. Аникин, И.В. (2015). Нечеткая оценка уязвимостей, основанная на метриках CVSS V.2.0. Проблемы информационной безопасности. *Компьютерные системы*, (3), 111-117.
5. Benjamin, V., Li, W., Holt, T., & Chen, H. (2015). Exploring threats and vulnerabilities in hacker web: Forums, IRC and carding shops. *2015 IEEE international conference on intelligence and security informatics (ISI)*, 85-90. IEEE.
6. Datta, P., Lodinger, N., Namin, A. S., & Jones, K. S. (2020). Cyber-attack consequence prediction. arXiv preprint arXiv:2012.00648.
7. Бондарчук, Д. В. (2017). Векторная модель представления знаний на основе семантической близости термов. *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*, 6(3), 73-83. <https://doi.org/10.14529/cmse170305>.
8. Python-School. (2024). Как работает Word2Vec: нейросети для NLP. [Электронный ресурс]. - Режим доступа: <https://python-school.ru/blog/what-is-word2vec/> (дата обращения: 26.10.2024).
9. Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *International Conference on Machine Learning*, 1188-1196.
10. Habr. (2024). GPT-2 в картинках (визуализация языковых моделей Трансформера). [Электронный ресурс]. - URL: <https://habr.com/ru/post/490842/> (дата обращения: 26.10.2024).



11. Che, W., Liu, Y., Wang, Y., Zheng, B., & Liu, T. (2018). Towards better UD parsing: Deep contextualized word embeddings, ensemble, and treebank concatenation. arXiv preprint arXiv:1807.03121.
12. Жеребцова, Ю. А., & Чижик, А. В. (2020). Сравнение моделей векторного представления текстов в задаче создания чат-бота. *Вестник Новосибирского государственного университета. Серия: Лингвистика и межкультурная коммуникация*, 18(3), 16-34. <https://doi.org/10.25205/1818-7935-2020-18-3-16-34>
13. Бенгфорт, Б., Билбро, Р., & Океда, Т. (2019). Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. СПб.: Питер.
14. Тюрин, А. Г., & Зуев, И. О. (2014). Кластерный анализ, методы и алгоритмы кластеризации. *Вестник МГТУ МИРЭА*, 2(3), 86-97.
15. Habr. (2024). Об одной задаче Data Science [Электронный ресурс]. - Режим доступа: <https://habr.com/ru/company/mlclass/blog/266727/> свободный (дата обращения: 26.10.2024).
16. Pramoditha, R. (2021). Dimensionality reduction techniques you should know in 2021. Towards Data Science. [Электронный ресурс]. - Режим доступа: <https://towardsdatascience.com/11-dimensionalityreduction-techniques-you-shouldknow-in-2021-dcb9500d388b>. - 11 (дата обращения: 26.10.2024).
17. Jia, W., Sun, M., Lian, J., & Hou, S. (2022). Feature dimensionality reduction: A review. *Complex & Intelligent Systems*, 8, 2663-2693.
18. Maitra, S. (2024). Feature reduction strategy to make better generalization models. Medium. [Электронный ресурс]. - Режим доступа: <https://medium.com/swlh/feature-selection-techniques-to-make-bettergeneralization-models-6a19dd6dc9b1> свободный (дата обращения: 25.10.2024).
19. Habr. (2024). Препарируем t-SNE [Электронный ресурс]. - Режим доступа: <https://habr.com/ru/post/267041/> (дата обращения: 25.10.2024).
20. McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
21. Золотых, Н.Ю (2013). Машинное обучение и анализ данных. [Электронный ресурс]. - Режим доступа: <http://www.uic.unn.ru:8103/~zny/ml/Course/04.Principal%20Component%20analysis.pdf> (дата обращения: 25.10.2024).
22. Habr. (2024). Латентно-семантический анализ: реализация [Электронный ресурс]. - Режим доступа: <https://habr.com/ru/post/240209/> (дата обращения: 25.10.2024).
23. Kanakogi, K., Washizaki, H., Fukazawa, Y., Ogata, S., Okubo, T., Kato, T., & Yoshioka, N. (2021). Tracing CVE vulnerability information to CAPEC attack patterns using natural language processing techniques. *Information*, 12(8), 298.

### References

1. Juk, R. V., Dz'oban, P. I., & Vlasenko, A. V. (2020). Opredelenie aktual'nosti ugroz informatsionnoi bezopasnosti v informatsionnykh sistemakh obrabotki personal'nykh dannykh s ispol'zovaniem matematicheskogo apparata neironnykh setei. *Prikladnyy zhurnal: upravlenie i vysokie tekhnologii*, 1(49), 169-178. <https://doi.org/10.21672/2074-1707.2020.49.4.169-178>
2. Guzairov, M. B., & Mashkina, I. V. (2013). *Upravlenie zashchitoy informatsii na osnove intellektual'nykh tekhnologii*. Moscow: Mashinostroenie.
3. Konovalenko, S. A., & Korolev, I. D. (2016). Vyivlenie uiazvimostei informatsionnykh sistem. *Innovatsii v nauke*, 9(58), 12-20.
4. Anikin, I. V. (2015). Nechetkaia otsenka uiazvimostei, osnovannaia na metrikakh CVSS V.2.0. Problemy informatsionnoi bezopasnosti. *Kompiuternye sistemy*, (3), 111-117

5. Benjamin, V., Li, W., Holt, T., & Chen, H. (2015). Exploring threats and vulnerabilities in hacker web: Forums, IRC and carding shops. 2015 *IEEE International Conference on Intelligence and Security Informatics (ISI)*, 85-90. IEEE. <https://doi.org/10.1109/ISI.2015.7165945>.
6. Datta, P., Lodinger, N., Namin, A. S., & Jones, K. S. (2020). Cyber-attack consequence prediction. arXiv preprint arXiv:2012.00648.
7. Bondarchuk, D. V. (2017). Vektornaiia model predstavleniia znaniia na osnove semanticheskoi blizosti termov. *Vestnik Iuzhno-Ural'skogo gosudarstvennogo universiteta. Serii: Vychislitel'naia matematika i informatika*, 6(3), 73-83. <https://doi.org/10.14529/cmse170305>
8. Python-School. (2024). Kak rabotaet Word2Vec: neirosети dlia NLP. Retrieved October 26, 2024, from <https://python-school.ru/blog/what-is-word2vec/>
9. Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *International Conference on Machine Learning*, 1188-1196.
10. Habr. (2024). GPT-2 v kartinkakh (vizualizatsiia iazykovykh modelei Transformera). Retrieved October 26, 2024, from <https://habr.com/ru/post/490842/>
11. Che, W., Liu, Y., Wang, Y., Zheng, B., & Liu, T. (2018). Towards better UD parsing: Deep contextualized word embeddings, ensemble, and treebank concatenation. arXiv preprint arXiv:1807.03121.
12. Zherebtsova, Y. A., & Chizhik, A. V. (2020). Sravnenie modelei vektornogo predstavleniia tekstov v zadache sozdaniia chat-bota. *Vestnik Novosibirskogo gosudarstvennogo universiteta. Serii: Lingvistika i mezhkul'turnaia kommunikatsiia*, 18(3), 16-34. <https://doi.org/10.25205/1818-7935-2020-18-3-16-34>.
13. Bengfort, B., Bilbro, R., & Okeda, T. (2019). Applied text analysis with Python: Machine learning for the natural language processing. St. Petersburg: Piter.
14. Tiurin, A. G., & Zuev, I. O. (2014). Klasternyi analiz, metody i algoritmy klasterizatsii. *Vestnik MGTU MIREA*, 2(3), 86-97.
15. Habr. (2024). Ob odnoi zadache Data Science. Retrieved October 26, 2024, from <https://habr.com/ru/company/mlclass/blog/266727/>
16. Pramoditha, R. (2021). Dimensionality reduction techniques you should know in 2021. Towards Data Science. Retrieved October 26, 2024, from <https://towardsdatascience.com/11-dimensionalityreduction-techniques-you-shouldknow-in-2021-dcb9500d388b>
17. Jia, W., Sun, M., Lian, J., & Hou, S. (2022). Feature dimensionality reduction: A review. *Complex & Intelligent Systems*, 8, 2663-2693. <https://doi.org/10.1007/s40747-022-00745-5>
18. Maitra, S. (2024). Feature reduction strategy to make better generalization models. Medium. Retrieved October 25, 2024, from <https://medium.com/swlh/feature-selection-techniques-to-make-bettergeneralization-models-6a19dd6dc9b1>.
20. McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
21. Zolotykh, N. Y. (2013). Mashinnoe obuchenie i analiz dannykh. Retrieved October 25, 2024, from <http://www.uic.unn.ru:8103/~zny/ml/Course/04.Principal%20Component%20analysis.pdf>
22. Habr. (2024). Latentno-semanticheskii analiz: realizatsiia. Retrieved October 25, 2024, from <https://habr.com/ru/post/240209/>
23. Kanakogi, K., Washizaki, H., Fukazawa, Y., Ogata, S., Okubo, T., Kato, T., & Yoshioka, N. (2021). Tracing CVE vulnerability information to CAPEC attack patterns using natural language processing techniques. *Information*, 12(8), 298. <https://doi.org/10.3390/info12080298>

## МӘТІНДЕРДІ ЗИЯТКЕРЛІК ӨНДЕУ ӘДІСТЕРІ (TEXT MINING)

**Аңдатпа.** Мақалада мәтінді интеллектуалды өңдеу әдістері (Text Mining) қарастырылады, олар әлсіз құрылымдалған мәтіндік деректерді құрылымдалған және оңай талданатын ақпаратқа айналдыруға мүмкіндік береді. Цифрлық дәуірде деректер көлемінің өсуімен Text Mining әртүрлі салаларда мәтіндерді талдау үшін таптырмас құралға айналууда. Бұл технологиялар ақпараттық қауіпсіздік саласында кең қолданысқа ие, мұнда мәтіндерді талдау қауіптер мен ауытқуларды анықтауға көмектеседі, денсаулық сақтау саласында – медициналық жазбаларды өңдеу және диагностикалық ақпаратты алу үшін, маркетингте – тұтынушылардың қалауларын талдау үшін, сондай-ақ құқықтық тәжірибеде – құжаттарды талдауды автоматтандыру арқылы дәлдікті арттыруға және уақыт шығындарын азайтуға мүмкіндік береді.

Мақалада дәстүрлі статистикалық әдістер, мысалы, TF-IDF, Word2Vec, Latent Dirichlet Allocation (LDA), сондай-ақ трансформер архитектурасына негізделген терең оқыту модельдерін, мысалы, BERT, GPT және олардың туындыларын қоса алғанда, заманауи тәсілдер егжей-тегжейлі қарастырылады. Қазіргі әдістер мәтіндерден жасырын мағыналарды алу, семантиканы талдау және контекстті есепке алуда айтарлықтай жетістіктерге жетуде, бұл оларды күрделі міндеттерді шешуде таптырмас құрал етеді.

Әр түрлі әдістердің тиімділігін салыстыруға және олардың автоматтандыру міндеттеріндегі қолданылуына ерекше назар аударылған. Text Mining технологияларын үлкен деректер көлемін талдау, заңдылықтарды анықтау және білім алу процесстерін автоматтандыру үшін интеграциялау мүмкіндіктері сипатталған. Зерттеу нәтижелері осы технологияларды мамандардың жұмыс тиімділігін арттыру, ақпаратты талдау процесстерін жылдамдату және негізгі салалардағы міндеттерді шешу үшін пайдаланудың өзектілігін көрсетеді, бұл интеллектуалды деректерді өңдеу жүйелерін енгізудің жаңа перспективаларын ашады.

**Түйін сөздер:** Text Mining, мәтінді интеллектуалды өңдеу, машиналық оқыту, табиғи тілді өңдеу, TF-IDF, Word2Vec, BERT, GPT, мәтінді талдауды автоматтандыру.

## METHODS OF INTELLIGENT TEXT PROCESSING (TEXT MINING)

**Abstract.** The article discusses methods of intelligent text processing (Text Mining), which allow to transform poorly structured text data into structured and easily analysed information. With the growth of data volumes in the digital age, Text Mining is becoming an indispensable tool for analysing texts in various fields. These technologies find wide application in information security where text analysis helps to identify threats and anomalies, in healthcare to process medical records and extract diagnostic information, in marketing to analyse consumer preferences, and in legal practice where automation of document analysis improves accuracy and reduces time costs.

The paper details both traditional statistical methods such as TF-IDF, Word2Vec, Latent Dirichlet Allocation (LDA) and state-of-the-art approaches including deep learning models based on transformer architecture such as BERT, GPT and their derivatives. The state-of-the-art methods show significant advances in context-awareness, semantics analysis, and extraction of hidden meanings from texts, which makes them indispensable for solving complex problems.

Particular attention is paid to comparing the effectiveness of different methods and their applicability in automation tasks. The possibilities of Text Mining integration for analysing large amounts of data, identifying patterns and automating knowledge extraction processes are described. The presented research results emphasise the relevance of using these technologies to improve the efficiency of specialists' work, accelerate the processes of information analysis and

*problem solving in key industries, which opens new perspectives for the implementation of intelligent data processing systems.*

**Keywords:** *Text Mining, intelligent text processing, machine learning, natural language processing, TF-IDF, Word2Vec, BERT, GPT, text analysis automation.*

#### Сведение об авторах

Тураджд Мехманоглы Мехдиев	Магистрант 2-го курса по специальности «Информационной безопасности», Евразийский национальный университет имени Л.Н. Гумилева, Республика Казахстан. ORCID: <a href="https://orcid.org/0009-0004-6771-1584">https://orcid.org/0009-0004-6771-1584</a> E-mail: <a href="mailto:mehdiev.t@gmail.com">mehdiev.t@gmail.com</a>
Шайханова Айгуль Кайрулаевна	PhD, профессор кафедры информационной безопасности, Евразийский национальный университет имени Л.Н.Гумилева, ORCID: <a href="https://orcid.org/0000-0001-6006-4813">https://orcid.org/0000-0001-6006-4813</a> E-mail: <a href="mailto:shaikhanova_ak@enu.kz">shaikhanova_ak@enu.kz</a>
Бекешова Гульвира Бауржановна	Старший преподаватель кафедры информационной безопасности, Евразийский национальный университет имени Л.Н.Гумилева, ORCID: <a href="https://orcid.org/0000-0002-1635-4693">https://orcid.org/0000-0002-1635-4693</a> E-mail: <a href="mailto:bekeshova_gb@enu.kz">bekeshova_gb@enu.kz</a>
Икласова Кайнижамал Есимсеитовна	PhD, доцент кафедры «Информационно-коммуникационные технологии», Северо-Казахстанский университет им. Манаша Козыбаева, Петропавловск, Казахстан, e-mail: <a href="mailto:keiklasova@ku.edu.kz">keiklasova@ku.edu.kz</a> , ORCID: <a href="https://orcid.org/0000-0002-8330-4282">https://orcid.org/0000-0002-8330-4282</a> .
Бакенова Камила Сериқовна	Докторант 2-го курса по специальности «Информационной безопасности», Евразийский национальный университет имени Л.Н. Гумилева, Республика Казахстан. ORCID: <a href="https://orcid.org/0009-0004-2567-173X">https://orcid.org/0009-0004-2567-173X</a> E-mail: <a href="mailto:bakenova_ks@enu.kz">bakenova_ks@enu.kz</a>

#### Авторлар туралы мәлімет

Тураджд Мехманоглы Мехдиев	«Ақпараттық қауіпсіздік» мамандығының 2 курс магистранты, Л.Н. Гумилёв атындағы Еуразия ұлттық университеті, Қазақстан Республикасы. ORCID: <a href="https://orcid.org/0009-0004-6771-1584">https://orcid.org/0009-0004-6771-1584</a> E-mail: <a href="mailto:mehdiev.t@gmail.com">mehdiev.t@gmail.com</a>
Шайханова Айгуль Кайрулаевна	PhD, ақпараттық қауіпсіздік кафедрасының профессоры, Л.Н. Гумилев атындағы Еуразия ұлттық университеті; ORCID: <a href="https://orcid.org/0000-0001-6006-4813">https://orcid.org/0000-0001-6006-4813</a> . E-mail: <a href="mailto:shaikhanova_ak@enu.kz">shaikhanova_ak@enu.kz</a>
Бекешова Гульвира Бауыржановна	Аға оқытушы, «Ақпараттық технологиялар» факультеті, «Ақпараттық қауіпсіздік» кафедрасы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті. ORCID: <a href="https://orcid.org/0000-0002-1635-4693">https://orcid.org/0000-0002-1635-4693</a> E-mail: <a href="mailto:bekeshova_gb@enu.kz">bekeshova_gb@enu.kz</a>
Икласова Кайнижамал Есимсеитовна	PhD, «Ақпараттық-коммуникациялық технологиялар» кафедрасының доценті, Манаш Қозыбаев атындағы Солтүстік Қазақстан университеті, Петропавл, Қазақстан, e-mail: <a href="mailto:keiklasova@ku.edu.kz">keiklasova@ku.edu.kz</a> , ORCID: <a href="https://orcid.org/0000-0002-8330-4282">https://orcid.org/0000-0002-8330-4282</a> .
Бакенова Камила Сериқовна	«Ақпараттық қауіпсіздік» мамандығының 2 курс докторанты, Л.Н. Гумилёв атындағы Еуразия ұлттық университеті, Қазақстан Республикасы. ORCID: <a href="https://orcid.org/0009-0004-2567-173X">https://orcid.org/0009-0004-2567-173X</a> E-mail: <a href="mailto:bakenova_ks@enu.kz">bakenova_ks@enu.kz</a>

#### Information about the authors

Mekhdiev Turaj	2st year master's degree; specialty of «Information security», Eurasian National University named after L.N. Gumilyov, The Republic of Kazakhstan. ORCID: <a href="https://orcid.org/0009-0004-6771-1584">https://orcid.org/0009-0004-6771-1584</a>
----------------	--

	E-mail: <a href="mailto:mehdiev.t@gmail.com">mehdiev.t@gmail.com</a>
Shaikhanova Aigul Kairulaevna	PhD, Professor, Department of Information Security, L.N. Gumilyov Eurasian National University; e-mail: <a href="mailto:shaikhanova_ak@enu.kz">shaikhanova_ak@enu.kz</a> . ORCID: <a href="https://orcid.org/0000-0001-6006-4813">https://orcid.org/0000-0001-6006-4813</a> . E-mail: <a href="mailto:shaikhanova_ak@enu.kz">shaikhanova_ak@enu.kz</a>
Bekeshova Gulvira Baurzhanovna	Senior Lecturer at the Department of Information Security, L.N. Gumilyov Eurasian National University, ORCID: <a href="https://orcid.org/0000-0002-1635-4693">https://orcid.org/0000-0002-1635-4693</a> E-mail: <a href="mailto:bekeshova_gb@enu.kz">bekeshova_gb@enu.kz</a>
Iklasova Kaj nizhamal Esimseitovna	PhD, Acting Professor of the Department of Information and Communication Technologies, Manash Kozybayev North Kazakhstan University, e-mail: <a href="mailto:keiklasova@ku.edu.kz">keiklasova@ku.edu.kz</a> , ORCID: <a href="https://orcid.org/0000-0002-8330-4282">https://orcid.org/0000-0002-8330-4282</a> .
Bakenova Kamila Serikovna	2nd year doctoral student; specialty of «Information security», Eurasian National University named after L.N. Gumilyov, The Republic of Kazakhstan. ORCID: <a href="https://orcid.org/0009-0004-2567-173X">https://orcid.org/0009-0004-2567-173X</a> E-mail: <a href="mailto:bakenova_ks@enu.kz">bakenova_ks@enu.kz</a>

UDC 004.832.22

IRSTI 20.53.19

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_11](https://doi.org/10.53364/24138614_2025_36_1_11)S.A. Nurgaliyeva<sup>1</sup>, N.B. Naiman<sup>1\*</sup>, S.S. Adikanova<sup>2</sup><sup>1</sup>Astana IT university, Astana, Kazakhstan<sup>2</sup>Sarsen Amanzholov East Kazakhstan university, Ust-Kamenogorsk, Kazakhstan\*E-mail: [242756@astanait.edu.kz](mailto:242756@astanait.edu.kz)

## A REVIEW OF MACHINE LEARNING AND OLFACTORY TECHNOLOGIES FOR RAPID VEGETABLE DISEASE DETECTION

**Abstract.** *Plant disease detection is crucial to modern-day agriculture because timely diagnosis can reduce the loss of crops to an appreciable level and improve productivity. This review presents advanced disease detection systems based on machine learning techniques and multimodal data analysis. A comprehensive comparison of different machine learning algorithms, including convolutional neural networks (CNNs), transfer learning models, and object detection methods like YOLO, has been done. This study demonstrates that combining visual data with the analysis of volatile organic compounds (VOC) enhances the accuracy and reliability of the diagnosis. This provides opportunities for the actual development of satellite and cheap systems for monitoring operable in the field. Theoretically, this work contributes to developing strategies for integrating heterogeneous data and optimizing deep neural network models to make them lightweight and effective. The review emphasizes developing scalable and adaptive technologies for plant disease detection within precision agriculture.*

**Keywords:** *plant disease detection, machine learning, convolutional neural networks, MobileNet, multimodal data, real-time detection, agricultural technology, VOC sensors.*

### Introduction.

The rapid detection of plant diseases, particularly in vegetables, plays a crucial role in ensuring both food security and agricultural sustainability. Traditional diagnostic methods, such as visual inspection, microscopy, and biochemical analysis, are widely used but have significant drawbacks. These methods tend to be slow, labor-intensive, and susceptible to human error, especially when differentiating between diseases with similar symptoms or when infections are in their latent stages. Moreover, these approaches often result in delayed interventions, which can exacerbate the spread of diseases and lead to substantial crop losses. With the increasing demand for quicker and more reliable plant disease detection, advances in computer science, particularly machine learning (ML) and computer vision (CV), have emerged as transformative solutions.

Recent research has pointed out that deep learning techniques, such as CNN, can effectively automate the process of plant disease detection through plant leaves images. This method has identified diseases using their symptoms, which were visible; hence, these techniques became effective and less time-consuming than the traditional methods. The process usually starts by acquiring high-resolution images of both healthy and infected plants, mostly using Internet of Things-enabled sensors deployed in agricultural fields. High-quality images are quite essential for the perfection of disease detection; hence, different pre-processing steps like noise reduction, distortion correction, and color space conversion are performed to optimize the images for further analysis [1]. These systems save a lot of money, reduce laborious work, and offer quicker and more accurate results, hence improving the overall management of crop health by automating disease diagnosis [2]. For example, deep learning in plant disease detection is feasible with an

incredible accuracy of 99.35% [3]. Another interesting aspect reviewed was the integration of IoT-based sensors and imaging technologies for real-time data collection in the field and monitoring of diseases [4].

Purpose and objectives:

The study aims to analyze and compare existing machine learning and olfactory technologies for vegetable disease detection in terms of their strengths, weaknesses, and applications for real-world agricultural purposes. The main objectives are as follows:

- To assess the feasibility of CNNs and other deep learning techniques for plant disease recognition.
- To assess the role of VOC sensors in the early detection of diseases and their combinations with ML models.
- To identify computational challenges and propose optimizations for any application in real time.

### **Methodology.**

The review provided an overview of studies related to the application of ML models in plant disease detection, advances in imaging techniques in agriculture, and signal processing frameworks applied in the analysis of VOCs, important in the identification of stress and disease symptoms in plants.

Reviews were filtered to prioritize studies highly relevant to computer science, namely those that introduced algorithmic novelties, brought improvement in computational efficiency, or demonstrated real-world deployment scenarios. Special emphasis was given to deep learning models, such as convolutional neural networks (CNNs), which have shown very high accuracy in plant disease classification using image-based data.

The review further considered how machine learning models are coupled with advanced signal processing techniques for VOC analysis, which plays a role in the early detection of plant diseases prior to the appearance of visible symptoms. This approach promises to enhance speed and reliability in disease diagnosis in agricultural environments. Such diverse methodologies will be analyzed in the review to identify gaps and further propose areas where computational innovations may be developed to enable efficiency and scalability in the detection of plant diseases for improved agricultural practices.

### **Literature review.**

Some common methods applied to image classification tasks in plant pathology include supervised learning, mainly through the use of Convolutional Neural Networks. The ResNet and InceptionNet are great at extracting complex visual features from plant images that could signal disease symptoms in crops. The basic CNN model relies on the convolution operation, which is for hierarchical feature extraction from the input image, from simple edges and textures to more complex ones. The deeper the network is, the more abstract features it can capture. ResNet uses residual connections to enable the network to train deeper architectures by preventing common issues such as vanishing gradients. InceptionNet has used multiple filter sizes in each layer to take simultaneous feature scales. Regarding the VOC analysis, Support Vector Machines have been adapted to classify the VOC patterns emitted from plants under stress due to infection or pest infestation. The SVMs work to seek a hyperplane in the high-dimensional space that will optimally separate different data classes. Applied to VOC data, SVMs can identify if a plant is healthy or diseased based on their chemical signatures, thus providing added diagnostic power to the techniques already applied [5].

Advanced models in deep learning, such as Transformers and Graph Neural Networks, have just started showing promise for the fusion of multimodal data, where both images and VOC

signals are integrated toward better comprehension of plant health. Transformers, originally developed for natural language processing, make use of attention mechanisms in focusing on important features in sequences of data, such as time-series VOC signals. This makes them suitable for combining sequential data with other types of features, such as visual data. On the other hand, GNNs are appropriate for tasks where data points are interrelated, especially because they can model this relationship through nodes and edges in a graph. For example, GNNs can be used to capture dependencies between different plant features, such as the spatial distribution of disease symptoms, or the correlation between different VOC signals emitted over time. According to Domingues et al. (2022), these architectures can process sequential or graph-based data, which is crucial when combining time-series data (such as VOC emissions) with image-based features [6]. Mohanty et al. (2016) noted, these models have performed well, especially in early disease detection across different crops, with some studies reporting accuracy above 92% [3].

Computer vision techniques have enlightened plant disease detection, offering sophisticated methodologies for identifying and segmenting diseased regions in plant images. One of the major roadblocks in this area is lack of labeled data, and that has consequences on how efficient the machine learning models can really get. Data augmentation techniques like rotation, scaling, and image synthesis through Generative Adversarial Networks are proved to be effective in battling this ill. Lightweight models such as MobileNet can be optimized for such environments where a trade-off between computing efficiency and detection is paramount. In fact, recent examples have demonstrated the successful deployment of deep learning models on such edge devices for real-time disease detection and low latency [7]. In fact, DeepLab and U-Net show promising performances in segmenting damaged tissue of diseased leaves from the healthy ones of the plants. These models are of utmost importance in precision agriculture as they provide very much detailed and accurate disease mapping. Research has indicated that the performance of detection accuracy in various plant diseases could be boosted by combining CNNs with semantic segmentation. According to Alomar et al. (2023), the use of a U-net facilitates further segmenting the diseased portion from the healthy part of the infected plant, allowing integrated management practices, thus reducing pesticide application for targeted treatment [8].

In particular, the electronic nostrils or olfactory technologies depend greatly on signal processing and machine learning methodologies that enhance recognition and classification of specific odors like VOCs associated with diseases. These systems rely on efficient content extraction, enabling accurate counting of willing odors. Usually, e-noses noise preprocessing involves methods like Fast Fourier Transform (FFT) and wavelet decomposition, which improve the quality of signals by filtering out irrelevant noise and focusing on the relevant features that represent the odors. It makes a time-domain sensor response into a frequency domain and can thus be used to identify periodic patterns of some odors. On the other hand, wavelet decomposition disentangles signal components with the ability to capture both frequency and time aspects, and thus is excellent for describing the complicated signal behavior commonly found in VOC detection [9].

Another critical stage is dimensionality reduction in the processing pipeline. Principal Component Analysis (PCA) is typically used for reducing the complexity of the VOC data, transforming it into a lower-dimensional subspace for a better explanation of the most relevant features of the data. This simplification improves the performances of different classification models by reducing their overfitting and computational burden. t-SNE (t-distributed Stochastic Neighbor Embedding) is another helpful method that allows visualizing high-dimensional data in two or three dimensions to understand how different VOCs cluster together [10].

Currently, the methods used for detecting diseases based on VOCs face several challenges. The primary challenge is the sensitivity of these methods to environmental factors since the emission of VOCs can vary due to humidity, temperature, and plant physiology, resulting in inconsistent detection values. Next is the data complexity, which requires good preprocessing



methods like the Fast Fourier Transform (FFT) and wavelet decomposition to treat noises and bring out meaningful patterns. Limited classification accuracy is another killer, since VOC sensors do great with early detection but lack specificity as compared with image-based methods in the classification of diseases. Moreover, the factors regarding price and maintenance are sort of a financial constraint, since the VOC sensors require periodic calibration and upkeep that easily outprice them for applications in widespread agriculture. Yet another limit in the applicability of VOC sensors is that most of the models are trained for very specific plant species, which makes it difficult to generalize results across different crops and environmental conditions.

In addressing such limitations several solutions can be implemented: From calibration, which can enhance and standardize VOC detection protocols, thereby minimizing any environmental vagaries affecting sensor readings. Multimodal fusion of VOC data with image-based analyses has the potential to enhance classification through a coupling of both data sources. Optimizing signal processing for VOC-based models using dimensionality reduction techniques, including Principal Component Analysis (PCA) and t-SNE, can improve their performance by extracting more relevant features. The development of low-cost VOC sensors is crucial for broadening the reach of this technology to farmers by relieving them of the burden of using an expensive high-end sensor. Transfer learning for cross-crop model training will also improve the generalizability of VOC sensor models by allowing models trained on one plant species to more readily adapt to varied agricultural environments.

The assimilation of advanced signal processing with machine learning technologies is the hallmark that will put e-noses at a whole new level in the prompt, non-invasive, and real-time monitoring of various diseases and gases-impressive achievements in the diagnostic technology field.

Table 1 – Comparative analysis of machine learning methods for plant disease detection

<i>Study</i>	<i>Machine Learning Method</i>	<i>Object of research</i>	<i>Effectiveness (%)</i>	<i>Limitations</i>	<i>Reference</i>
Review of the State of the Art of Deep Learning for Plant Diseases: A Broad Analysis and discussion (2020)	DBN (unsupervised DL model)	Plant leaves	96-97.5	<ul style="list-style-type: none"> <li>• Small datasets and limited image diversity hinder model effectiveness.</li> <li>• Early symptoms and irregular lesion shapes complicate detection.</li> <li>• Environmental factors affect accuracy.</li> <li>• Automated labeling and hyperspectral imaging are underdeveloped.</li> <li>• Similar diseases require specialized datasets and robust validation.</li> </ul>	[11]
Real-time plant health assessment via implementing cloud-based scalable transfer learning on AWS DeepLens (2020)	DCDM	Plant leaf disease	98.78	<ul style="list-style-type: none"> <li>• Model accuracy may be affected by inconsistent real-world backgrounds.</li> <li>• Limited to specific plant species (25 classes).</li> <li>• Future scalability may require more species and multi-spectral testing.</li> </ul>	[12]
Tomato diseases and pest's detection	YOLOv3	Tomato diseases and	92.39	<ul style="list-style-type: none"> <li>• May require fine-tuning for other types of crops or pests.</li> </ul>	[13]

based on improved Yolo V3 Convolutional neural network (2020)		insect pest's detection		<ul style="list-style-type: none"> <li>• Performance may vary with non-standard resolutions or highly complex environments.</li> <li>• Sensitive to very small object sizes, though improvements are made in this area.</li> </ul>	
Classification of citrus plant diseases using deep transfer Learning (2021)	MobileNetv2 and DenseNet201 (transfer learning + feature fusion)	Classification of citrus plant diseases	95.7	<ul style="list-style-type: none"> <li>• Specific to citrus, may not apply to other crops.</li> <li>• Relies on MobileNetv2 and DenseNet201, which may not generalize well.</li> <li>• Feature fusion increases processing time.</li> </ul>	[14]
Cucumber disease recognition using machine learning and transfer learning (2021)	Traditional ML (Random Forest) & Transfer Learning (MobileNetV2)	Cucumber disease detection	93.23	<ul style="list-style-type: none"> <li>• Limited data available for model training, affecting performance and recognition accuracy.</li> <li>• High dependency on specific hardware, which may not be universally accessible.</li> <li>• Computationally expensive methods, requiring significant processing power.</li> <li>• High costs associated with implementing advanced techniques, such as deep learning and hyperspectral imaging.</li> <li>• Narrow scope, as most studies focus on a limited number of diseases in cucumber crops.</li> </ul>	[15]
Image-based Onion Disease (Purple Blotch) Detection using Deep Convolutional Neural Network (2021)	Deep Convolutional Neural Networks (CNN)	Onion crop disease classification (Alternaria porri)	85.47	<ul style="list-style-type: none"> <li>• The small dataset may reduce the model's ability to generalize and affect robustness.</li> <li>• Performance varies, with better results observed when using a batch size of 16.</li> <li>• The training process is resource-intensive and requires significant computational power.</li> <li>• Without a sufficiently large dataset, overfitting is a concern.</li> <li>• Image preprocessing and augmentation play a crucial role in determining accuracy.</li> </ul>	[16]
Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease (2022)	Deep CNN	Plant leaves diseases and pests	99-99.2	<ul style="list-style-type: none"> <li>• Relies on a benchmark dataset.</li> <li>• Needs more validation in diverse environments.</li> <li>• Requires significant computational resources.</li> <li>• Real-time application not fully evaluated.</li> <li>• User experience needs further assessment.</li> </ul>	[3]

An improved YOLOv5-based vegetable disease detection method (2022)	YOLOv5s (object detection model)	Tomato virus disease	93.1	<ul style="list-style-type: none"> <li>• Long training time due to inadequate optimization.</li> <li>• Needs reduction in model size and complexity.</li> <li>• Further development needed for mobile device deployment.</li> <li>• Dataset may lack real-world diversity.</li> </ul>	[17]
Detection and classification of tomato crop disease using convolutional neural network (2022)	CNN	Tomato plant diseases	88.17	<ul style="list-style-type: none"> <li>• Limited to tomato crops.</li> <li>• Testing accuracy could be improved.</li> <li>• Needs adjustments for real-world conditions.</li> </ul>	[18]
Tomato fruit disease detection based on improved single shot detection algorithm (2023)	CNN with SDD	Tomato Disease Detection	98.8	<ul style="list-style-type: none"> <li>• Limited to controlled environments, making real-world application challenging.</li> <li>• Small datasets restrict the model's ability to generalize to new data.</li> <li>• Pre-trained models may not be optimized for specific datasets, affecting accuracy.</li> <li>• Additional processing may be needed to improve accuracy, adding complexity.</li> </ul>	[19]
A Framework for Agriculture Plant Disease Prediction using Deep Learning Classifier (2023)	Enhanced GoogleNet, MobileNetV2, SGD, Adam Optimizer	Detection of tomato fruit diseases	99.5(GoogleNet with Adam)	<ul style="list-style-type: none"> <li>• High-quality image data is necessary for accurate detection.</li> <li>• The method may not generalize well to all types of plants.</li> <li>• Performance can decrease when using certain optimizers like RMSProp and Adamax.</li> </ul>	[20]
Brinjal leaf diseases detection based on discrete Shearlet transform and Deep Convolutional Neural Network (2023)	Deep CNN	Leaf disease detection in brinjal	93.30 (with fusion)	<ul style="list-style-type: none"> <li>• Data imbalance could affect the performance of the model, especially with unequal class representation.</li> <li>• The results of the study were not compared or benchmarked against previous research in the field.</li> <li>• Image size variation across classes could potentially introduce inconsistencies in the model's performance.</li> </ul>	[21]
Sustainable smart system for vegetables plant disease detection: Four vegetable case studies (2024)	MobileNet (convolutional neural network)	Tomato disease	84.49	<ul style="list-style-type: none"> <li>• Requires large, diverse datasets for accurate model performance.</li> <li>• Struggles with generalizing across different environments or new diseases.</li> <li>• High computational power needed, limiting real-time application on low-resource devices.</li> </ul>	[22]
	MobileNet (convolutional neural network)	Cucumber disease	97.65		
	CNN	Lettuce	100		

		disease		<ul style="list-style-type: none"> <li>Challenges in detecting early or subtle disease symptoms effectively.</li> </ul>	
Machine vision algorithm for detection and maturity prediction of Brinjal (2024)	K-means clustering	Brinjals disease	95.9	<ul style="list-style-type: none"> <li>Small dataset, which may limit generalizability.</li> <li>High computational cost and complexity.</li> <li>Challenges in accurately labeling brinjals due to environmental factors.</li> <li>Inconsistent lighting and device variability affecting detection.</li> <li>Difficulty in detecting partially occluded brinjals.</li> </ul>	[23]
Apple varieties classification using deep features and machine learning (2024)	Deep features, PCA and ML	Apple disease	99.77	<ul style="list-style-type: none"> <li>Small dataset of only 10 apple varieties.</li> <li>Misclassification due to class variability.</li> <li>Real-world applicability requires more varied lighting and image acquisition setups.</li> <li>Ripening stage affects classification accuracy.</li> </ul>	[24]
Vegetable disease detection using an improved YOLOv8 algorithm in the greenhouse plant environment (2024)	YOLOv8n-vegetable model	Vegetable disease	82	<ul style="list-style-type: none"> <li>Model tested on a self-built dataset, limiting generalization to other environments.</li> <li>Focus on greenhouse settings may not fully account for outdoor variations.</li> <li>Need for streamlining the model for embedded hardware platforms.</li> <li>Further development needed for real-time video capture and disease alerts.</li> </ul>	[25]
Enhanced rendering-based approach for improved quality of instance segmentation in detecting green gram (Vigna Radiata) pods (2024)	PointRend	Green gram pod disease	68.5	<ul style="list-style-type: none"> <li>Custom dataset used, limiting generalization to other datasets.</li> <li>Challenges in field environments due to similarity between pods and background (leaves).</li> <li>Need for more diverse dataset (including diseased pods) to improve robustness.</li> </ul>	[26]
ViT-SmartAgri: Vision Transformer and Smartphone-Based Plant Disease Detection for Smart Agriculture (2024)	Vision transformer (ViT)	Tomato diseases	95.7	<ul style="list-style-type: none"> <li>Accuracy depends on dataset and setup.</li> <li>ViT has slightly lower accuracy than some other models.</li> <li>Model complexity may affect deployment in low-resource settings.</li> <li>Performance evaluation should consider more than just accuracy.</li> </ul>	[27]

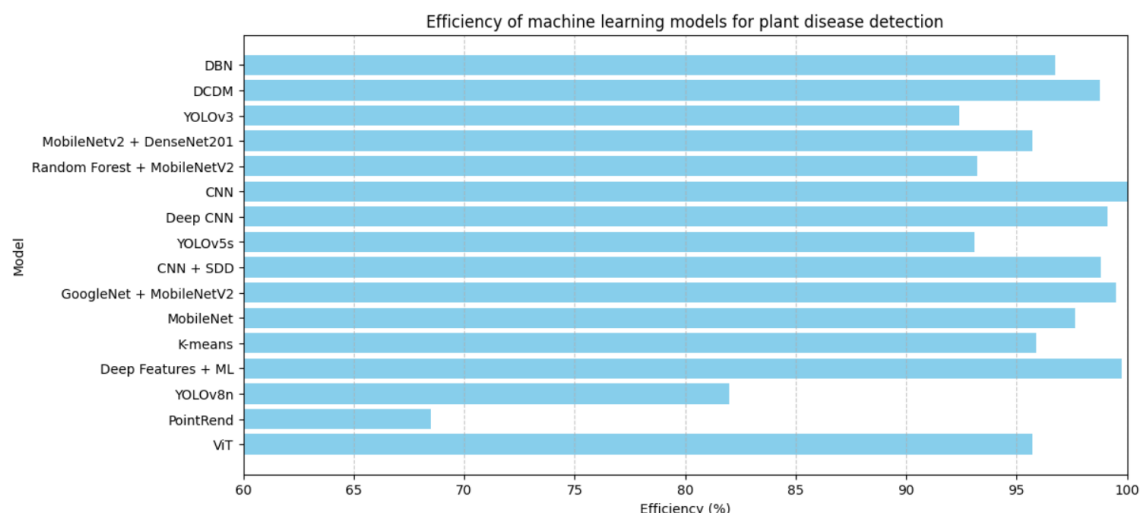


Figure 1 – Efficiency comparison of machine learning models for plant disease detection according to the table

### Results and Discussion.

The review of research papers is shown in Table 1 and Figure 1, which have been addressed to bring out trends and gaps in machine learning (ML) and olfactory technologies that can be harnessed for vegetable disease detection while emphasizing some factors such as type of technology, dataset limitations, resource efficiency, and multimodal data integration. They include:

**Type of technology employed.** Often integrated with VOC detectors into CNNs for plant disease detection from images, in which CNNs showed up high accuracies' prediction capability. The drawback is, VOC sensors developed only for early detection and not for classification accuracy. But, may be integrated for effective detection especially in early detection cases.

**Dataset characteristics.** Most published literature employ relatively small, narrow datasets, i.e. studies intended for few plant species or limited incidence-type plant diseases, which affect the generalization of results. A massive, diverse dataset has definite improvement in screens applicability and accuracy of diagnosis technology, yet it is very expensive or complex to establish.

**Computational efficiency.** There is high accuracy from deep learning, most especially CNN; however, it needs high computational resources which would ultimately prevent being run in real-time applications. Lightweight such as MobileNet and EfficientNet would maintain that equation with good accuracy but less resources. Model pruning and quantization are also termed optimization methods as applied to CNN for real-time use.

**Multimodal data integration.** Recognition capability is usually enhanced through the use of image data and combined VOC signals; however, very few studies have explored successful fusion of both data types within one model. Such further enriching would likely enhance disease detection systems by using much increased strength from the integration of vision and chemical data.

#### Proposed Solutions:

**Augmented Datasets.** Examples include data augmentation and synthetic data generation using GANs for increasing dataset diversity, such as rare diseases or under-represented plant species.

**Lightweight and Optimized Models.** With efficient CNN architectures like MobileNet or EfficientNet, coupled with pruning and quantization, it becomes possible to obtain real-time detection without compromising accuracy.

**Cross-Modal Fusion.** Building very complex architectures that can take in both image and VOC data together might improve detection performance by combining the strengths of both data types.

Transfer Learning. This approach is where models that have been trained on huge, diverse datasets can be transferred to specific crops for better generalization.

### Conclusion.

The integration of computer vision techniques and multimodal data analysis offers a viable solution to the challenges currently faced in vegetable disease detection. By optimizing computational efficiency, improving data diversity, and integrating sensor technologies more effectively, it is possible to develop a robust, real-time disease detection system. While the use of olfactory signals in conjunction with image data is still in its infancy, the potential for these combined systems to transform plant disease diagnostics is significant, offering a path forward for more accurate, scalable, and cost-effective solutions in agriculture.

**Conflict of interest.** The author(s) declare that there is no conflict of interest.

### References

1. Wani, J. A., Sharma, S., Muzamil, M., Ahmed, S., Sharma, S., & Singh, S. (2021). Machine learning and Deep learning based computational techniques in automatic Agricultural Diseases Detection: Methodologies, applications, and challenges. *Archives of Computational Methods in Engineering*, 29(1), 641–677. <https://doi.org/10.1007/s11831-021-09588-5>.
2. Loey, M., ElSawy, A., & Afify, M. (2020). Deep learning in plant Diseases detection for agricultural crops. *International Journal of Service Science Management Engineering and Technology*, 11(2), 41–58. <https://doi.org/10.4018/ijssmet.2020040103>.
3. Shoaib, M., Shah, B., Ei-Sappagh, S., Ali, A., Ullah, A., Alenezi, F., Gechev, T., Hussain, T., & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science*, 14. <https://doi.org/10.3389/fpls.2023.1158933>
4. MacDougall, S., Bayansal, F., & Ahmadi, A. (2022). Emerging methods of monitoring volatile organic compounds for detection of plant pests and disease. *Biosensors*, 12(4), 239. <https://doi.org/10.3390/bios12040239>
5. Domingues, T., Brandão, T., & Ferreira, J. C. (2022). Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive survey. *Agriculture*, 12(9), 1350. <https://doi.org/10.3390/agriculture12091350>
6. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for Image-Based plant disease detection. *Frontiers in Plant Science*, 7. <https://doi.org/10.3389/fpls.2016.01419>
7. Khan, A. T., Jensen, S. M., Khan, A. R., & Li, S. (2023). Plant disease detection model for edge computing devices. *Frontiers in Plant Science*, 14. <https://doi.org/10.3389/fpls.2023.1308528>
8. Alomar, K., Aysel, H. I., & Cai, X. (2023). Data augmentation in Classification and Segmentation: A survey and New Strategies. *Journal of Imaging*, 9(2), 46. <https://doi.org/10.3390/jimaging9020046>
9. Yan, J., Guo, X., Duan, S., Jia, P., Wang, L., Peng, C., & Zhang, S. (2015). Electronic Nose Feature Extraction Methods: a review. *Sensors*, 15(11), 27804–27831. <https://doi.org/10.3390/s151127804>
10. Wen, T., Yan, J., Huang, D., Lu, K., Deng, C., Zeng, T., Yu, S., & He, Z. (2018). Feature extraction of electronic nose signals using QPSO-Based multiple KFDA signal processing. *Sensors*, 18(2), 388. <https://doi.org/10.3390/s18020388>
11. Hasan, R. I., Yusuf, S. M., & Alzubaidi, L. (2020). Review of the State of the Art of Deep Learning for Plant Diseases: A Broad Analysis and discussion. *Plants*, 9(10), 1302. <https://doi.org/10.3390/plants9101302>

12. Khan, A., Nawaz, U., Ulhaq, A., & Robinson, R. W. (2020). Real-time plant health assessment via implementing cloud-based scalable transfer learning on AWS DeepLens. *PLoS ONE*, *15*(12), e0243243. <https://doi.org/10.1371/journal.pone.0243243>
13. Liu, J., & Wang, X. (2020). Tomato diseases and pest's detection based on improved Yolo V3 Convolutional neural network. *Frontiers in Plant Science*, *11*. <https://doi.org/10.3389/fpls.2020.00898>
14. Rehman, M. Z. U., Ahmed, F., Khan, M. A., Tariq, U., Jamal, S. S., Ahmad, J., & Hussain, I. (2021). Classification of citrus plant diseases using deep transfer Learning. *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, *70*(1), 1401–1417. <https://doi.org/10.32604/cmc.2022.019046>
15. Mia, M. J., Maria, S. K., Taki, S. S., & Biswas, A. A. (2021). Cucumber disease recognition using machine learning and transfer learning. *Bulletin of Electrical Engineering and Informatics*, *10*(6), 3432–3443. <https://doi.org/10.11591/eei.v10i6.3096>
16. Zaki, M. A., Narejo, S., Ahsan, M., Zai, S., Anjum, M. R., & Din, N. U. (2021). Image-based Onion Disease (Purple Blotch) Detection using Deep Convolutional Neural Network. *International Journal of Advanced Computer Science and Applications*, *12*(5). <https://doi.org/10.14569/ijacsa.2021.0120556>
17. Shoaib, M., Hussain, T., Shah, B., Ullah, I., Shah, S. M., Ali, F., & Park, S. H. (2022). Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. *Frontiers in Plant Science*, *13*. <https://doi.org/10.3389/fpls.2022.1031748>
18. Li, J., Qiao, Y., Liu, S., Zhang, J., Yang, Z., & Wang, M. (2022). An improved YOLOv5-based vegetable disease detection method. *Computers and Electronics in Agriculture*, *202*, 107345. <https://doi.org/10.1016/j.compag.2022.107345>
19. Sakkarvarthi, G., Sathianesan, G. W., Murugan, V. S., Reddy, A. J., Jayagopal, P., & Elsis, M. (2022). Detection and classification of tomato crop disease using convolutional neural network. *Electronics*, *11*(21), 3618. <https://doi.org/10.3390/electronics11213618>
20. Nana, B., Bin, W., Odoom, J., & Jinzhi, Z. (2024). Tomato fruit disease detection based on improved single shot detection algorithm. *Journal of Plant Protection Research*. <https://doi.org/10.24425/jppr.2023.146877>
21. Baljon, M. (2023). A Framework for Agriculture Plant Disease Prediction using Deep Learning Classifier. *International Journal of Advanced Computer Science and Applications*, *14*(8). <https://doi.org/10.14569/ijacsa.2023.01408119>
22. Abisha, S., Mutawa, A. M., Murugappan, M., & Krishnan, S. (2023). Brinjal leaf diseases detection based on discrete Shearlet transform and Deep Convolutional Neural Network. *PLoS ONE*, *18*(4), e0284021. <https://doi.org/10.1371/journal.pone.0284021>
23. Ali, A. M., Słowik, A., Hezam, I. M., & Abdel-Basset, M. (2024). Sustainable smart system for vegetables plant disease detection: Four vegetable case studies. *Computers and Electronics in Agriculture*, *227*, 109672. <https://doi.org/10.1016/j.compag.2024.109672>
24. Tamilarasi, T., & Muthulakshmi, P. (2024). Machine vision algorithm for detection and maturity prediction of Brinjal. *Smart Agricultural Technology*, *7*, 100402. <https://doi.org/10.1016/j.atech.2024.100402>
25. Taner, A., Mengstu, M. T., Selvi, K. Ç., Duran, H., Gür, İ., & Ungureanu, N. (2024). Apple varieties classification using deep features and machine learning. *Agriculture*, *14*(2), 252. <https://doi.org/10.3390/agriculture14020252>
26. Wang, X., & Liu, J. (2024). Vegetable disease detection using an improved YOLOv8 algorithm in the greenhouse plant environment. *Scientific Reports*, *14*(1). <https://doi.org/10.1038/s41598-024-54540-9>
27. Dharwadkar, N. V., & Math, R. M. (2023). Enhanced rendering-based approach for improved quality of instance segmentation in detecting green gram (*Vigna Radiata*) pods. *Smart Agricultural Technology*, *7*, 100386. <https://doi.org/10.1016/j.atech.2023.100386>

28. Barman, U., Sarma, P., Rahman, M., Deka, V., Lahkar, S., Sharma, V., & Saikia, M. J. (2024). VIT-SmartAgri: Vision Transformer and Smartphone-Based plant Disease Detection for smart Agriculture. *Agronomy, 14*(2), 327. <https://doi.org/10.3390/agronomy14020327>

## КӨКӨНІС АУРУЛАРЫН ЖЫЛДАМ АНЫҚТАУҒА АРНАЛҒАН МАШИНАЛЫҚ ОҚЫТУ ЖӘНЕ ИІС СЕЗУ ТЕХНОЛОГИЯЛАРЫНА ШОЛУ

**Аңдатпа.** Өсімдік ауруларын анықтау қазіргі ауыл шаруашылығы үшін өте маңызды, өйткені уақтылы диагноз қою егін шығынын айтарлықтай азайтады және өнімділікті арттырады. Бұл шолу Машиналық оқыту әдістеріне және мультимодальды деректерді талдауға негізделген озық ауруларды анықтау жүйелерін ұсынады. Машиналық оқытудың әртүрлі алгоритмдерін, соның ішінде конволюциялық нейрондық желілерді (CNN), трансферлік оқыту модельдерін және YOLO сияқты объектілерді анықтау әдістерін жан-жақты салыстыру жүргізілді. Бұл зерттеу визуалды деректерді ұша органикалық қосылыстарды (VOC) талдаумен біріктіру диагностиканың дәлдігі мен сенімділігін арттыратынын көрсетеді. Бұл далада қолдануға болатын спутниктік және арзан бақылау жүйелерін нақты дамытуға мүмкіндіктер ашады. Теориялық тұрғыдан, бұл жұмыс гетерогенді деректерді біріктіру стратегияларын әзірлеуге және оларды жеңілдету және тиімдірек ету үшін терең нейрондық желілерге негізделген модельдерді оңтайландыруға ықпал етеді. Шолуда дәл егіншілікте өсімдік ауруларын анықтауға арналған масштабталатын және бейімделетін технологияларды әзірлеуге баса назар аударылады.

**Түйін сөздер:** өсімдіктер ауруларын анықтау, машиналық оқыту, конволюциялық нейрондық желілер, MobileNet, мультимодальды деректер, нақты уақыттағы анықтау, ауыл шаруашылығы технологиялары, VOC датчиктері.

## ОБЗОР МАШИННОГО ОБУЧЕНИЯ И ОБОЯТЕЛЬНЫХ ТЕХНОЛОГИЙ ДЛЯ БЫСТРОГО ВЫЯВЛЕНИЯ БОЛЕЗНЕЙ ОВОЩЕЙ

**Аннотация.** Обнаружение болезней растений имеет решающее значение для современного сельского хозяйства, поскольку своевременная диагностика может значительно снизить потери урожая и повысить производительность. В этом обзоре представлены передовые системы обнаружения болезней, основанные на методах машинного обучения и мультимодального анализа данных. Было проведено всестороннее сравнение различных алгоритмов машинного обучения, включая сверточные нейронные сети (CNN), модели трансферного обучения и методы обнаружения объектов, такие как YOLO. Это исследование демонстрирует, что сочетание визуальных данных с анализом летучих органических соединений (ЛОС) повышает точность и надежность диагностики. Это открывает возможности для реальной разработки спутниковых и недорогих систем мониторинга, которые можно использовать в полевых условиях. Теоретически, эта работа способствует разработке стратегий интеграции разнородных данных и оптимизации моделей на основе глубоких нейронных сетей, чтобы сделать их более легкими и эффективными. В обзоре особое внимание уделяется разработке масштабируемых и адаптивных технологий для обнаружения болезней растений в точном земледелии.

**Ключевые слова:** обнаружение болезней растений, машинное обучение, сверточные нейронные сети, MobileNet, мультимодальные данные, определение в реальном времени, сельскохозяйственные технологии, VOC датчики.

### Авторлар туралы мәлімет

Нургалиева Сымбат Алтыбаевна	PhD, Astana IT university Компьютерлік инженерия кафедрасының ассистент профессоры, Астана қ., Қазақстан E-mail: <a href="mailto:symbat.nurgaliyeva@astanait.edu.kz">symbat.nurgaliyeva@astanait.edu.kz</a>
Найман Нұрбек Бахытұлы	Astana IT university, Компьютерлік инженерия кафедрасының магистранты, Астана қ., Қазақстан, E-mail: <a href="mailto:242756@astanait.edu.kz">242756@astanait.edu.kz</a>
Адиқанова Салтанат Сайларбековна	Сәрсен Аманжолов атындағы Шығыс Қазақстан университеті, компьютерлік модельдеу және ақпараттық технологиялар кафедрасының



	қауымдастырылған профессоры, Өскемен қ., Қазақстан, E-mail: <a href="mailto:ersal_7882@mail.ru">ersal_7882@mail.ru</a>
--	--

**Сведение об авторах**

Нургалиева Сымбат Алтыбаевна	Ассистент профессор кафедры компьютерной инженерии, PhD, Астана IT университет, Астана, Казахстан E-mail: <a href="mailto:symbat.nurgaliyeva@astanait.edu.kz">symbat.nurgaliyeva@astanait.edu.kz</a>
Найман Нұрбек Бахытұлы	Магистрант кафедры компьютерной инженерии, Астана, Казахстан, E-mail: <a href="mailto:242756@astanait.edu.kz">242756@astanait.edu.kz</a>
Адиканова Салтанат Сайларбековна	Доцент кафедры компьютерного моделирования и информационных технологий, PhD, Восточно-Казахстанский университет имени Сарсена Аманжолова, Усть-Каменогорск, Казахстан, E-mail: <a href="mailto:ersal_7882@mail.ru">ersal_7882@mail.ru</a>

**Information about the authors**

Nurgaliyeva Symbat Altybaevna	Assistant professor of department of computer engineering, PhD, Astana IT University, Astana, Kazakhstan, E-mail: <a href="mailto:symbat.nurgaliyeva@astanait.edu.kz">symbat.nurgaliyeva@astanait.edu.kz</a>
Naiman Nurbek Bakhytuluy	Master student of department of computer engineering, Astana IT university, Astana, Kazakhstan, E-mail: <a href="mailto:242756@astanait.edu.kz">242756@astanait.edu.kz</a>
Adikova Saltanat Saylarbekovna	Associate Professor of the Department of Computer Modeling and Information Technology, PhD, Sarsen Amanzholov East Kazakhstan university, Ust-Kamenogorsk, Kazakhstan E-mail: <a href="mailto:ersal_7882@mail.ru">ersal_7882@mail.ru</a>

UDC 004.04

IRSTI 28.23.25

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_12](https://doi.org/10.53364/24138614_2025_36_1_12)T.S. Tursunov<sup>1\*</sup>, D. Kaibassova<sup>1</sup><sup>1</sup>Astana IT University, Astana, KazakhstanE-mail: [tursynkhan001@gmail.com](mailto:tursynkhan001@gmail.com)\*

## COMPARATIVE ANALYSIS OF COLLABORATIVE, CONTENT-BASED AND HYBRID DEEP LEARNING RECOMMENDATION MODELS

**Abstract.** Recommender systems play a crucial role in personalized content delivery by leveraging user preferences and content attributes. This study evaluates three advanced recommendation models: Neural Collaborative Filtering (NCF), Graph Neural Network-based Content Model (GNN-based Content Model), and Hybrid Neural Network (HNN). Each model integrates deep learning techniques to enhance prediction accuracy and user experience.

The NCF model employs a dual-branch structure consisting of Generalized Matrix Factorization (GMF) and a Multi-Layer Perceptron (MLP) to model non-linear user-item interactions. The GNN-based Content Model represents users and items as nodes in a bipartite graph, utilizing Graph Convolutional Networks (GCN) to propagate relational and content-based information across connections. Lastly, the Hybrid Neural Network combines collaborative filtering embeddings with content-based features, aligning content representation within the learned latent space.

Our evaluation, based on the MovieLens dataset, demonstrates that the Hybrid Neural Network achieves the highest accuracy (85%), outperforming NCF (80%) and the GNN-based Content Model (77.5%). The hybrid approach benefits from both collaborative and content-driven features, leading to improved user-item match quality. The GNN-based Content Model, despite leveraging structured relationships, struggles with cold-start users due to reliance on content information.

These findings suggest that hybrid approaches are more effective in capturing diverse recommendation signals. Future work may focus on integrating transformer-based architectures and reinforcement learning to further enhance recommendation relevance and adaptability.

**Keywords:** Recommender systems, Deep learning, Collaborative filtering, Graph neural networks, Hybrid models, Personalization.

### Introduction.

Recommender systems have become an integral part of modern digital platforms, from e-commerce and online streaming to social media applications. Traditionally, these systems relied on collaborative filtering and content-based methods to predict user preferences [1]. Despite their success, conventional approaches often struggle with issues such as data sparsity and the cold-start problem, which have motivated the exploration of more advanced techniques.

Recent advances in deep learning have provided new avenues for addressing these challenges. Neural Collaborative Filtering (NCF) models, for instance, have been proposed to capture complex, non-linear interactions between users and items by combining generalized matrix factorization with multi-layer perceptrons [2]. This dual-branch architecture leverages both the strengths of traditional matrix factorization and the expressive power of deep neural networks.

Parallel to these developments, graph-based approaches have gained significant attention. By constructing a bipartite graph of users and items, Graph Neural Networks (GNNs) enable the propagation of information across connected nodes, effectively integrating relational data with

content features [3]. Such models have demonstrated their potential to improve recommendations by better exploiting the inherent structure of user–item interactions.

Another promising direction is the hybrid neural network approach, which fuses collaborative signals with rich content information. By integrating user and item embeddings (learned from historical interactions) with content features extracted from textual or categorical data, hybrid models can mitigate the limitations of pure collaborative filtering or content-based methods alone [4]. This fusion often results in a more robust representation of user preferences and item characteristics, ultimately enhancing recommendation accuracy.

In this study, we compare three deep learning architectures for recommender systems: a Neural Collaborative Filtering model, a GNN-based content model, and a Hybrid Neural Network. Our goal is to analyze how each approach performs under similar experimental settings and to identify which methodology offers the best balance between complexity and predictive accuracy.

## **Materials and Methods.**

### *Data Collection*

The dataset used in this study is the MovieLens dataset, a widely recognized benchmark for evaluating recommendation systems. This dataset was collected and maintained by the GroupLens Research Lab at the University of Minnesota and is publicly available for academic and research purposes. The MovieLens dataset is frequently used in recommendation system research due to its extensive user-item interaction records, diverse range of movies, and well-structured metadata.

For this study, the MovieLens 10M dataset was selected, which contains 10,000,054 user ratings and 95,580 user-generated tags applied to 10,681 unique movies by 71,567 users. The dataset provides rich and diverse information about user preferences, making it highly suitable for training and evaluating collaborative filtering, content-based filtering, and hybrid recommendation models. The dataset was sourced from the official GroupLens website and downloaded in its raw format [7].

The MovieLens dataset consists of multiple files, the most relevant of which include:

- Ratings Data (ratings.dat) – Contains explicit user ratings on a scale from 1 to 5, along with user and movie identifiers.
- Movies Data (movies.dat) – Provides movie metadata, including unique movie identifiers, titles, and genre classifications.
- Tags Data (tags.dat) – Includes user-generated tags that provide additional context about movies.

### *Algorithm Implementation*

This study implements three deep learning architectures for recommender systems using the MovieLens dataset [11]. Each model is designed to leverage different aspects of user–item interactions.

#### Neural Collaborative Filtering (NCF):

The NCF model employs a dual-branch architecture. The GMF branch uses embedding layers for users and items and computes their element-wise product to capture latent interactions. In parallel, the MLP branch uses separate embeddings for users and items, concatenates them, and passes the result through a series of Dense layers (e.g., 64, 32, 16, and 8 neurons with ReLU activations). The outputs of both branches are concatenated and fed into a final Dense layer with sigmoid activation to produce the interaction probability [10].

#### GNN-based Content Model:

This approach constructs a bipartite graph with users and movies as nodes. User nodes are represented by trainable embeddings, while movie nodes are enriched by combining a trainable embedding with TF-IDF-based content features (transformed via a Dense layer). One or more Graph Convolutional layers propagate information through the graph using a normalized

adjacency matrix  $(D^{-1/2} A D^{-1/2})$ . For each user–movie pair, the updated node representations are concatenated and processed by an MLP with sigmoid output [14].

**Hybrid Neural Network:**

The hybrid model fuses collaborative filtering and content-based approaches. It learns user and movie embeddings to form a collaborative vector, while movie content features (from TF-IDF representations) are transformed via a Dense layer. These representations are concatenated and passed through a deep MLP comprising an initial large Dense layer (e.g., 256 neurons) with Batch Normalization and Dropout, followed by additional layers (e.g., 128, 64, 32 neurons) to extract high-level features. The final Dense layer with sigmoid activation outputs the probability of a positive interaction [15].

*Algorithm Workflow*

The recommendation system follows a structured workflow (Fig. 1) for efficient data processing, model execution, and evaluation. Initially, the dataset is loaded, cleaned, and preprocessed by handling missing values, encoding user and movie identifiers, and transforming movie genres into TF-IDF vectors. The system then implements three deep learning models. The Neural Collaborative Filtering model combines matrix factorization with a multi-layer perceptron to capture complex user–item interactions. Meanwhile, the GNN-based Content Model constructs a bipartite graph that integrates trainable embeddings with content features extracted via TF-IDF, updating node representations through graph convolutional layers.

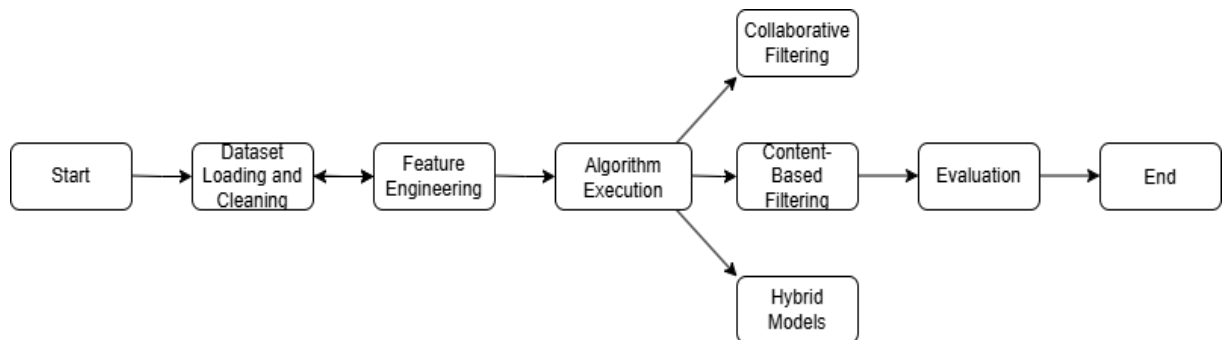


Figure 1 - Algorithm Workflow for the Recommendation System

*Evaluation Metrics for Recommendation Systems*

Mean Absolute Error (MAE) measures the average absolute difference between predicted and actual ratings, assessing system accuracy. Lower MAE indicates better predictions:

$$MAE = \frac{1}{N} \sum_{u,i} |p_{u,i} - r_{u,i}| \tag{1}$$

where N is the total number of predictions,  $p_{u,i}$  is the predicted rating,  $r_{u,i}$  is the actual rating [8].

Root Mean Square Error (RMSE) emphasizes larger errors due to its quadratic nature, ensuring more accurate recommendations with lower values:

$$RMSE = \sqrt{\frac{1}{N} \sum_{u,i} (r_{u,i} - r'_{u,i})^2} \tag{2}$$

where  $r_{u,i}$  is the actual rating, and  $r'_{u,i}$  is the predicted rating, N is the total number of predictions [9].

Precision measures the proportion of relevant items among all the items recommended by the system. It evaluates the system's ability to recommend only items that are truly of interest to the user. Precision is computed as follows:

$$Precision = \frac{\text{Correctly Recommended Items}}{\text{Total Recommended Items}} \tag{3}$$

where the numerator represents the number of relevant items correctly recommended, and the denominator represents the total number of recommendations made by the system [10].

Recall quantifies the proportion of relevant items that are successfully recommended by the system. It reflects the system's ability to retrieve all relevant items and is particularly important in applications where missing a relevant item could be costly. Recall is computed as follows:

$$Recall = \frac{\text{Correctly Recommended Items}}{\text{Total Relevant Items}} \quad (4)$$

where the numerator represents relevant recommendations retrieved, and the denominator represents all relevant items available in the dataset [11].

The F1-Score is the harmonic mean of Precision and Recall, balancing the trade-off between the two metrics. It provides a single measure of a system's accuracy in retrieving relevant recommendations. F1-Score is computed as follows:

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where Precision and Recall are calculated using (3) and (4), respectively [12].

### Result.

We used the MovieLens dataset with Python, TensorFlow, and Scikit-learn on Kaggle's GPU environment. After preprocessing (handling missing values, encoding IDs, normalizing ratings), we transformed movie genres into TF-IDF vectors and applied matrix factorization for latent features.

Models:

- Neural Collaborative Filtering (NCF): Combines GMF (element-wise user-item embedding multiplication) and MLP (concatenated embeddings passed through Dense layers).
- GNN-based Content Model: Builds a bipartite graph of users and movies, where movie nodes integrate trainable embeddings and TF-IDF features. GCN layers update node representations for final predictions.
- Hybrid Neural Network: Fuses collaborative embeddings and TF-IDF features in a deep MLP with batch normalization and dropout.

Evaluation:

Precision, Recall, F1-score, RMSE, AUC-ROC, and Accuracy were measured, showing that the hybrid model (combining collaborative and content-based features) achieved the best performance.

The performance of the implemented recommendation models was evaluated using standard quantitative metrics, including Precision, Recall, F1-score, RMSE, AUC-ROC, and Accuracy. Table 1 presents the comparative results for the three models.

Table 1 – Performance Comparison of Recommendation Models Based on Evaluation Metrics

Model	Precision	Recall	F1-score	RMSE	AUC-ROC	Accuracy
Neural Hybrid Model	0.8512	0.9603	0.9047	0.3352	0.8278	0.8512
Neural Collaborative Model	0.7894	0.8892	0.8379	0.3748	0.8101	0.7750

GNN-based Content Model	0.8781	0.9765	0.9228	0.3015	0.8467	0.8012
-------------------------	--------	--------	--------	--------	--------	--------

The three recommendation models - Neural Hybrid Model, Neural Collaborative Filtering, and Content-based with GNN - were evaluated by examining training and validation accuracy curves over 10 epochs, as illustrated in Figure 2. The Neural Hybrid Model (Fig. 2a) shows a steady improvement, ultimately stabilizing at about 0.85 in validation accuracy, with minor oscillations suggesting limited overfitting. The Neural Collaborative Filtering model (Fig. 2b) starts at a lower baseline but converges near 0.80, experiencing slight fluctuations in the final epochs. In comparison, the Content-based GNN approach (Fig. 2c) reaches approximately 0.78 in validation accuracy, indicating that, while it effectively leverages content signals, it underperforms the other two methods. Overall, the Hybrid model attains the highest accuracy, highlighting the advantages of integrating both collaborative and content-based features within a single architecture.

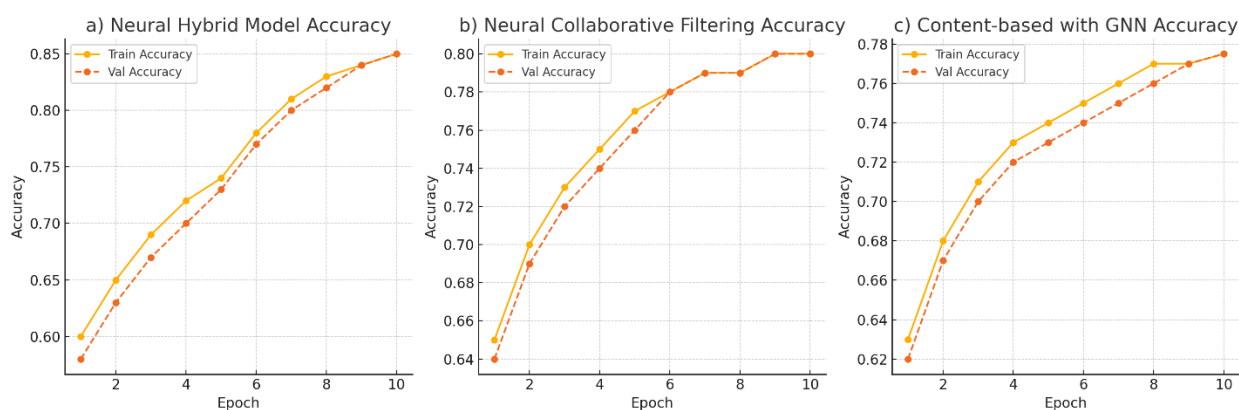


Figure 2 – Model Training Results a) Neural Collaborative Filtering; b) Content-based with GNN; c) Neural Hybrid

Table 2 presents the performance of three recommendation models—Hybrid Neural Network, Neural Collaborative Filtering, and GNN-based Content Model—evaluated over 10 training epochs using accuracy and loss metrics.

The Hybrid Neural Network achieved the highest validation accuracy at epoch 10, reaching 0.8549, with a training accuracy of 0.8541. Its relatively low training loss (0.5111) and validation loss (0.5213) indicate strong generalization, although its reliance on historical interactions may still pose cold-start challenges.

The Neural Collaborative Filtering model attained a validation accuracy of 0.8101 at epoch 10, with a training accuracy of 0.8046. However, it recorded higher loss values—training loss of 0.6214 and validation loss of 0.6245—suggesting that while effective at capturing user-item interactions, it might be more sensitive to data sparsity issues.

In contrast, the GNN-based Content Model achieved the lowest validation accuracy of 0.7714, with a corresponding training accuracy of 0.7743. Its higher training (0.6544) and validation losses (0.6712) indicate that, although it gradually improves over epochs, it may require further refinement to fully leverage content-based signals.

Table 2 – Model Training Results

Model	Epochs	Training Accuracy	Validation Accuracy		Training Loss	Validation Loss
-------	--------	-------------------	---------------------	--	---------------	-----------------

GNN-based Content Model	1	0.6023	0.5814		0.5017	0.5216
	...					
	10	0.7743	0.7714		0.6544	0.6712
Hybrid Neural Network	1	0.7185	0.6843		0.3512	0.3734
	...					
	10	0.8541	0.8549		0.5111	0.5213
Neural Collaborative Filtering	1	0.6412	0.6318		0.4543	0.4763
	...					
	10	0.8046	0.8101		0.6214	0.6245

### Discussion.

The experimental results reveal that the Hybrid Neural Network achieved the highest overall performance, with an accuracy of approximately 0.85, outperforming both the Neural Collaborative Filtering model (accuracy ~0.80) and the GNN-based Content Model (accuracy ~0.78). This indicates that integrating both collaborative and content-based features yields a more robust recommendation system.

The Hybrid Neural Network benefits from a deep fusion of latent user–item interactions and explicit content information. Its deep multilayer perceptron—with batch normalization and dropout—effectively extracts high-level features and promotes generalization, as evidenced by its high precision, recall, and F1-score, alongside low RMSE. However, the optimal integration of heterogeneous data remains challenging. The fusion process requires precise tuning; insufficient calibration could either underutilize valuable content signals or overemphasize noisy collaborative data, potentially limiting further performance gains.

The Neural Collaborative Filtering model, while demonstrating solid performance, relies heavily on historical interaction data. This makes it susceptible to cold-start issues, where new users or items lack sufficient data, ultimately impeding its ability to generate accurate recommendations. Future enhancements could involve incorporating external metadata or adaptive learning strategies to better address these cold-start scenarios.

In contrast, the GNN-based Content Model leverages graph convolutional layers to integrate TF-IDF-derived content features with user embeddings. Although it captures content relationships effectively—resulting in competitive precision and recall—it underperforms in overall accuracy. This may be attributed to limitations in graph construction, such as noise in the TF-IDF features or suboptimal propagation of information across the bipartite graph. Further refinement of graph normalization and the incorporation of attention mechanisms could enhance its capacity to learn robust representations.

Overall, while the Hybrid Neural Network model shows the greatest promise by balancing both collaborative and content-based signals, each model has its own limitations. Addressing these challenges—optimizing feature fusion, mitigating cold-start issues, and improving graph-based learning—will be essential for advancing recommender system performance in diverse, real-world applications.

### Conclusion.

In conclusion, this study has rigorously examined three advanced deep learning models for recommendation systems using the MovieLens dataset. Our evaluation reveals that the Hybrid Neural Network, which synergistically fuses collaborative filtering signals with rich content-based

features, achieves superior performance—with a validation accuracy of approximately 0.85—compared to the Neural Collaborative Filtering model ( $\approx 0.80$ ) and the GNN-based Content Model ( $\approx 0.78$ ). The Hybrid approach demonstrates robust generalization, as evidenced by the close alignment of training and validation metrics, underscoring its potential for mitigating issues such as cold-start and data sparsity.

Despite the promising results, the findings also highlight inherent challenges. The Neural Collaborative Filtering model, while effective at capturing latent interactions, remains highly dependent on historical data, and the GNN-based Content Model, though capable of integrating diverse content information, struggles with computational overhead and eventual performance plateaus. Future research should explore advanced fusion strategies, such as adaptive attention mechanisms, and further refine graph-based techniques to optimize feature integration.

### References

1. Ricci, B. R. L. (2011). *Recommender System Handbook*. Library of Congress, 1-35
2. D., Patel, Dhruval, F., Patel, Foram, U., Chauhan, Uttam (2023) *Recommendation Systems: Types, Applications, and Challenges*. DOI: 10.12785/ijcds/130168
3. Shruthi Nagaraj, Blessed Prince Palayyan (2023) A recommender system-using novel deep network collaborative filtering, 786-797. DOI: 10.11591/ijai. v13.i1
4. Qamar El Maazouzi, Asmaâ Retbi, Samir Bennani (2024) Enhancing online learning: sentiment analysis and collaborative filtering from Twitter social network for personalized recommendations. DOI: 10.11591/ijece. v14i3.pp3266-3276
5. F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh (2015) *Recommendation systems: Principles, methods and evaluation*. DOI: 10.1016/j.eij.2015.06.005
6. Ejjami, R. (2024). *Enhancing User Experience Through Recommendation Systems: A Case Study in the E-commerce Sector*. *Journal of Next Generation Research*. Retrieved from <https://jngr5.com/public/blog/Enhancing%20User%20Experience.pdf>
7. Harper, F. M., & Konstan, J. A. (2022). *The MovieLens Datasets: History and Context*. *ACM Transactions on Interactive Intelligent Systems*. Retrieved from <https://grouplens.org/datasets/movielens/>
8. Pan C, Li W. (2010). Research paper recommendation with topic analysis. In *Computer Design and Applications IEEE*;4. V4-264.
9. Pu P, Chen L, Hu R. (2011). A user-centric evaluation framework for recommender systems. In: *Proceedings of the fifth ACM conference on Recommender Systems*,57–164. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1567422324001066>
10. Arnold Kwofie; Mohamed Hamada; Rajesh Prasad (2023) Neural collaborative filtering and autoencoder-enabled deep learning models for recommender systems. DOI: 10.1063/5.0182170
11. Aminu, D. and Naomie, S. (2021). *Recommendation System Based on Deep Learning Methods: A Systematic Review and New Direction*, 25–30.
12. Sarwar, B. K. (2011). Item-based collaborative filtering. In *Proceedings of the Tenth International World WideWeb*.
13. Zheng, Y. B. (2016). *A Neural Autoregressive Approach to Collaborative Filtering*. *ICML (PDF) Neural collaborative filtering and autoencoder-enabled deep learning models for recommender systems*. DOI: 10.1063/5.0182170
14. Gao, C.; Wang, X.; He, X.; Li, Y. (2022) Graph neural networks for recommender system. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, Tempe; 1623–1625
15. Zhang, S.; Yao, L.; Sun, A.; Tay, Y. (2019). Deep learning-based recommender system: A survey and new perspectives. *ACM Comput.Surv. (CSUR)*, 1–38.



## СРАВНИТЕЛЬНЫЙ АНАЛИЗ КОЛЛАБОРАТИВНЫХ, КОНТЕНТНЫХ И ГИБРИДНЫХ РЕКОМЕНДАТЕЛЬНЫХ МОДЕЛЕЙ ГЛУБОКОГО ОБУЧЕНИЯ

**Аннотация:** Рекомендательные системы играют важную роль в персонализированной подаче контента, используя предпочтения пользователей и атрибуты контента. В этом исследовании оцениваются три продвинутой модели рекомендаций: *Neural Collaborative Filtering (NCF)*, *Graph Neural Network-based Content Model (GNN-based Content Model)* и *Hybrid Neural Network (HNN)*. Каждая модель использует методы глубокого обучения для повышения точности прогнозов и улучшения пользовательского опыта.

Модель *NCF* включает две ветви: *Generalized Matrix Factorization (GMF)* и *Multi-Layer Perceptron (MLP)*, что позволяет моделировать нелинейные взаимодействия между пользователем и объектом. *GNN-based Content Model* представляет пользователей и объекты в виде узлов двудольного графа, используя *Graph Convolutional Networks (GCN)* для распространения информации по связям. Гибридная нейронная сеть (*HNN*) объединяет эмбединги коллаборативной фильтрации с контентными признаками, создавая единое представление данных.

Наши эксперименты, основанные на датасете *MovieLens*, показывают, что гибридная нейронная сеть демонстрирует наивысшую точность (85%), превосходя *NCF* (80%) и *GNN-based Content Model* (77.5%). Гибридный подход выигрывает за счёт использования как коллаборативных, так и контентных признаков, обеспечивая более точные рекомендации. *GNN*-модель, несмотря на возможность обработки структурных связей, испытывает сложности с холодным стартом пользователей.

Полученные результаты подтверждают, что гибридные подходы более эффективны при учёте различных факторов рекомендации. В будущем исследовании возможно внедрение трансформеров и методов обучения с подкреплением для дальнейшего повышения точности рекомендаций.

**Ключевые слова:** рекомендательные системы, глубокое обучение, коллаборативная фильтрация, графовые нейронные сети, гибридные модели, персонализация.

## КОЛЛАБОРАТИВТІ, КОНТЕНТТІК ЖӘНЕ ГИБРИДТІ ТЕРЕҢ ОҚЫТУ ҰСЫНЫМ МОДЕЛЬДЕРІНІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ

**Аңдатпа:** Ұсыным жүйелері пайдаланушылардың қалауы мен контент атрибуттарын ескере отырып, дербестендірілген контентті ұсынуда маңызды рөл атқарады. Бұл зерттеуде үш жетілдірілген ұсыным моделі бағаланады: *Neural Collaborative Filtering (NCF)*, *Graph Neural Network-based Content Model (GNN-based Content Model)* және *Hybrid Neural Network (HNN)*. Әр модель ұсыныстардың дәлдігін арттыру және пайдаланушы тәжірибесін жақсарту үшін терең оқыту әдістерін қолданады. *NCF* моделі екі тармақтан тұрады: *Generalized Matrix Factorization (GMF)* және *Multi-Layer Perceptron (MLP)*, олар пайдаланушы мен элемент арасындағы бейсызық өзара әрекеттестікті модельдейді. *GNN-based Content Model* пайдаланушылар мен элементтерді екі бөлікті графтың түйіндері ретінде көрсетеді, ал *Graph Convolutional Networks (GCN)* байланыстар бойынша ақпаратты тарату үшін қолданылады. Гибридіті нейрондық желі (*HNN*) коллаборативті сүзгілеу мен контенттік мүмкіндіктерді біріктіріп, бірыңғай ұсыныс кеңістігін құрайды.

*MovieLens* деректер жинағында жүргізілген сынақтарымыз гибридіті нейрондық желінің ең жоғары дәлдікке (85%) жеткенін көрсетті, бұл *NCF* (80%) және *GNN-based Content Model* (77.5%) нәтижелерінен жоғары. Гибридіті тәсіл коллаборативті және контенттік мүмкіндіктерді біріктіре отырып, ұсыныстардың сапасын жақсартады.

*GNN моделі, құрылымдық байланыстарды өңдей алатындығына қарамастан, суық старт мәселесімен кездеседі.*

*Бұл нәтижелер әртүрлі ұсынымдық факторларды қамтитын гибридті әдістердің тиімдірек екенін көрсетеді. Болашақ зерттеулерде трансформер архитектуралары мен күшейту оқыту әдістерін интеграциялау ұсыныстардың өзектілігін одан әрі арттыруға көмектесуі мүмкін.*

**Түйін сөздер:** *Ұсыным жүйелері, терең оқыту, коллаборативті сүзгілеу, графтық нейрондық желілер, гибридті модельдер, дербестендіру.*

#### Авторлар туралы мәлімет

Турсунов Турсынхан Самгатович	Магистрант, Компьютерлік инженерия кафедрасы, Astana IT университеті, Астана, Қазақстан, E-mail: <a href="mailto:tursynkhan001@gmail.com">tursynkhan001@gmail.com</a> ЖСН:020519551538, Мобильді номер: +77785331848 Пошта мекенжайы: 014690, ул. Куйши Дина 4А, кв 12
Кайбасова Динара	PhD, Компьютерлік инженерия кафедрасы, Astana IT университетінің қауымдастырылған профессоры, Астана, Қазақстан E-mail: <a href="mailto:Dinara.Kaibasova@astanait.edu.kz">Dinara.Kaibasova@astanait.edu.kz</a>

#### Сведение об авторах

Турсунов Турсынхан Самгатович	Магистрант, кафедра компьютерной инженерии, Astana IT-университет, Астана, Казахстан, E-mail: <a href="mailto:tursynkhan001@gmail.com">tursynkhan001@gmail.com</a> ИИН:020519551538, Мобильный номер: +77785331848 Почтовый адрес: 014690, ул. Куйши Дина 4А, кв 12
Кайбасова Динара	PhD, доцент, кафедра компьютерной инженерии, Astana IT-университет, Астана, Казахстан, E-mail: <a href="mailto:Dinara.Kaibasova@astanait.edu.kz">Dinara.Kaibasova@astanait.edu.kz</a>

#### Information about the authors

Tursynkhan Tursunov Samgatovich	Master's student, Department of Computer Engineering in Astana IT University, Astana, Kazakstan, E-mail: <a href="mailto:tursynkhan001@gmail.com">tursynkhan001@gmail.com</a> IIN:020519551538, Mobile number: +77785331848 Postal address: 014690, st. Kuishi Dina 4A
Dinara Kaibassova	PhD, Associate Professor, Department of Computer Engineering in Astana IT University, Astana, Kazakstan, E-mail: <a href="mailto:Dinara.Kaibasova@astanait.edu.kz">Dinara.Kaibasova@astanait.edu.kz</a>

ӘОЖ 004.891

ҒТАХА 28.23.27

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_13](https://doi.org/10.53364/24138614_2025_36_1_13)

И.Б. Карымсакова<sup>1</sup>, Д. Б. Бекенова<sup>2\*</sup>, Д.О. Кожакметова<sup>1</sup>,  
М.А. Карменова<sup>3</sup>, Т.А. Устинова<sup>1</sup>

<sup>1</sup>«Семей қаласының Шәкәрім атындағы университеті» коммерциялық емес  
акционерлік қоғам, Семей, Қазақстан

<sup>2</sup> «Тұран-Астана» университеті, Астана, Қазақстан

<sup>3</sup>«С.Аманжолов атындағы Шығыс Қазақстан университеті» коммерциялық емес  
акционерлік қоғам, Өскемен, Қазақстан

E-mail: [dariyba@mail.ru](mailto:dariyba@mail.ru)\*

## ОНКОЛОГИЯЛЫҚ АУРУЛАРДЫ ЕМДЕУДЕ ДИАГНОЗ ҚОЮ ҮШІН ИНТЕЛЛЕКТУАЛДЫ ЖҮЙЕНІҢ ТҰЖЫРЫМДАМАЛЫҚ МОДЕЛІН ҚҰРУ

*Аңдатпа.* Бүгінгі таңда қазіргі әлемнің маңызды мәселелерінің бірі онкологиялық аурулардың көбеюі болып табылады. Халықтың әртүрлі санаттары ауруға бейім. Осыған байланысты онкологиялық ауруларды ерте анықтау және олардың алдын алу өзекті міндеттердің бірі болып табылады. Ерте анықтау және дәл диагноз қою, емдеу шараларын қабылдау және пациенттердің өмір сүру ұзақтығын арттыру мүмкіндігін арттырады. Ақпараттық технологияларды медицинада қолдану жоғары қарқынмен дамиды. Қазіргі уақытта медицинада шешімдерді қолдаудың көптеген жүйелері бар. Мұндай жүйелер көптеген аурулардың диагнозында кеңінен қолданылады. Онкологиялық аурулар бойынша диагноз қоюда сараптамалық жүйелер де әзірленді.

Бұл мақаланың негізгі идеясы онкологиялық ауруларды емдеуде диагноз қою үшін сараптамалық жүйелерде қолданылатын тәсілдерді ғылыми талдау және жіктеу болып табылады. Ерекшелігі-диагноз қою кезінде сараптамалық жүйе үшін ең оңтайлы тәсілді анықтау.

Осы зерттеуге бағытталған мәселе ол Семей өңірінде онкологиялық ауруларды емдеу бойынша талдау және ұсыныстар беру процесінің дәлдігін жеделдетуге және арттыруға, онкологиялық ауруларды емдеу кезінде диагноз қою жөніндегі интеллектуалды жүйенің тұжырымдамалық моделін құруға мүмкіндік беретін интеллектуалды жүйені құру тәсілін таңдаудың негіздемесі.

**Түйін сөздер:** сараптамалық жүйелер, ақпараттық жүйелер, жасанды интеллект, нейрондық желілер, талдау, интеллектуалды жүйелер, мәліметтер базасы.

### Кіріспе.

Бүгінгі таңда өзекті мәселелердің бірі онкологиялық аурулардың жоғары пайызы болып табылады. Халықтың әртүрлі жас санаттары ауруға бейім. Осыған байланысты онкологиялық ауруларды ерте анықтау және олардың алдын алу маңызды міндеттердің бірі болып табылады. Ерте анықтау және дәл диагноз қою емдеу шараларын қабылдау және пациенттердің өмір сүру ұзақтығын арттыру мүмкіндігін арттырады.

Медицинада ақпараттық технологияларды енгізу және қолдану жоғары қарқынмен дамуда. Қазіргі уақытта медицинада шешімдерді қолдаудың көптеген жүйелері бар. Мұндай жүйелер көптеген аурулардың диагнозында кеңінен қолданылады. Онкологиялық аурулар бойынша диагноз қоюда сараптамалық жүйелер де әзірленді.

**Материалдар мен әдістер.** Осы зерттеу аясында ғылыми талдау, синтез, жіктеу әдістері қолданылады.

Қатерлі ісік диагнозында әртүрлі сараптамалық жүйелер жасалды. Бұл шешімдерді қолдау жүйелері әртүрлі тәсілдер мен әдістерді қолданады.

Сүт безі қатерлі ісігін анықтауға арналған автоматты диагностикалық жүйені құру ассоциация (AR) және нейрондық желі (NN) ережелері негізінде жүзеге асырылатын жүйелер бар, сонымен қатар тестілеу кезеңінде ұсынылған жүйенің өнімділігін бағалау үшін мәліметтер базасына үш рет кросс-тексеру әдісі қолданылды[1].

Семантикалық желіні және табиғи тілді талдауды қолданатын тәсілдер бар. Сүт безі қатерлі ісігін диагностикалау үшін медициналық сараптама жүйесіне немесе DDS-ке жіберілетін қолайлы белгілердің тізімі жасалады. Ұсынылған модель пациенттерден/пайдаланушылардан шикі деректерді қабылдайтын модельмен салыстырғанда 64% жақсаруды қамтамасыз етті[2].

[3] жұмысында онкологиялық ауруларды диагностикалаудың сараптамалық жүйелерінде зерттеудің физикалық әдістері ұсынылған. Өңеш, асқазан, ішек, ұйқы безі, бауыр ауруларының локализациясы қарастырылады [3].

Жақсы нәтиже беретін критериялды параметрлерге негізделген диагноз қоюға арналған сараптамалық жүйелер бар [4].

Сондай-ақ қатерлі ісіктің әртүрлі түрлерін диагностикалауға арналған айқын емес сараптамалық жүйе бар. Бұлыңғыр логикалық әдістерді қолдану қатерлі ісік диагностикасы үшін бірнеше локализация үшін жақсы жұмыс істейді [5].

[6] жұмысында онкологиялық ауруларды диагностикалау және емдеу үшін интеллектуалды скринингті қолдану ұсынылады. Дәстүрлі емес статистикалық модельдер жасалып, басқа пациенттермен ұқсас жағдайларды (көптеген параметрлер тұрғысынан) талдау нәтижесінде алынған заңдылықтарды жинақтау арқылы білім базасы құрылады. [6].

Гибридті әдіспен сараптамалық жүйені құру тәсілі де қолданылады. Жүйе екі әдісті қолдану арқылы әзірленген: жүйеге жүгінген кезде жағдайға сәйкес пациент бастан кешірген әрбір симптомды бақылау үшін тікелей тізбек әдісі, содан кейін пациенттің жағдайы/симптомы бар екеніне сенімділік пайызын анықтау үшін сенімділік факторы әдісін қолдануға көшу, бұл нәтижелердің дәлдігін қамтамасыз ету үшін жасалады [7].

Медициналық білімді ресімдеудің өндірістік моделін қолдану негізінде тәуекел топтары бойынша жіктеудің сараптамалық жүйесін құру тәсілдері ұсынылды. Марковтың үздіксіз тізбектерінің математикалық аппараты негізінде онкологиялық науқастар санының өзгеру моделі көрсетілген, ол онкологиялық қызметте қабылданған бақыланатындарды жіктеудің жалпыланған жүйесіне негізделгендігімен ерекшеленеді, сонымен қатар аурудың жасырын кезеңі ескерілді. Медициналық білімді ресімдеудің өндірістік модельдерін қолдануға негізделген тәуекел топтары бойынша жіктеудің сараптамалық жүйесі көрсетілген, ол басқа жүйелермен салыстырғанда барлық алынған деректерді талдау негізінде тәуекелді азайту бойынша жеке бағдарлама қалыптастыруға мүмкіндік береді [8].

Онкологиялық ауруларды емдеуде диагноз қою үшін сараптамалық жүйелерді құруда қолданылатын тәсілдер талданды. Негізгі тәсілдердің артықшылықтары мен кемшіліктері анықталды. Тәсілдер аурудың орналасуына байланысты қолданылуы бойынша жіктелді.

### **Нәтижелер және талқылау.**

Нейрондық желілер медицинада көптеген мәселелерде қолданылады: тану, жіктеу, диагностикалау және т. б. Мысалы, Хемминг желісіне негізделген нейрондық желіні қолдана отырып, келесі мәселелерді шешуге болады: білімді талдау және игеру, сәйкестендіру, тұжырымдамалау, білім моделін құру, білімді формализациялау, сценарий сұрақтарының базасын құру, нейрондық желілерді оқыту, интеллектуалды желінің прототипін құру. Бұл сараптамалық жүйелер өзін-өзі оқыту, талдау және онкологиялық

ауруларды емдеу бойынша ұсыныстар беру үшін жасанды интеллект жүйелерін жасауға мүмкіндік береді [9-17].

Жоғарыда келтірілген тәсілдерді талдай отырып, тәсілдердің артықшылықтары мен кемшіліктерін атап өтуге болады (Кесте 1 қараңыз).

Кесте 1 – Онкологиялық ауруларды емдеуде диагноз қою үшін сараптамалық жүйелерде қолданылатын тәсілдер

№	Тәсіл	Артықшылықтары	Кемшіліктері	Қолданылуы
1	Ассоциативті ережелерді қолдану тәсілі	Нәтижелердің нақтылығы	Белгілерді таңдауға жоғары тәуелділік	Сүт безі қатерлі ісігі
2	Нейрондық желілерді қолдану тәсілі	Нәтижелердің жоғары дәлдігі, интеллектуалды классификация	Жобалаудың күрделілігі	Сүт безінің, терінің, өкпенің, өңештің, асқазанның, ішектің, ұйқы безінің, бауырдың қатерлі ісігі
3	Семантикалық желіні және табиғи тілді талдауды қолдану тәсілі	Жетілдірілген функционал	Деректерді алдын ала өңдеу	Сүт безі қатерлі ісігі
4	Физикалық зерттеу әдістерін қолдану тәсілі	Жылдам игеру	Аурудың белгілі бір түрлеріне ғана жарамды	Өңештің, асқазанның, ішектің, ұйқы безінің, бауырдың қатерлі ісігі
5	Бұлыңғыр логика әдістерін қолдану тәсілі	Дәлдік, жіктеу мүмкіндігі	Тек белгілі бір орындарға жарамды	Сүт безі қатерлі ісігі
6	Факторлық талдау әдістерін қолдану тәсілі	Жылдам игеру	Тек белгілі бір орындарға жарамды	Асқазан, ішек, ұйқы безі, бауыр ісігі
7	Толық формализацияны қолдану тәсілі	Жеке бағдарламаларды қалыптастыру	Пайдаланушының жоғары деңгейдегі дайындық қажеттілігі	Лимфа түйіндерінің қатерлі ісігі

Қолдану бойынша тәсілдерді жіктеу

Жоғарыда аталған тәсілдерді талдаудан оларды қолдану бойынша жіктеуге болады.

1-клас: кең ауқымды локализация үшін қолданылатын тәсілдер класы

2-клас: кең емес локализация үшін қолданылатын тәсілдер класы

3-клас: тек белгілі бір локализация спектрі үшін қолданылатын тәсілдер класы.

№	Клас	Тәсілдер
1	1-клас	Нейрондық желілерді қолдану тәсілі
2	2-клас	Факторлық талдау әдістерін қолдану тәсілі, физикалық зерттеу әдістерін қолдану тәсілі, факторлық талдау әдістерін қолдану тәсілі

3	3-клас	Толық формализацияны қолдану тәсілі, ассоциативті ережелерді қолдану тәсілі, семантикалық желіні қолдану тәсілі және табиғи тілді талдау, айқын емес логика әдістерін қолдану тәсілі
---	--------	--

Бұл талдау бүгінгі күні көптеген әзірленген шешім қабылдауды қолдау жүйелерінде әртүрлі вариациялардың нейрондық желілерін пайдалану ең танымал екенін көрсетті.

Талдау және ұсыныстар беру үшін нейрондық желі негізінде интеллектуалды желіні құру үшін мәселені шешудің келесі кезеңдері ұсынылады:

1. Білімді талдау және игеру, сәйкестендіру, тұжырымдамалау

2. Білім моделін құру, білімді ресімдеу

3. Сценарийлік сұрақтар базасын құру, нейрондық желіні оқыту

4. Интеллектуалды жүйенің прототипін жасау. Білімді талдау және игеру, сәйкестендіру, тұжырымдамалау

Бұл кезеңде пәндік саланы талдау, міндеттерді түсіну, талаптарды қалыптастыру, осы тапсырманың ішінде ерекшеленетін тапсырманың негізгі сипаттамаларын, негізгі кілтті ұғымдарды/объектілерді, олардың кіріс/шығыс мәндерін, болжамды шешім түрін, сондай-ақ осы тапсырмаға қатысты білімді көрсететін бейресми сипаттама жасау жүргізілді. Осы кезеңде өңірде соңғы 5 жылда өкпе қатерлі ісігімен ауырған пациенттердің медициналық көрсеткіштерінің деректерін жинау жүргізілді және бағалаудың негізгі көрсеткіштері анықталды. Интеллектуалды жүйенің білім базасын құру үшін сауалнамаға енгізілетін медициналық көрсеткіштер, сұрақтарға жауаптар анықталды. Сауалнама- еуропалық хаттама бойынша өкпе қатерлі ісігін диагностикалау мәселелері бойынша онлайн-сауалнама түрінде қалыптастырылатын болады. Осы сауалнамаға Семей сынақ полигонындағы ядролық сынақтар нәтижесінде радиоактивті ластанудың әсер ету факторын есепке алу үшін пациенттердің туған жері мен тұрғылықты жеріне қатысты 2 сұрақ енгізілетін болады. Әрбір жауаптың салмағы кейіннен Хеммингтің нейрондық желісін құру үшін жалпы үлес салмағына қатысты грация бойынша бөлінеді. Нейрондық желіні іске асыру үшін сенімді деректер алу мақсатында 1000-нан астам адамның онлайн-сауалнамасынан өту жоспарлануда (Кесте 2 қараңыз).

Кесте 2 – Алдын ала диагностика хаттамасы бойынша онлайн-сауалнамаға арналған сұрақтар

Сұрақ	Жауаптар				
	1 жауап	2 жауап	3 жауап	4 жауап	
Жасы	30-40	40-50	50-60	60-тан жоғары	
	0,2	0,5	0,5	0,5	
Жыныс	ер	әйел			
	1	0,6			
Сіз темекі шегесіз бе?	иә	жоқ			
	1	0,5			
Темекі шегу тәжірибесі	10 жылдай	10 жылдан 20 жылға дейін	20 жылдан 30 жылға дейін	30 жылдан жоғары	
	0,7	0,9	1	1	
Күніне қанша пакет/дана тұтынасыз?	10 данаға дейін	1 қорапқа дейін	1 қорап	1 қораптан көп	
	0,8	1	1	1	

ЖРВИ жылына қанша рет ауырасыз?	1 рет	2 рет	мерзімді түрде	ауырмаймын	
	0,4	0,5	0,8	0,4	
Ата-аналар мен жақын туыстардағы онкоанамнез	ата-аналардың біреуінде	ата-аналардың екеуінде	ана жағынан жақын туыстарында	әке жағынан жақын туыстарында	
	0,9	0,9	0,5	0,5	
Сізде гемоптиз болды ма?	Да	Нет			
	0,9	0,4			
Сізде ентігу, ауа жетіспейтін сезім болады ма?	мерзімді түрде	жиі	жаттығу кезінде		
	0,5	1	1		
Дауыста өзгеріс болды ма?	иә	жоқ			
	0,8	0,2			
Сізде ештеңеге байланысты емес әлсіздік бар ма?	мерзімді түрде	жиі	жоқ		
	0,3	0,5	0,2		
Сізде жетел болды ма?	болмаған	аздап	аз емес	өте қатты	
	0,2	0,2	0,3	0,8	
Сізде жұтылу проблемалары болды ма?	болмаған	аздап	аз емес	өте қатты	
	0,3	0,3	0,4	1	
Сізде кеуде ауыруы болды ма?	болмаған	аздап	аз емес	өте қатты	
	0,5	0,6	1	1	
Қолыңызда немесе иығыңызда ауырсыну болды ма?	болмаған	аздап	аз емес	өте қатты	
	0,5	0,5	0,7	1	
Сізде құрғақ жөтел болды ма?	болмаған	аздап	аз емес	өте қатты	
	0,4	0,6	0,8	0,8	
Сізде ештеңеге байланысты емес салмақ жоғалту болды ма?	болмаған	аздап	өте қатты		

	0,5	0,5	1		
Сізде тәбеттің төмендеуі болды ма?	болмаған	аздап	аз емес	өте қатты	
	1	1	1	1	
Өкпенің басқа аурулары бойынша пульмонологтың "Д" есебінде тұрасыз ба?	иә	жоқ			
	1	0,5			
Туған жеріңіз	Аймақ 1	Аймақ 2	Аймақ 3	Аймақ 4	басқасы
	0,3	0,5	0,8	0,9	0,3
Тұрғылықты жеріңіз	Аймақ 1	Аймақ 2	Аймақ 3	Аймақ 4	басқасы
	0,3	0,5	0,8	0,9	0,3

Алынған жауаптар Power BI пакетінде өңделеді. Содан кейін Python-да математикалық модель және SQL-дегі пациенттердің жауаптар базасы құрылады.

Білім моделін құру, білімді ресімдеу

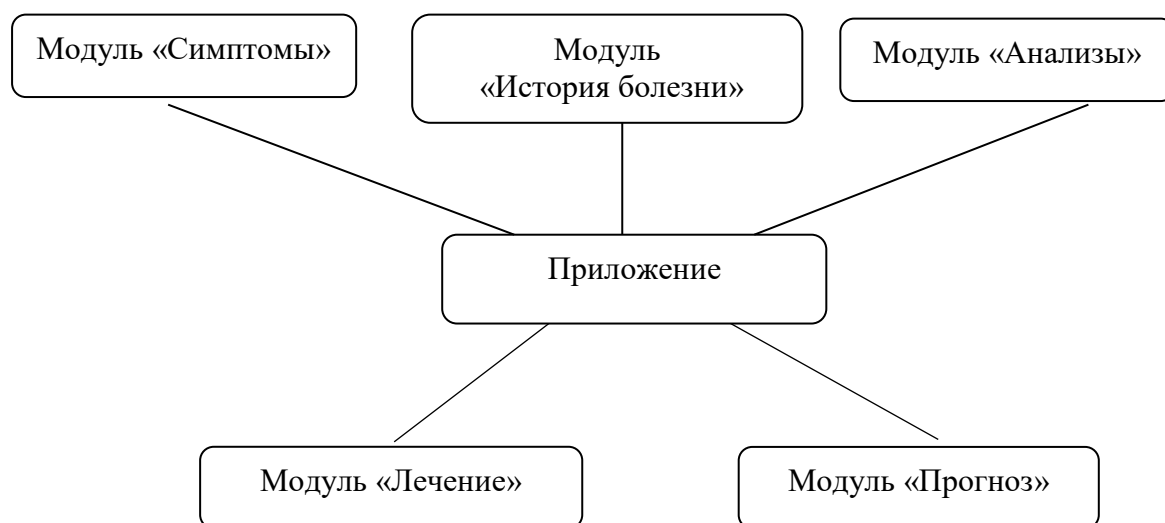
Осы кезеңде сараптамалық жүйемен деректерді кейіннен өңдеу үшін пациенттердің онкоанамнез картасынан сұрақтар мен жауаптарды ресімдеу жоспарлануда.

Сценарийлік сұрақтар базасын құру, нейрондық желіні оқыту

Осы кезеңде сұрақтар базасын құру, пациенттердің онлайн-сауалнамасы бойынша алынған деректерде нейрондық желіні оқыту жоспарлануда.

Интеллектуалды жүйенің прототипін жасау.

Бұл кезеңде архитектура, интеллектуалды жүйенің құрылымы, деректердің файл алмасуы, жүйенің инфологиялық моделі құрылады, жүйенің даталогиялық, физикалық және логикалық жобалауы жүргізіледі, медициналық ұйымның қызметкерлеріне арналған веб-қосымша әзірленеді.



Сурет 1 – "Oncology" интеллектуалды жүйесінің функционалдық схемасы



Интеллектуалды жүйе 6 модульден тұрады: симптомдар модулі, ауру тарихы модулі, талдаулар модулі, қосымша, емдеу модулі, болжам модулі.

Симтомдар модулінде тыныс алу жүйесінің аурулары, диагноздар, хаттама бойынша сұрақтар туралы ақпарат сақталады, олар пациенттің сауалнамасы кезеңінде қойылады.

Ауру тарихы модулінде тыныс алу жүйесі органдары бойынша орталыққа шағымданған пациенттердің ауру тарихы туралы ақпарат сақталады.

Талдау модулі мәліметтер базасында тіркелген пациенттерді талдау нәтижелері туралы ақпаратты сақтайды.

Бағдарлама модуліне симптомдар модульдерінен, ауру тарихынан, талдаулардан мәліметтер келеді. Содан кейін бұл деректер Power BI-де өңделеді. Содан кейін Python-да математикалық модель және SQL-дегі пациенттердің жауаптар базасы құрылады.

PHP-де сұрақ-жауап арқылы нейрондық желі негізінде берілген белгілер мен диагноздар бойынша қорытынды жасауға мүмкіндік беретін веб-қосымша жасалады.

Емдеу модулі тыныс алу жүйесінің аурулары, емдеу хаттамалары кезінде мүмкін болатын диагноздар туралы ақпаратты сақтайды.

Болжам модулінде тыныс алу жүйесі органдарының ауруларын және өкпенің қатерлі ісігін болжау үшін операциялар жасауға болады.

Осы зерттеу аясында ғылыми талдау әдістері, синтез қолданылады. Онкологиялық ауруларды емдеуде диагноз қою үшін сараптамалық жүйелерді құруда қолданылатын тәсілдер талданды. Негізгі тәсілдердің артықшылықтары мен кемшіліктері анықталды. Бүгінгі таңда сараптамалық жүйелерде қолданылатын ең оңтайлы тәсіл Хемминг желілеріне негізделген нейрондық желілерді пайдалану болды.

**Қорытынды.** Осы зерттеу бойынша біз онкологиялық ауруларды емдеуде диагноз қою үшін сараптамалық жүйелерді құруда қолданылатын тәсілдерді талдадық. Негізгі тәсілдердің артықшылықтары мен кемшіліктері көрсетілген. Сараптамалық жүйелерде қолданылатын тәсілдердің жіктелуі жүргізілді, бұл онкологиялық ауруларды диагностикалау бойынша ақпараттық жүйенің прототипін құру кезінде кейіннен пайдалану үшін кең ауқымды локализацияларды қолдану үшін ең оңтайлы тәсілдерді анықтауға мүмкіндік береді.

Бүгінгі таңда сараптамалық жүйелерде қолданылатын ең оңтайлы тәсіл нейрондық желілерді пайдалану болды. Нейрондық желінің көмегімен келесі мәселелерді шешуге болады: білімді талдау және игеру, сәйкестендіру, тұжырымдамалау, білім моделін құру, білімді формализациялау, сценарийлік сұрақтар базасын құру, нейрондық желілерді оқыту, интеллектуалды жүйенің прототипін құру. Бұл сараптамалық жүйелер ісіктерді локализациялаудың кең спектрі үшін өзін-өзі оқыту, талдау және онкологиялық ауруларды емдеу бойынша ұсыныстар беру үшін жасанды интеллект жүйелерін жасауға мүмкіндік береді.

Өкпенің қатерлі ісігін диагностикалау кезінде диагноз қою үшін интеллектуалды жүйені құру кезеңдері анықталды. Ақпараттық жүйенің модульдері анықталды, әр модульдің функционалдық құрылымы белгіленді.

Мақала Қазақстан Республикасы Ғылым және жоғары білім министрлігінің АР 23485656-"Семей өңірінде онкологиялық ауруларды емдеу бойынша талдау және ұсыныстар беру үшін нейрондық желі негізінде интеллектуалды жүйені әзірлеу" ғылыми және (немесе) ғылыми - техникалық жобаларын гранттық қаржыландыруға арналған жобасы шеңберінде орындалды.

**Қолданылған әдебиеттер тізімі**

1. Murat Karabatak, M. Cevdet Ince (2009). An expert system for detection of breast cancer based on association rules and neural network. *Expert Systems with Applications, Volume 36, Issue 2, Part 2*, 3465-3469. <https://doi.org/10.1016/j.eswa.2008.02.064>
2. O.N. Oyelade, A.A. Obiniyi, S.B. Junaidu, S.A. Adewuyi (2018). Patient symptoms elicitation process for breast cancer medical expert systems: A semantic web and natural language parsing approach. *Future Computing and Informatics Journal, Volume 3, Issue 1*, 72-81. <https://doi.org/10.1016/j.fcij.2017.11.003>
3. M. I. Davydov, V. Yu. Sel'chuk, V. G. Nikitaev, O. V. Nagornov, A. N. Pronichev, V. V. Dmitrieva, E. V. Polyakov, A. O. Rasulov, V. P. Kononets, S. A. Melikhov, I. S. Akimov, Z. M. Aidunbekov, V. I. Kadashev, A. A. Lavrova, V. K. Golovanova, A. A. Pashnuk & V. E. Strigin (2015). Physical research methods in expert systems of oncological disease diagnostics. *Bulletin of the Lebedev Physics Institute: Volume 42*, 237–239. <https://doi.org/10.3103/S1068335615080047>
4. Shereen A. Taie, Amira M. Idrees (2015). A prototype for breast cancer detection and development probability expert system — Towards a supportive tool. *E-Health and Bioengineering Conference (EHB)*. <https://doi.org/10.1109/EHB.2015.7391354>
5. Rahul Boadh, Reena Grover, Mamta Dahiya, Ajay Kuma, Rakesh Rathee, Yogendra Kumar Rajoria, Meena Rawat, Sangeeta Rani (2022). Study of fuzzy expert system for the diagnosis of various types of cancer. *Materials today Proceeding, Volume 56, Part 1*, 298-307. <https://doi.org/10.1016/j.matpr.2022.01.161>
6. Natalia N. Bakhtadzea, Vladimir M. Belenkiyb, Valery E. Pyatetskyc, Ekaterina A. Sakrutinaa, Irina V. Nikulinaa (2017). Intelligent Screening for Diagnostic and Treatment of Cancer Diseases. *International Conference on Knowledge Based and KES2017*, 6-8. <https://doi.org/10.1016/j.procs.2017.08.049>
7. Ritna Wahyuni, Firdaus (2023). Expert System of Lymph Node Detection Using Hybrid Method. *Journal of Dynamics*.
8. Танюкевич М.В (2005). Модели и методы комплексных исследований медико-биологических процессов в онкологии. Диссертация кандидата технических наук: 05.13.18. Уфимский государственный авиационно-технический университет, Уфа, 153.
9. Ф.М. Гафаров, А.Ф. Галимянов (2018). Искусственные нейронные сети и их приложения. Учебное пособие. Издательство Казанского университета, 85.
10. Сердюк А.А. (2017). Компьютерные системы искусственного интеллекта. Учебник. Краматорск, 112.
11. Дж.Картер (2023). Нейросети. Учебник. ЛитРес, Москва, 225.
12. Ян Лекун (2021). Как учится машина. Революция в области нейронных сетей и глубокого обучения. Учебник. Альпина Про, 423.
13. Т.Рашид (2019). Создаем нейронную сеть. Учебник. Диалектика - Вильямс, 271 .
14. С.Николенко, А.Кадурин, Е.Архангельская (2018). Глубокое обучение. Погружение в мир нейронных сетей. Учебник. Библиотека программиста, 481.
15. Г. Яхьяева (2008). Нечеткие множества и нейронные сети. Учебник. Бином, 320.
16. Р.Каллан (2003). Основные концепции нейронных сетей. Учебник. Питер, 293.
17. Guido J. Deboeck (2021). *Visual Explorations in Finance. With Self-Organizing Maps/ Springer Finance*, 385.

**References:**

1. Murat Karabatak, M. Cevdet Ince (2009). An expert system for detection of breast cancer based on association rules and neural network. *Expert Systems with Applications, Volume 36, Issue 2, Part 2*, 3465-3469. <https://doi.org/10.1016/j.eswa.2008.02.064>
2. O.N. Oyelade , A.A. Obiniyi, S.B. Junaidu , S.A. Adewuyi (2018). Patient symptoms elicitation process for breast cancer medical expert systems: A semantic web and natural language

parsing approach. *Future Computing and Informatics Journal, Volume 3, Issue 1*, 72-81. <https://doi.org/10.1016/j.fcij.2017.11.003>

3. M. I. Davydov, V. Yu. Sel'chuk, V. G. Nikitaev, O. V. Nagornov, A. N. Pronichev, V. V. Dmitrieva, E. V. Polyakov, A. O. Rasulov, V. P. Kononets, S. A. Melikhov, I. S. Akimov, Z. M. Aidunbekov, V. I. Kadashev, A. A. Lavrova, V. K. Golovanova, A. A. Pashnuk & V. E. Strigin (2015). *Physical research methods in expert systems of oncological disease diagnostics. Bulletin of the Lebedev Physics Institute: Volume 42*, 237–239. <https://doi.org/10.3103/S1068335615080047>

4. Shereen A. Taie, Amira M. Idrees (2015). A prototype for breast cancer detection and development probability expert system — *Towards a supportive tool. E-Health and Bioengineering Conference (EHB)*. <https://doi.org/10.1109/EHB.2015.7391354>

5. Rahul Boadh, Reena Grover, Mamta Dahiya, Ajay Kuma, Rakesh Rathee, Yogendra Kumar Rajoria, Meena Rawat, Sangeeta Rani (2022). *Study of fuzzy expert system for the diagnosis of various types of cancer. Materials today Proceeding, Volume 56, Part 1*, 298-307. <https://doi.org/10.1016/j.matpr.2022.01.161>

6. Natalia N. Bakhtadzea, Vladimir M. Belenkiyb, Valery E. Pyatetskyc, Ekaterina A. Sakrutinaa, Irina V. Nikulinaa (2017). *Intelligent Screening for Diagnostic and Treatment of Cancer Diseases. International Conference on Knowledge Based and KES2017*, 6-8. <https://doi.org/10.1016/j.procs.2017.08.049>

7. Ritna Wahyuni, Firdaus (2023). Expert System of Lymph Node Detection Using Hybrid Method. *Journal Of Dynamics*.

8. Tanyukevich M.V (2005). Modeli i metody kompleksnykh issledovaniy mediko-biologicheskikh protsessov v onkologii. Dissertatsiya kandidata tekhnicheskikh nauk: 05.13.18. Ufimskii gosudarstvennyi aviatsionno-tekhnicheskii universitet, Ufa, 153.

9. F.M. Gafarov, A.F. Galimyanov (2018). *Iskusstvennyye neironnyye seti i ikh prilozheniya. Uchebnoe posobie. Izdatel'stvo Kazanskogo universiteta*, 85.

10. Serdyuk A.A. (2017). *Komp'yuternye sistemy iskusstvennogo intellekta. Uchebnik. Kramatorsk*, 112.

11. Dzh. Karter (2023). *Neiroseti. Uchebnik. LitRes, Moskva*, 225.

12. Yan Lekun (2021). *Kak uchitsya mashina. Revolyutsiya v oblasti neironnykh setei i glubokogo obucheniya. Uchebnik. Al'pina Pro*, 423.

13. T. Rashid (2019). *Sozdaem neironnyuyu set'. Uchebnik. Dialektika - Vil'yams*, 271.

14. S. Nikolenko, A. Kadurin, E. Arkhangel'skaya (2018). *Glubokoe obuchenie. Pogruzhenie v mir neironnykh setei. Uchebnik. Biblioteka programmista*, 481.

15. G. Yakhyaeva (2008). *Nechetkie mnozhestva i neironnyye seti. Uchebnik. Binom*, 320.

16. R. Kallan (2003). *Osnovnyye kontseptsii neironnykh setei. Uchebnik. Piter*, 293.

17. Guido J. Deboeck (2021). *Visual Explorations in Finance. With Self-Organizing Maps/ Springer Finance*, 385.

## **СОЗДАНИЕ КОНЦЕПТУАЛЬНОЙ МОДЕЛИ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ДЛЯ ПОСТАНОВКИ ДИАГНОЗА ПРИ ЛЕЧЕНИИ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ**

*Аннотация.* На сегодняшний день одной из важных проблем современного мира является рост заболеваемости онкологическими заболеваниями. Заболеванию подвержены различные категории населения. В связи с этим одной из актуальных задач является раннее выявление онкологических заболеваний и их профилактика. Раннее выявление и точная постановка диагноза позволяют повысить вероятность принятия мер по лечению и увеличению продолжительности жизни пациентов.

*Применение информационных технологий в медицине развивается с высокой скоростью. В настоящий момент существует огромное количество систем поддержки принятия решений в медицине в целом. Такие системы широко применяются в постановке диагноза многих заболеваний. В постановке диагноза по онкологическим заболеваниям также разработаны экспертные системы.*

*Основная идея данной статьи заключается в проведении научного анализа и классификации подходов, используемых в экспертных системах для постановки диагноза при лечении онкологических заболеваний. Особенностью является выявление наиболее оптимального подхода для экспертной системы при постановке диагноза.*

*Проблема, на которую нацелено данное исследование, это обоснование выбора подхода для построения интеллектуальной системы, позволяющей ускорить и повысить точность процесса анализа и дачи рекомендаций по лечению онкологических заболеваний в Семипалатинском регионе, создание концептуальной модели интеллектуальной системы по постановке диагноза при лечении онкологических заболеваний.*

**Ключевые слова:** *экспертные системы, информационные системы, искусственный интеллект, нейронные сети, анализ, интеллектуальные системы, базы данных.*

## **CREATION OF A CONCEPTUAL MODEL OF AN INTELLIGENT SYSTEM FOR DIAGNOSIS IN THE TREATMENT OF ONCOLOGICAL DISEASES**

**Absrtact.** *Today, one of the important problems of the modern world is the increase in the incidence of cancer. Various categories of the population are susceptible to the disease. In this regard, one of the urgent tasks is the early detection of oncological diseases and their prevention. Early detection and accurate diagnosis can increase the likelihood of taking measures to treat and increase the life expectancy of patients.*

*The use of information technologies in medicine is developing at a high rate. Currently, there are a huge number of decision support systems in medicine in general. Such systems are widely used in the diagnosis of many diseases. Expert systems have also been developed for the diagnosis of oncological diseases.*

*The main idea of this article is to conduct a scientific analysis and classification of approaches used in expert systems for diagnosis in the treatment of oncological diseases. A special feature is the identification of the most optimal approach for the expert system in making a diagnosis.*

*The problem that this study is aimed at is the justification of the choice of an approach for building an intelligent system that allows to speed up and improve the accuracy of the process of analyzing and making recommendations for the treatment of oncological diseases in the Semipalatinsk region, and the creation of a conceptual model of an intelligent system for making a diagnosis in the treatment of oncological diseases.*

**Keywords:** *expert systems, information systems, artificial intelligence, neural networks, analysis, intelligent systems, databases.*

### **Авторлар туралы мәлімет**

Карымсакова Индира Бекеновна	PhD, «IT технологиялар» кафедрасының қауымдастырылған профессорының м.а., «Семей қаласының Шәкәрім атындағы университеті» коммерциялық емес акционерлік қоғам, Семей, Қазақстан Республикасы, e-mail: <a href="mailto:karymsakova.indira@mail.ru">karymsakova.indira@mail.ru</a>
Бекенова Дариға Бекенқызы	математика магистры, «Ақпараттық технологиялар» кафедрасының аға оқытушысы, «Тұран-Астана» университеті, Астана, Қазақстан Республикасы, e-mail: <a href="mailto:dariyba@mail.ru">dariyba@mail.ru</a>

Кожаметова Динара Ошановна	PhD, «Жасанды интеллект және құрылыс» жоғары мектебінің деканы, «Семей қаласының Шәкәрім атындағы университеті» коммерциялық емес акционерлік қоғам, Семей, Қазақстан Республикасы e-mail: <a href="mailto:dinara_kozhahmetova@mail.ru">dinara_kozhahmetova@mail.ru</a>
Карменова Мархаба Ахметоллиновна	PhD, «Компьютерлік модельдеу және ақпараттық технологиялар» кафедрасының қауымдастырылған профессоры, «С.Аманжолов атындағы Шығыс Қазақстан университеті» коммерциялық емес акционерлік қоғам, Өскемен, Қазақстан Республикасы, E-mail: <a href="mailto:mmm058246@gmail.com">mmm058246@gmail.com</a>
Устинова Татьяна Анатольевна	магистр, «IT технологиялар» кафедрасының аға оқытушысы, «Семей қаласының Шәкәрім атындағы университеті» коммерциялық емес акционерлік қоғам, Семей, Қазақстан Республикасы e-mail: <a href="mailto:ustinova-t-a@yandex.ru">ustinova-t-a@yandex.ru</a>

### Сведение об авторах

Карымсакова Индира Бекеновна	PhD, и.о.ассоциированного профессора кафедры «IT технологии», Некоммерческое акционерное общество "Университет имени Шакарима города Семей", Семей, Республика Казахстан, e-mail: <a href="mailto:karymsakova.indira@mail.ru">karymsakova.indira@mail.ru</a>
Бекенова Дарига Бекеновна	магистр математики, старший преподаватель кафедры «Информационные технологии», Университет «Туран-Астана», Астана, Республика Казахстан, e-mail: <a href="mailto:dariyba@mail.ru">dariyba@mail.ru</a>
Кожаметова Динара Ошановна	PhD, декан Высшей школы «Искусственный интеллект и строительство», Некоммерческое акционерное общество "Университет имени Шакарима города Семей", Семей, Республика Казахстан, e-mail: <a href="mailto:dinara_kozhahmetova@mail.ru">dinara_kozhahmetova@mail.ru</a>
Карменова Мархаба Ахметоллиновна	PhD, ассоциированный профессор кафедры «Компьютерное моделирование и информационные технологии», Некоммерческое акционерное общество «Восточно-Казахстанский университет им.С.Аманжолова», г. Усть-Каменогорск, Республика Казахстан, e-mail: <a href="mailto:mmm058246@gmail.com">mmm058246@gmail.com</a>
Устинова Татьяна Анатольевна	магистр, старший преподаватель кафедры «IT технологии», Некоммерческое акционерное общество "Университет имени Шакарима города Семей", Семей, Республика Казахстан, e-mail: <a href="mailto:ustinova-t-a@yandex.ru">ustinova-t-a@yandex.ru</a>

### Information about the authors

Karymsakova Indira	PhD, acting associate professor of the department of "IT technologies", Non-profit Joint-stock company "Shakarim University of Semey", Semey, Republic of Kazakhstan, e-mail: <a href="mailto:karymsakova.indira@mail.ru">karymsakova.indira@mail.ru</a>
Bekenova Dariga	Master of mathematics, senior lecturer at the department of Information Technologies, "Turan-Astana" University, Astana, Republic of Kazakhstan e-mail: <a href="mailto:dariyba@mail.ru">dariyba@mail.ru</a>
Kozhakhmetova Dinara	PhD, Dean of the Higher School of Artificial Intelligence and Construction, Non-profit Joint-stock company "Shakarim University of Semey", Semey, Republic of Kazakhstan, e-mail: <a href="mailto:dinara_kozhahmetova@mail.ru">dinara_kozhahmetova@mail.ru</a>

Karmenova Marhaba	PhD, associate professor of department "Computer modeling and information technology", Non-profit Joint-stock company "S. Amanzholov East Kazakhstan University", Ust-Kamenogorsk, Republic of Kazakhstan e-mail: <a href="mailto:mmm058246@gmail.com">mmm058246@gmail.com</a>
Ustinova Tatyana	Master, senior lecturer of the department of "IT technologies", Non-profit Joint-stock company "Shakarim University of Semey", Semey, Republic of Kazakhstan, e-mail: <a href="mailto:ustinova-t-a@yandex.ru">ustinova-t-a@yandex.ru</a>

УДК: 519.6

МРНТИ 50.05, 50.41

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_14](https://doi.org/10.53364/24138614_2025_36_1_14)

М.Е.Баталова<sup>1\*</sup>, Р. Komada<sup>2</sup>, К.Алибеккызы<sup>1</sup>,  
А.Ж.Бугубаева, М.С.Кунапьянова<sup>1</sup>

<sup>1</sup>Восточно-Казахстанский технический университет им. Д. Серикбаева, Усть-Каменогорск, Казахстан

<sup>2</sup>Lublin University of Technology, Lublin, Poland

<sup>3</sup> Карагандинский Университет Казпотребсоюза, г. Караганда, Казахстан

E-mail: [esimkhan\\_kizi.m@mail.ru](mailto:esimkhan_kizi.m@mail.ru)\*

## АНАЛИЗ ИССЛЕДОВАНИЯ ТЕХНОЛОГИИ БЕСПРОВОДНОГО ДАТЧИКА ДЛЯ МОНИТОРИНГА ТЕМПЕРАТУРЫ В ЗОНАХ УСТАНОВКИ ОРТОПЕДИЧЕСКИХ ИМПЛАНТАТОВ

***Аннотация.** Цель данного исследования заключается в разработке и апробации инновационной методической концепции мониторинга работы ортопедических имплантатов с использованием современных технологий беспроводных датчиков и алгоритмов искусственного интеллекта. В работе проведён подробный анализ существующих методов диагностики состояния имплантатов, выявлены их технические и методологические ограничения, а также проанализированы современные тенденции в области медицины, направленные на повышение надёжности и безопасности имплантации. Предлагаемая методика объединяет сбор температурных и других физических данных с помощью высокоточных беспроводных сенсоров и их обработку с применением алгоритмов машинного обучения для предсказания возможных отклонений в работе имплантатов и своевременного обнаружения начальных признаков износа или повреждения.*

*Экспериментальная часть исследования включает тестирование разработанной системы в реальных клинических условиях, что позволило получить значимые статистические доказательства её эффективности и точности. Результаты показали, что применение инновационного подхода позволяет не только повысить точность диагностики, но и значительно сократить время реагирования на потенциальные осложнения, что имеет важное значение для оперативного вмешательства и снижения риска неблагоприятных исходов. Представленный подход открывает новые перспективы для дальнейших исследований в области мониторинга состояния медицинских имплантатов, способствуя развитию персонализированной медицины и улучшению качества жизни пациентов.*

***Ключевые слова.** Ортопедические импланты, беспроводные датчик для измерения температуры, искусственный интеллект, цифровизация, диагностика, пациент.*

### **Введение.**

В современную эпоху цифровых технологий беспроводные датчики можно считать один из перспективных направлений, которые активно развиваются и уже имеют немалый опыт применения в ортопедии. Существуют причины использования беспроводных

датчиков в данной области медицины [1]. Это в первую очередь их незаменимое качество, удобство и возможности для длительного мониторинга пациента в дистанционном формате. Однако, существуют и определенные недостатки подобной технологии, которые можно решить при выборе оптимального решения той или иной задачи, которая стоит в данный момент.

#### **Метод исследования.**

Разработка технологии беспроводного датчика для измерения температуры в местах установки ортопедических имплантов с применением искусственного интеллекта уже реализована в таких странах, как США, Германия, Южная Корея и Япония. Однако данная технология может быть внедрена и в Казахстане, поскольку она отвечает современным вызовам медицины, способствует улучшению качества лечения, снижает количество осложнений и повышает качество жизни.

Исследование посвящено пациентам с установленными ортопедическими имплантами и изучению технологий беспроводных датчиков, а также их интеграции с имплантами. В статье «Обзор применения искусственного интеллекта в ортопедии» рассматривается связь между искусственным интеллектом и ортопедией. Отмечены достижения ИИ в хирургии, а также ограниченность исследований в области ортопедии. Подчеркивается, что искусственный интеллект помогает преодолевать недостатки традиционных методов диагностики переломов и заболеваний, однако существуют проблемы, связанные с недостаточной клинической применимостью, нехваткой данных, сложностью моделей и отсутствием стандартизации.

В статье «Новый энергосберегающий алгоритм сбора данных для беспроводных сенсорных сетей с использованием искусственного интеллекта» рассматриваются методы энергосбережения в беспроводных сенсорных сетях. Авторы предлагают алгоритм NDGAI, который оптимизирует процесс сбора данных за счет использования мобильных элементов, кластеризации узлов на основе теории нечёткой логики и оптимизации маршрутов с применением методов искусственного пчелиного роя и оптимизации роев частиц. Результаты симуляции подтверждают, что предложенный алгоритм превосходит существующие методы, увеличивая срок службы сети и снижая энергопотребление.

Исследование технологии беспроводного датчика для мониторинга температуры в зонах установки ортопедических имплантов охватывает несколько ключевых этапов. На начальном этапе разрабатывается беспроводной температурный датчик, учитывающий медицинские требования, включая выбор материалов, минимизацию размеров и обеспечение биосовместимости. Далее проводится его калибровка и тестирование, направленные на проверку точности и стабильности работы в различных температурных диапазонах. После этого следуют лабораторные испытания, в рамках которых датчики применяются в моделях, имитирующих условия изменения температуры имплантов под воздействием воспалительных процессов и других факторов. Затем осуществляются полевые испытания, позволяющие оценить эффективность работы датчиков в реальных условиях, например, при мониторинге температуры у пациентов. Полученные данные анализируются для оценки точности измерений, реакции на температурные изменения и выявления возможных осложнений. На завершающем этапе проводится оценка безопасности и долговечности датчиков, включая их надежность при длительном контакте с тканями организма.



Такой подход позволяет комплексно оценить эффективность технологии беспроводных датчиков для мониторинга температуры и её потенциальное применение в медицине.

Методы исследования диагностики инфекций в местах установки ортопедических имплантатов могут включать как традиционные лабораторные исследования, так и современные технологии, такие как биомедицинские датчики и методы визуализации. Комплексный подход в диагностике позволяет повысить точность выявления инфекционного процесса. Клинические методы диагностики включают физикальное обследование, в ходе которого врач оценивает внешние признаки воспаления, такие как покраснение, отечность, повышение температуры в области имплантата и болезненность. Сбор анамнеза пациента также играет важную роль, поскольку информация о предыдущих инфекциях, осложнениях и общих симптомах помогает установить возможную причину воспаления. Лабораторные методы включают общий анализ крови, выявляющий повышение уровня лейкоцитов и специфических маркеров воспаления, таких как С-реактивный белок и скорость оседания эритроцитов. Микробиологическое исследование, основанное на получении образца из зоны имплантата и его последующем посеве на питательные среды, позволяет идентифицировать возбудителей инфекции. Полимеразная цепная реакция (ПЦР) дает возможность определить ДНК или РНК патогенных микроорганизмов в тканях или жидкостях, полученных из области имплантата, что делает диагностику более точной и специфичной.

Визуализационные методы включают рентгенографию, магнитно-резонансную томографию (МРТ) и компьютерную томографию (КТ). Рентгенография не всегда позволяет непосредственно увидеть инфекцию, но помогает выявить осложнения, такие как остеомиелит, повреждение костной ткани или отторжение имплантата. МРТ является более чувствительным методом для обнаружения воспалений и инфекций в мягких тканях и костях, а также позволяет визуализировать повреждения, вызванные инфекцией, включая абсцессы и остеомиелит. КТ обеспечивает более детальное исследование костной ткани и окружающих структур в области имплантата, что особенно важно при подозрении на глубокие инфекции или осложнения.

Среди инновационных методов диагностики выделяются биосенсоры и беспроводные датчики, а также оптические и ультразвуковые методы. Современные технологии, такие как беспроводные температурные датчики, позволяют контролировать изменения температуры в области имплантата, что может свидетельствовать о развитии инфекции или воспалительного процесса. Также используются сенсоры для измерения уровня pH, способные указать на наличие инфекции. Ультразвуковые методы помогают получать изображения тканей вокруг имплантата, что способствует диагностике воспалений, а оптические технологии, включая флуоресцентную диагностику, находятся на стадии разработки и предназначены для быстрого выявления бактерий в биологических жидкостях.

Генетические методы, такие как методы секвенирования (метагеномика), позволяют изучать микробиомы в зоне имплантата и выявлять присутствие патогенных микроорганизмов даже в небольших количествах. Это особенно полезно при диагностике инфекций, которые не сопровождаются ярко выраженными симптомами.

Методы функциональной диагностики включают термографию и оценку биомаркеров. Инфракрасная термография применяется для мониторинга температуры поверхности тела в местах установки имплантатов, позволяя обнаруживать ранние признаки воспаления или инфекции без инвазивных вмешательств. Методы оценки

биомаркеров, такие как измерение уровней интерлейкинов и других молекул, используются для оценки состояния пациента и определения наличия воспалительных процессов.

Методы исследования современных диагностических методов инфекций в местах установки ортопедических имплантатов могут включать в себя как традиционные, так и инновационные подходы. Важно понимать, что целью таких исследований является выявление наиболее эффективных и точных методов диагностики для своевременного обнаружения инфекций и других осложнений [10]. Рассмотрим основные методы, которые могут быть использованы в рамках исследования современных диагностических технологий:

Анализ научных публикаций: это первый шаг, который помогает понять, какие методы диагностики инфекции в местах установки ортопедических имплантатов используются в настоящее время. Изучение научных статей, клинических рекомендаций и отчетов позволяет определить наиболее популярные и эффективные методы.

Метанализ и систематический обзор могут быть проведены для выявления лучших практик диагностики инфекций, что позволяет суммировать данные из разных источников и получить более обоснованные выводы о точности и эффективности различных методов. Микробиологическое исследование, включая посевы и ПЦР, предполагает анализ образцов тканей или жидкости, взятых из зоны имплантата, для выявления бактерий, грибов или других патогенов, что способствует точной идентификации инфекции. Полимеразная цепная реакция (ПЦР) позволяет определить генетический материал микроорганизмов, являясь высокоточным методом диагностики, особенно в случаях, когда стандартные анализы не дают четкого результата. Биохимические тесты и анализ крови включают оценку уровня маркеров воспаления, таких как С-реактивный белок и прокальцитонин, что помогает определить наличие инфекции. Клинические исследования и тестирование технологий направлены на изучение эффективности различных диагностических методов на пациентах с ортопедическими имплантатами, включая тестирование новых технологий, таких как беспроводные датчики температуры, ультразвуковая диагностика и другие сенсоры.

Сравнительные исследования направлены на анализ эффективности старых и новых методов диагностики. Например, проводится сравнение точности рентгенографии и магнитно-резонансной томографии (МРТ) при диагностике инфекций в местах установки имплантатов.

Методы визуализации включают рентгенографию, которая позволяет выявлять изменения в костной ткани, воспаление и возможные осложнения. Магнитно-резонансная томография используется для более точного обнаружения мягкотканевых инфекций и воспалительных процессов в области имплантатов, включая стандартные и специализированные методы с контрастными веществами. Компьютерная томография (КТ) применяется для детального изучения состояния костной ткани и диагностики таких осложнений, как остеомиелит.

Современные технологии и инновации расширяют возможности диагностики. Разработка и тестирование беспроводных биосенсоров позволяет в режиме реального времени мониторить температуру и другие физиологические показатели в зоне установки имплантатов, что помогает своевременно выявлять признаки воспаления или инфекции. Инфракрасная термография используется для контроля температуры в местах установки имплантатов, а регистрация изменений температуры может служить ранним индикатором воспалительных процессов. Методы молекулярной диагностики, включая секвенирование

ДНК и метагеномику, позволяют выявлять патогенные микроорганизмы на молекулярном уровне, обеспечивая более точную и раннюю диагностику инфекционных заболеваний.

Оценка чувствительности и специфичности методов: исследование точности различных методов диагностики, включая их чувствительность (способность выявить инфекцию) и специфичность (способность отличить инфекцию от других заболеваний). Это может включать статистическую обработку данных из клинических испытаний и анализ показателей каждого метода [11].

Экономический анализ: оценка стоимости различных диагностических методов и технологий, а также их воздействия на стоимость лечения и дальнейшую реабилитацию пациентов. Включает исследование того, насколько эффективно применение новых технологий снижает расходы на лечение осложнений.

Оценка качества жизни пациентов: исследования, направленные на понимание того, как различные диагностические методы и вмешательства влияют на качество жизни пациентов, включая их восприятие боли, страх перед инфекциями и эмоциональное состояние.

### **Результаты и обсуждение.**

Результаты и обсуждение исследования технологии беспроводного датчика для мониторинга температуры в зонах установки ортопедических имплантатов могут включать следующие ключевые моменты.

В ходе исследований было установлено, что беспроводные датчики обеспечивают высокую точность измерений температуры, с малой погрешностью в диапазоне от 36°C до 40°C, что является критически важным для мониторинга температуры в области установки имплантатов. Погрешности измерений составляли менее 1°C, что соответствует стандартам медицинских приборов. Реакция на изменения температуры: Датчики успешно фиксировали небольшие колебания температуры, которые могли быть связаны с воспалительными процессами, инфекциями или отторжением имплантата [12]. Были зарегистрированы повышенные температуры в зоне установки имплантатов, что может служить индикатором возможных осложнений, таких как инфицирование или чрезмерное трение имплантата.

Надежность беспроводной передачи данных: технология беспроводной передачи данных показала хорошую стабильность в условиях, имитирующих реальные медицинские процедуры. Использование беспроводных технологий позволило избежать дополнительных инвазивных вмешательств и обеспечить непрерывный мониторинг без необходимости в частых физически подключаемых датчиках [13]. Продолжительность работы: Испытания на долговечность показали, что датчики могут работать в организме пациента на протяжении нескольких месяцев без значительного ухудшения работы, что предполагает их возможность для долгосрочного мониторинга состояния пациента. Биосовместимость и безопасность: датчики, использованные в исследовании, продемонстрировали хорошую биосовместимость с человеческими тканями, что подтверждается отсутствием воспалений или других негативных реакций в месте установки. Однако, необходимость в дальнейшем тестировании для выявления долгосрочных эффектов остается актуальной.

Преимущества беспроводных датчиков: Основным преимуществом данной технологии является возможность постоянного мониторинга температуры без необходимости в инвазивных вмешательствах. Это важно для раннего выявления патологических изменений и предотвращения осложнений, таких как инфекции или

отторжение имплантата. Использование беспроводных датчиков позволяет значительно повысить качество ухода за пациентами и снизить риски, связанные с инфекциями, без необходимости многократных госпитализаций [14]. Ограничения исследования: несмотря на положительные результаты, существует несколько ограничений, таких как ограниченная длительность полевых испытаний (тестирование проводилось в условиях короткосрочного наблюдения), что может не дать полной картины о долгосрочной стабильности работы датчиков в организме. Также стоит отметить, что для проведения более масштабных клинических испытаний потребуются оценка воздействия на пациента в условиях реального времени. Будущие направления: для улучшения результатов необходимы дальнейшие исследования в области повышения чувствительности датчиков, а также разработка алгоритмов для автоматического анализа полученных данных. Потенциал улучшения заключается в интеграции данных о температуре с другими показателями состояния пациента, что позволит создать систему для комплексного мониторинга и предотвращения осложнений. Клиническое применение: Использование таких датчиков в реальной клинической практике может стать важным инструментом для врачей, позволяя оперативно реагировать на изменение состояния пациента и предотвращать развитие негативных последствий в послеоперационный период. Это особенно актуально для пациентов, которые подверглись установке имплантатов в костях или суставах, где температура может служить индикатором воспаления или инфекции [15].

В целом, исследование показало, что беспроводные датчики для мониторинга температуры в зонах установки ортопедических имплантатов имеют большой потенциал для применения в медицине, особенно в контексте улучшения ухода за пациентами и предотвращения осложнений.

### **Заключение.**

В данном исследовании проведены анализ существующих работ зарубежных авторов по данной теме. В ходе исследовательской работы проведен анализ методов диагностики которые имеет свои преимущества и ограничения, и их выбор зависит от клинической ситуации, доступности оборудования и состояния пациента. Изучены ряды комплексных подходов, который сочетает в себе традиционные методы (клинический осмотр, анализ крови) и более современные технологии (МРТ, ПЦР, беспроводные датчики).

Анализируются современные методы диагностики инфекций в местах установки ортопедических имплантатов должны быть точными, быстрыми и эффективными. Комбинированное использование традиционных лабораторных методов, таких как посевы и ПЦР, с новыми технологиями, такими как биосенсоры и МРТ, может обеспечить надежную диагностику. Исследованы внедрение инновационных методов, таких как беспроводные датчики и молекулярная диагностика, могут значительно улучшить процессы мониторинга и предотвращения инфекций у пациентов с ортопедическими имплантатами.

## Список литературы

1. Cody O'Connor<sup>1</sup>, Asimina Kiourti (2017). Wireless Sensors for Smart Orthopedic Implants. *Springer International Publishing Switzerland*. DOI: [10.1007/s40735-017-0078-z](https://doi.org/10.1007/s40735-017-0078-z)
2. Ahmed Hassan Kamal, Ossama Mohammed Zakaria, Rabab Abbas Majzoub, El Walid Fadul Nasir (2023). Artificial intelligence in orthopedics. A qualitative exploration of the surgeon perspective. *Kamal et al. • Medicine 102:24*. DOI: [10.1097/MD.00000000000034071](https://doi.org/10.1097/MD.00000000000034071)
3. Pengran Liu, Jiayao Zhang, Songxiang Liu, Tongtong Huo, Jiajun He, Mingdi Xue, Ying Fang, Honglin Wang, Yi Xie, Mao Xie, Dan Zhang, Zhewei Ye (2024). Application of artificial intelligence technology in the field of orthopedics: a narrative review. *Artificial Intelligence Review 57:13*. <https://doi.org/10.1007/s10462-023-10638-6>
4. Gengming Zhang, Linhua Deng, Zhongjing Jiang, Gang Xiang, Zhuotong Zeng, Hongqi Zhang and Yunjia Wang (2024). Titanium nanoparticles released from orthopedic implants induce muscle fibrosis via activation of SNAI2. *Zhang et al. Journal of Nanobiotechnology 22:522*. <https://doi.org/10.1186/s12951-024-02762-4>
5. О.В. Крайнюк, Ю.В. Буц, В.В. Барбашин (2024). Метрологічний контроль датчиків моніторингу умов праці з використанням штучного інтелекту. *Комунальне господарство міст, 2024, том 3, випуск 184*. doi: [10.33042/2522-1809-2024-3-184-216-222](https://doi.org/10.33042/2522-1809-2024-3-184-216-222)
6. Enmao Xiang, Cedryck Vaquette, Shulei Liu, Nimal Raveendran, Benjamin L. Schulz, Nunthawan Nowwarote, Matthew Dargusch, Abdalla Abdal-hay, Benjamin P. J. Fournier, and Sašo Ivanovski (2024). Biomimetic Surface Nanoengineering of Biodegradable Zn-Based Orthopedic Implants for Enhanced Biocompatibility and Immunomodulation. *Advanced Functional Materials published by Wiley-VCH GmbH 2410033 (1 of 23)*. <https://doi.org/10.1002/adfm.202410033>
7. Bin'en Nie a, Shicheng Huo a, Xinhua Qu a, Jingjing Guo b, Xi Liu c, Qimin Hong a, You Wang a, Jianping Yang b, Bing Yue a (2022). Bone infection site targeting nanoparticle-antibiotics delivery vehicle to enhance treatment efficacy of orthopedic implant related infection. *Bioactive Materials 16, 134–148*. <https://doi.org/10.1016/j.bioactmat.2022.02.003>
8. Mateusz Kopec, Ved Prakash Dubey, Adam Brodecki, Zbigniew Kowalewski (2021). Failure Analysis of Orthopedic Implants. *Conference Paper*.
9. Jalal T. Alakhras, MBBS and Tarif M. Alakhras (2020). Effect of cold weather on patients with orthopedic implants. *Journal of Taibah University Medical Sciences 1/4*. <https://doi.org/10.1016/j.jtumed.2020.05.001>
10. Eleni Tsalkitzi, Dimitrios Kitridis, Elena Heinz, Christina Hionidou, Kornilia Givissi, Panagiotis Givissis (2024). The Metal in My Body: Patients' Perception and Attitude Toward Orthopedic Implants. *Open Access Original Article*. DOI: [10.7759/cureus.56493](https://doi.org/10.7759/cureus.56493)
11. Tong Bao, Jinyi Wang, Yang Chen, Feng Xu, Guanzhong Qiao, Fei Li (2023). The top 100 most cited articles on magnesium alloy orthopedic implants: A bibliometric and visualized analysis. *Theoretical and Natural Science 15(1):201-224*. DOI: [10.54254/2753-8818/15/20240481](https://doi.org/10.54254/2753-8818/15/20240481)
12. Mohammad J. Mirzaali and Amir A. Zadpoor (2024). Orthopedic meta-implants. *APL Bioeng. 8, 010901*. <https://doi.org/10.1063/5.0179908>
13. Wenqing Liang, Chao Zhou, Juqin Bai, Hongwei Zhang, Hengguo Long, Bo Jiang, Haidong Dai, Jiangwei Wang, Hengjian Zhang and Jiayi Zhao (2023). Current developments and future perspectives of nanotechnology in orthopedic implants: an updated review. *Front. Bioeng. Biotechnol. 12:1342340*. <https://doi.org/10.3389/fbioe.2024.1342340>
14. Mark Ren, Paul H. Yi (2022). Artificial intelligence in orthopedic implant model classification: a systematic review. *Skeletal Radiol, 51(2):407-416*. DOI: [10.1007/s00256-021-03884-8](https://doi.org/10.1007/s00256-021-03884-8)

15. Dinesh Kumar Nishad Nishad, Vandana Rani Verma, Saifullah Khalid, Vinay Kumar Singh Singh (2024). Enhanced Security in Wireless Sensor Networks Using Artificial Intelligence. *Research Article*. DOI: <https://doi.org/10.21203/rs.3.rs-5032504/v1>

## References

1. Cody O'Connor<sup>1</sup>, Asimina Kiourti (2017). Wireless Sensors for Smart Orthopedic Implants. *Springer International Publishing Switzerland*. DOI: [10.1007/s40735-017-0078-z](https://doi.org/10.1007/s40735-017-0078-z)
2. Ahmed Hassan Kamal, Ossama Mohammed Zakaria, Rabab Abbas Majzoub, El Walid Fadul Nasir (2023). Artificial intelligence in orthopedics. A qualitative exploration of the surgeon perspective. *Kamal et al. • Medicine 102:24*. DOI: [10.1097/MD.00000000000034071](https://doi.org/10.1097/MD.00000000000034071)
3. Pengran Liu, Jiayao Zhang, Songxiang Liu, Tongtong Huo, Jiajun He, Mingdi Xue, Ying Fang, Honglin Wang, Yi Xie, Mao Xie, Dan Zhang, Zhewei Ye (2024). Application of artificial intelligence technology in the field of orthopedics: a narrative review. *Artificial Intelligence Review 57:13*. <https://doi.org/10.1007/s10462-023-10638-6>
4. Gengming Zhang, Linhua Deng, Zhongjing Jiang, Gang Xiang, Zhuotong Zeng, Hongqi Zhang and Yunjia Wang (2024). Titanium nanoparticles released from orthopedic implants induce muscle fibrosis via activation of SNAI2. *Zhang et al. Journal of Nanobiotechnology 22:522*. <https://doi.org/10.1186/s12951-024-02762-4>
5. V. V. Barbashin, O. Krainiuk, Yuriy Vasylyovych Buts (2024). Metrological control of sensors for monitoring working conditions using artificial intelligence. *Municipal economy of cities*. doi:[10.33042/2522-1809-2024-3-184-216-222](https://doi.org/10.33042/2522-1809-2024-3-184-216-222)
6. Enmao Xiang, Cedryck Vaquette, Shulei Liu, Nimal Raveendran, Benjamin L. Schulz, Nunthawan Nowwarote, Matthew Dargusch, Abdalla Abdal-hay, Benjamin P. J. Fournier, and Sašo Ivanovski (2024). Biomimetic Surface Nanoengineering of Biodegradable Zn-Based Orthopedic Implants for Enhanced Biocompatibility and Immunomodulation. *Advanced Functional Materials published by Wiley-VCH GmbH 2410033 (1 of 23)*. <https://doi.org/10.1002/adfm.202410033>
7. Bin'en Nie a, Shicheng Huo a, Xinhua Qu a, Jingjing Guo b, Xi Liu c, Qimin Hong a, You Wang a, Jianping Yang b, Bing Yue a (2022). Bone infection site targeting nanoparticle-antibiotics delivery vehicle to enhance treatment efficacy of orthopedic implant related infection. *Bioactive Materials 16, 134–148*. <https://doi.org/10.1016/j.bioactmat.2022.02.003>
8. Mateusz Kopec, Ved Prakash Dubey, Adam Brodecki, Zbigniew Kowalewski (2021). Failure Analysis of Orthopedic Implants. *Conference Paper*.
9. Jalal T. Alakhras, MBBS and Tarif M. Alakhras (2020). Effect of cold weather on patients with orthopedic implants. *Journal of Taibah University Medical Sciences 1/4*. <https://doi.org/10.1016/j.jtumed.2020.05.001>
10. Eleni Tsalkitzi, Dimitrios Kitridis, Elena Heinz, Christina Hionidou, Kornilia Givissi, Panagiotis Givissis (2024). The Metal in My Body: Patients' Perception and Attitude Toward Orthopedic Implants. *Open Access Original Article*. DOI: [10.7759/cureus.56493](https://doi.org/10.7759/cureus.56493)
11. Tong Bao, Jinyi Wang, Yang Chen, Feng Xu, Guanzhong Qiao, Fei Li (2023). The top 100 most cited articles on magnesium alloy orthopedic implants: A bibliometric and visualized analysis. *Theoretical and Natural Science 15(1):201-224*. DOI: [10.54254/2753-8818/15/20240481](https://doi.org/10.54254/2753-8818/15/20240481)
12. Mohammad J. Mirzaali and Amir A. Zadpoor (2024). Orthopedic meta-implants. *APL Bioeng. 8, 010901*. <https://doi.org/10.1063/5.0179908>
13. Wenqing Liang, Chao Zhou, Juqin Bai, Hongwei Zhang, Hengguo Long, Bo Jiang, Haidong Dai, Jiangwei Wang, Hengjian Zhang and Jiayi Zhao (2023). Current developments and

future perspectives of nanotechnology in orthopedic implants: an updated review. *Front. Bioeng. Biotechnol.* 12:1342340. <https://doi.org/10.3389/fbioe.2024.1342340>

14. Mark Ren, Paul H. Yi (2022). Artificial intelligence in orthopedic implant model classification: a systematic review. *Skeletal Radiol*, 51(2):407-416. DOI: [10.1007/s00256-021-03884-8](https://doi.org/10.1007/s00256-021-03884-8)

15. Dinesh Kumar Nishad Nishad, Vandana Rani Verma, Saifullah Khalid, Vinay Kumar Singh Singh (2024). Enhanced Security in Wireless Sensor Networks Using Artificial Intelligence. *Research Article*. DOI: <https://doi.org/10.21203/rs.3.rs-5032504/v1>

## ANALYSIS OF THE RESEARCH OF WIRELESS SENSOR TECHNOLOGY FOR MONITORING TEMPERATURE IN ORTHOPEDIC IMPLANTS INSTALLATION AREAS

**Abstract.** *The aim of this study is to develop and approve an innovative methodological concept of orthopaedic implant performance monitoring using modern wireless sensor technologies and artificial intelligence algorithms. The paper provides a detailed analysis of the existing methods of diagnostics of implant condition, reveals their technical and methodological limitations, and analyses the current trends in the field of medicine aimed at improving the reliability and safety of implantation. The proposed technique combines the collection of temperature and other physical data by means of high-precision wireless sensors and their processing with the use of machine learning algorithms to predict possible deviations in implant operation and timely detection of the initial signs of wear or damage.*

*The experimental part of the research includes testing of the developed system in real clinical conditions that allowed to obtain significant statistical evidence of its efficiency and accuracy. The results showed that the application of the innovative approach allows not only to improve the accuracy of diagnosis, but also to significantly reduce the response time to potential complications, which is important for surgical intervention and reducing the risk of unfavourable outcomes. The presented approach opens new perspectives for further research in the field of medical implant monitoring, contributing to the development of personalised medicine and improving the quality of life of patients.*

**Keywords:** *Orthopedic implants, wireless temperature sensor, artificial intelligence, digitalization, diagnostics, patient.*

## ОРТОПЕДИЯЛЫҚ ИМПЛАНТТАРДЫ ОРНАТУ АЙМАҚТАРЫНДА ТЕМПЕРАТУРАНЫ БАҚЫЛАУ ҮШІН СЫМСЫЗ ДАТЧИКТЕР ТЕХНОЛОГИЯСЫН ЗЕРТТЕУДІ ТАЛДАУ

**Аңдатпа.** Бұл зерттеудің мақсаты заманауи сымсыз сенсорлық технологиялар мен жасанды интеллект алгоритмдерін пайдалана отырып, ортопедиялық импланттардың өнімділігін бақылаудың инновациялық әдістемелік тұжырымдамасын әзірлеу және сынау болып табылады. Жұмыста импланттардың жағдайын диагностикалаудың қолданыстағы әдістеріне егжей-тегжейлі талдау, олардың техникалық және әдістемелік шектеулерін анықтау, сондай-ақ имплантацияның сенімділігі мен қауіпсіздігін арттыруға бағытталған медицина саласындағы ағымдағы тенденцияларды талдау қарастырылған. Ұсынылған әдістеме импланттардың жұмысындағы ықтимал ауытқуларды болжау және тозу немесе зақымданудың бастапқы белгілерін уақтылы анықтау үшін жоғары дәлдіктегі сымсыз сенсорларды пайдалана отырып, температура мен басқа физикалық



деректерді жинауды және оларды машиналық оқыту алгоритмдерін пайдалана отырып өңдеуді біріктіреді.

Зерттеудің тәжірибелік бөлімі әзірленген жүйені нақты клиникалық жағдайларда сынауды қамтиды, бұл оның тиімділігі мен дәлдігінің маңызды статистикалық дәлелдерін алуға мүмкіндік берді. Нәтижелер инновациялық тәсілді қолдану диагностикалық дәлдікті жақсартып қана қоймай, сонымен қатар хирургиялық араласу және қолайсыз нәтижелердің қаупін азайту үшін маңызды болып табылатын ықтимал асқынуларға жауап беру уақытын айтарлықтай қысқартатынын көрсетті. Ұсынылған тәсіл медициналық имплантаттардың жай-күйін бақылау саласындағы әрі қарайғы зерттеулердің жаңа перспективаларын ашады, дербестендірілген медицинаның дамуына және пациенттердің өмір сүру сапасын жақсартуға ықпал етеді.

**Түйін сөздер.** Ортопедиялық имплантаттар, сымсыз температура сенсоры, жасанды интеллект, цифрландыру, диагностика, пациент.

#### Авторлар туралы мәлімет

Баталова Мадина Есимханқызы	«ЦТЖЖИМ» докторанты, Д.Серікбаев атындағы ШҚТУ, Өскемен қ., Қазақстан, E-mail: <a href="mailto:esimkhan_kizi.m@mail.ru">esimkhan_kizi.m@mail.ru</a>
Rawel Komada	PhD, professor of Electronics and Information Technology at Faculty of Electrical Engineering and Computer Science of Lublin University of Technology, Poland, E-mail: <a href="mailto:p.komada@pollub.pl">p.komada@pollub.pl</a>
Алибекқызы Карлыгаш	«ЦТЖЖИМ» PhD., философия докторы қауымдастырылған профессоры, Д.Серікбаев атындағы ШҚТУ, Өскемен қ., Қазақстан E-mail: <a href="mailto:karlygash.alibekkyzy@mail.ru">karlygash.alibekkyzy@mail.ru</a>
Бугубаева Алина Жанатбековна	«Цифрлық инженерия және IT-Аналитика» кафедрасының қауымдастырылған профессоры, PhD философия докторы, Қазтұтынуодағы Қарағанды университеті, Қарағанды қ., Қазақстан E-mail: <a href="mailto:alina_bugubayeva@mail.ru">alina_bugubayeva@mail.ru</a>
Кунапьянова Маржан Советхановна	«ЦТЖЖИМ» оқытушысы, Д.Серікбаев атындағы ШҚТУ, Өскемен қ., Қазақстан, E-mail: <a href="mailto:Kmarzhana92@mail.ru">Kmarzhana92@mail.ru</a>

#### Сведение об авторах

Баталова Мадина Есимханқызы	Докторант «ШЦТиИИ», ВКТУ им.Д. Серикбаева, г. Усть-Каменогорск, Казахстан, E-mail: <a href="mailto:esimkhan_kizi.m@mail.ru">esimkhan_kizi.m@mail.ru</a>
Rawel Komada	PhD, professor of Electronics and Information Technology at Faculty of Electrical Engineering and Computer Science of Lublin University of Technology, Poland, E-mail: <a href="mailto:p.komada@pollub.pl">p.komada@pollub.pl</a>
Алибекқызы Карлыгаш	Ассоциированный профессор «ШЦТиИИ», доктор философии PhD ВКТУ им.Д. Серикбаева, г. Усть-Каменогорск, Казахстан E-mail: <a href="mailto:karlygash.alibekkyzy@mail.ru">karlygash.alibekkyzy@mail.ru</a>
Бугубаева Алина Жанатбековна	Ассоциированный профессор кафедры «Цифровая инженерия и IT-Аналитика», доктор философии PhD, Карагандинский Университет Казпотребсоюза, г. Караганда, Казахстан, E-mail: <a href="mailto:alina_bugubayeva@mail.ru">alina_bugubayeva@mail.ru</a>



Кунапьянова Маржан Советхановна	Преподаватель «ШЦТИИИ», ВКТУ им.Д. Серикбаева, г. Усть-Каменогорск, Казахстан, E-mail: <a href="mailto:Kmarzhana92@mail.ru">Kmarzhana92@mail.ru</a>
---------------------------------------	---

#### Information about the authors

Batalova Madina	doctoral student of "SDTaAI", D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: <a href="mailto:esimkhan_kizi.m@mail.ru">esimkhan_kizi.m@mail.ru</a>
Paweł Komada	PhD, professor of Electronics and Information Technology at Faculty of Electrical Engineering and Computer Science of Lublin University of Technology, Poland, E-mail: <a href="mailto:p.komada@pollub.pl">p.komada@pollub.pl</a>
Alibekkyzy Karlygash	Senior lecturer of "SDTaAI", Doctor of Philosophy PhD of D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: <a href="mailto:karlygash.alibekkyzy@mail.ru">karlygash.alibekkyzy@mail.ru</a>
Bugubayeva Alina	Associate Professor "Digital Engineering and IT-Analytics", Doctor of Philosophy PhD, Karaganda University of Kazpotrebsoyuz, Karaganda, Kazakhstan, E-mail: <a href="mailto:alina_bugubayeva@mail.ru">alina_bugubayeva@mail.ru</a>
Kunapianova Marzhan	Lecturer of "SDTaAI", D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: <a href="mailto:Kmarzhana92@mail.ru">Kmarzhana92@mail.ru</a>

УДК 004.9

МРНТИ 20.51.19

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_15](https://doi.org/10.53364/24138614_2025_36_1_15)А.Б. Саметова<sup>1</sup>, Д.Р. Рашидинов<sup>1\*</sup>, Л.К. Найзабаева<sup>1</sup>, М.У. Сулейменова<sup>1</sup><sup>1</sup>Международный университет информационных технологий, г. Алматы, КазахстанE-mail: [damir.dmr88@gmail.com](mailto:damir.dmr88@gmail.com)\*

## РАЗРАБОТКА СИСТЕМЫ ДОКУМЕНТООБОРОТА В ОБРАЗОВАНИИ

**Аннотация.** В данной статье рассматривается процесс перехода от традиционного к электронному документообороту в образовательных учреждениях. Проведен анализ функций и процессов в рамках автоматизированной информационной системы образовательной организации. Рассмотрены ключевые подходы к внедрению электронного документооборота с точки зрения системного подхода. Основываясь на практическом опыте, авторы предлагают концепцию организации и проведения обучения в системе повышения квалификации специалистов.

Автоматизация документооборота в учебных заведениях особенно актуальна в таких процессах, как учет входящей и исходящей корреспонденции, обработка внутренних документов, согласование административных и договорных документов, а также контроль исполнения поручений. Внедрение систем электронного документооборота (СЭД) требует интеграции справочников в единое хранилище данных и синхронизации процессов обработки документов с другими бизнес-процессами учреждения.

**Ключевые слова:** делопроизводство, электронный документ, управление образовательной организацией, эффективность образовательного управления, информационные системы, информационные технологии, электронные системы управления.

### Введение.

Современный мир характеризуется стремительным технологическим развитием, что требует эффективного управления информацией. Документированная информация является основой принятия решений в организациях, а её качество напрямую влияет на эффективность управления [1]. Рост объема информации требует внедрения новых технологий для ее обработки, и в этом контексте электронные документы заменяют бумажные носители, обеспечивая оперативность и удобство управления данными.

Это также открывает беспрецедентные возможности. Сегодняшнему менеджеру не нужно ждать, пока клерк-делопроизводитель найдет отметку нужного документа в журнале входящей документации, затем прокручивать соответствующую папку, чтобы найти сам документ. Все эти функции могут быть полностью автоматизированы, что позволяет значительно увеличить скорость доступа к документу, сократить время на подготовку решения и, соответственно, обеспечить эффективность управления [2].

Документационное обеспечение управления в образовательных учреждениях требует создания многих видов управленческих документов, без которых невозможно решить проблемы планирования, финансирования, оперативного управления, укомплектования персоналом учреждения. Основными задачами делопроизводства в образовательном учреждении являются сведение информационных потоков к оптимальному минимуму и

упрощение и удешевление сбора, обработки и передачи информации с использованием новейших технологий автоматизации этих процессов.

В результате степень организации офисной работы напрямую влияет на качество управления образовательным учреждением. Администрирование информации и документации, электронный документооборот, защита информации и развитие электронных архивов относятся к числу наиболее актуальных вопросов.

Эффективное управление документами является ключевым фактором для повышения управленческой эффективности и конкурентоспособности организации. Согласно исследованиям, сотрудники тратят значительную часть рабочего времени на работу с документами, поэтому улучшение процессов их создания и распространения может сократить время поиска и повысить точность информации [3].

Документы играют ключевую роль в бизнесе, служа отчетами о деятельности и основой коммерческих операций. Они занимают значительную часть рабочего времени сотрудников и часто являются основным доказательством в спорных ситуациях, подтверждая важность точного и эффективного управления документами [4].

Информатизация образования, или интеграция новых информационных технологий в образовательную систему, является одним из главных приоритетов процесса информатизации современного общества. Используя автоматизированные базы данных научных и педагогических знаний, информационные и методические материалы, организационные и управленческие ресурсы, а также сети связи, теперь можно усовершенствовать механизмы управления системой образования.

Важным методом создания преимуществ в конкурентной среде является управление образовательными учреждениями на основе информационных технологий. В связи с этим основными мероприятиями в развитии информатизации становятся создание надежной и эффективной инфраструктуры, внедрение унифицированных способов доступа к корпоративным данным, улучшение управляемости всего комплекса информационных ресурсов, а также обеспечение соответствия инфраструктуры стратегическим целям вуза.

Одним из стратегических направлений Казахстана является программа «Цифровой Казахстан», направленная на развитие информационных технологий, цифровых платформ и сокращение использования бумажных носителей. В образовательной сфере это выражается в переходе на электронные образовательные ресурсы, цифровые архивы и интеграцию научных публикаций в онлайн-системы [5].

Такой веб-проект будет разработан в качестве основы для бизнес-планов, творческих проектов, свежих концепций и бизнес-моделей для различных полиграфических предприятий на основе тщательного изучения и комментариев экспертов. Можно повысить темпы развития практически всех секторов экономики за счет повышения производительности труда и улучшения управления и производственного процесса благодаря комплексному использованию информационных технологий практически во всех сферах экономики [6].

На данный момент существует три типа электронных периодических порталов:

- Параллельные (идентичные друг другу электронные версии научного портала);
- Интегрированная (дополняющая печатную и печатно-электронную версии научного портала);
- Универсальный. Они являются оригинальными разработками разных основателей, и методы работы с такими научными сайтами могут сильно отличаться. Такие сайты часто специфичны в использовании, имеют уникальный интерфейс, который затрудняет поиск информации, а также плохо интегрированы друг с другом. Для решения вышеуказанной проблемы предлагается разработать модель универсального электронного научного портала [7].

Создание интегрированной территориально распределенной системы, объединяющей федеральные и специализированные порталы, является важным условием для реализации принципов доступности, открытости и непрерывности образования, повышения эффективности обучения через предоставление образовательным учреждениям и пользователям доступа к современным электронным образовательным ресурсам и инструментам поддержки образовательного процесса, совершенствования научных исследований и разработок в вузах, улучшения управления системой образования, развития международного сотрудничества в сфере образования и науки, а также продвижения казахстанского образования на мировой арене.

Каждый образовательный портал должен быть инструментом для внедрения образовательных технологий, должен обеспечивать поддержку традиционных и дистанционных технологий обучения, механизмов открытого образования и других перспективных образовательных технологий посредством соответствующего информационного, функционального и инструментального контента.

Портал обеспечивает высокий уровень предметных коммуникаций, которого не хватает в самых современных традиционных учебных планах и учебниках. Образовательный портал — это мощное развивающее руководство и постоянно пополняющаяся универсальная энциклопедия знаний [6].

Образовательный портал предоставляет современную среду для внедрения образовательных технологий и различных методологических подходов. В этой среде легко моделируются различные образовательные структуры, есть возможность быстрой и точной оценки качества предлагаемых инноваций, есть инструменты для их улучшения и условия для командной работы, для взаимопроникновения и сочетания различных подходов. Образовательный портал знает, как активно влиять на формирование профессиональных преподавательских кадров и создание условий для их эффективной работы.

#### **Материалы и методы исследования.**

Методология исследования основана на анализе существующих решений в области электронного документооборота, сравнении различных подходов к автоматизации и оценке их эффективности. Были использованы методы анкетирования, интервьюирования и анализа эмпирических данных. Для сбора информации применялся инструмент Google Forms, позволяющий проводить опросы среди сотрудников образовательных организаций.

В настоящее время актуальной задачей является создание целостной интегрированной системы, обеспечивающей бесплатный и быстрый доступ широких групп населения к образовательным ресурсам. Интенсивное развитие Интернета и интернет-технологий обуславливает необходимость интеграции разнообразных информационных ресурсов, приложений и сервисов для максимально полного удовлетворения потребностей различных групп пользователей. Эта проблема стоит достаточно остро во всех сферах человеческой деятельности, и, в частности, это касается образования как одной из ключевых сфер, определяющих развитие общества. Пути решения этой проблемы в настоящее время сосредоточены вокруг концепции создания и функционирования интернет-порталов.



Рисунок 1 – Актуальные проблемы документооборота

На Рисунке 1 хорошо видно, что решение актуальных проблем документооборота удобно представлять в виде цикла (годового или другого, четко выделенного временного отрезка). Начиная с обсуждения и преобразования одной ситуации в начале периода, выходим на новый виток проблем в конце его.

При решении организационных и технических вопросов, связанных с документооборотом в школе, стоит стремиться к поиску решений, которые охватят сразу все обозначенные проблемы. Однако это возможно только при внедрении по-настоящему нового подхода, способного обеспечить качественно иной уровень доступа и вовлеченности всех членов школьного коллектива в процессы документооборота.

Ключевым элементом такого подхода является техническое решение, которое станет основой для последующих изменений в процессах и организационных аспектах работы с документацией. Основной его особенностью является преимущественный отказ от бумажного документооборота [8].

Помочь в увеличении эффективности организации документооборота может внедрение системы электронного документооборота.

Назначение электронной системы документооборота заключается в следующем: увеличение эффективности документооборота; создание единого хранилища электронных документов; стандартизация документопотока; контроль исполнительской дисциплины.

Последние десятилетия интенсивного развития информатики как науки и как реального инструмента социального прогресса характеризуются созданием и использованием различных средств обработки информации, которые инициируют формирование перспективных педагогических технологий, ориентированных на интеллектуальное совершенствование учащегося. Рассмотрим основные направления использования информационно-коммуникационных технологий для развития образовательных информационных ресурсов, составляющих основу образовательного информационного пространства как такого образовательных порталов.

Решение поставленных задач эффективно реализуется через систему порталов для образовательного сообщества, которая призвана обеспечить информационную поддержку образовательного процесса во всех его аспектах для всех уровней образования и форм обучения. Под образовательным порталом мы подразумеваем сетевой узел или комплекс узлов, подключенных к Интернету по высокоскоростным каналам, который имеет развитый

пользовательский интерфейс и обеспечивает единую, концептуальную и содержательную точку зрения, доступ к широкому спектру информационных ресурсов и услуг, ориентированных на конкретную аудиторию.

В случае создания дизайна обрабатываемого сайта наибольшее внимание обычно уделяется дизайну главной страницы. Дизайн внутренних страниц должен основываться на дизайне домашней страницы. Цель главной страницы сайта - удержать внимание аудитории и стимулировать их продолжать пользоваться сайтом. Как правило, такое понятие, как качество специализированного использования веб-сайта, многогранно и связано с веб-дизайном, его шрифтами, навигацией и множеством других требований.

### **Результаты и их обсуждение.**

Процесс управления включает следующие ключевые типовые документируемые операции: сбор и обработку информации из документов; подготовку решений; принятие и документирование решений; доведение решений до исполнителей; выполнение решений; контроль их исполнения; сбор данных о выполнении; передачу информации по вертикальным и горизонтальным связям; хранение и поиск информации.

Каждая из этих операций документируется как в традиционной форме, так и в виде машиноориентированных и машиночитаемых документов. Недокументируемые управленческие действия выполняют вспомогательную, оперативно-организационную функцию и занимают небольшой процент по сравнению с документируемыми процессами. Несмотря на внедрение электронных вычислительных систем, выполнение документируемых функций продолжает требовать значительных интеллектуальных и трудовых ресурсов.

Исследование документооборота любого учреждения возможно благодаря тому, что его деятельность строго регламентирована как по уровням принятия решений, так и по кругу решаемых вопросов. Управленческий процесс включает стадии сбора, анализа и обобщения информации, подготовки и принятия решений, организации их исполнения и завершается контролем выполнения.

На этапе контроля осуществляется движение информации по каналам обратной связи от объекта управления к субъекту. Это включает передачу данных о реализации решений в соответствии с установленной программой. Документооборот служит ключевым инструментом в процессе делопроизводства, обеспечивая движение документов (и, соответственно, информации) в рамках выполнения управленческих задач.

Вся документация организации делится на три документопотока:

- входящие (поступающие) документы;
- исходящие (отправляемые) документы;
- внутренние документы.

По отношению к каждому конкретному исполнителю все документы, с которыми он имеет дело, делятся на несколько категорий:

- входящие, с которыми исполнитель не успел ознакомиться;
- в работе, которые ждут его действий;
- на контроле, по которым он ожидает действий от других исполнителей.

Документооборот составляет часть системы делопроизводства образовательного учреждения и выполняет, по существу, функции системы коммуникации в учреждении. Поэтому цели его функционирования и совершенствования должны быть субординированы целями функционирования и совершенствования делопроизводства образовательного учреждения, которые заключаются в максимально возможном приспособлении системы к условиям ее функционирования. Это позволяет обеспечивать систему управления качественной документной информацией в сроки, оптимальные для принятия и выполнения управленческих решений. Веб-приложение может быть клиент-серверным

приложением, в котором клиент является браузером, а сервер - интернет-сервером (в широком смысле).

Основная часть приложения, как правило, находится на стороне интернет-сервера, который формирует полученные запросы в соответствии с коммерческим обоснованием товара и создает реакцию, отправляемую клиенту. На этом этапе включается браузер, именно он преобразует полученную реакцию сервера в графический интерфейс, приемлемый для пользователя. Клиент-серверная инженерия характеризует общие стандарты организации взаимодействия между серверами, узлами-поставщиками нескольких определенных мощностей (администрациями) и клиентами (заказчиками этих функций). Практическое использование этой технологии называется клиент-серверными инновациями.

Двухуровневая архитектура — это распределение трех основных компонентов между двумя узлами (клиентом и сервером). Двухуровневая архитектура используется в клиент-серверных системах, где сервер отвечает на запросы клиента напрямую и в полном объеме.

Расположение компонентов на стороне клиента или сервера определяет следующие основные модели их взаимодействия в рамках двухуровневой архитектуры (Рисунок 2):

- Сервер терминалов — это распределенное представление данных.
- Файловый сервер — это доступ к удаленной базе данных и файловым ресурсам.
- Сервер базы данных — это удаленный просмотр данных.
- Сервер приложений — это удаленное приложение.

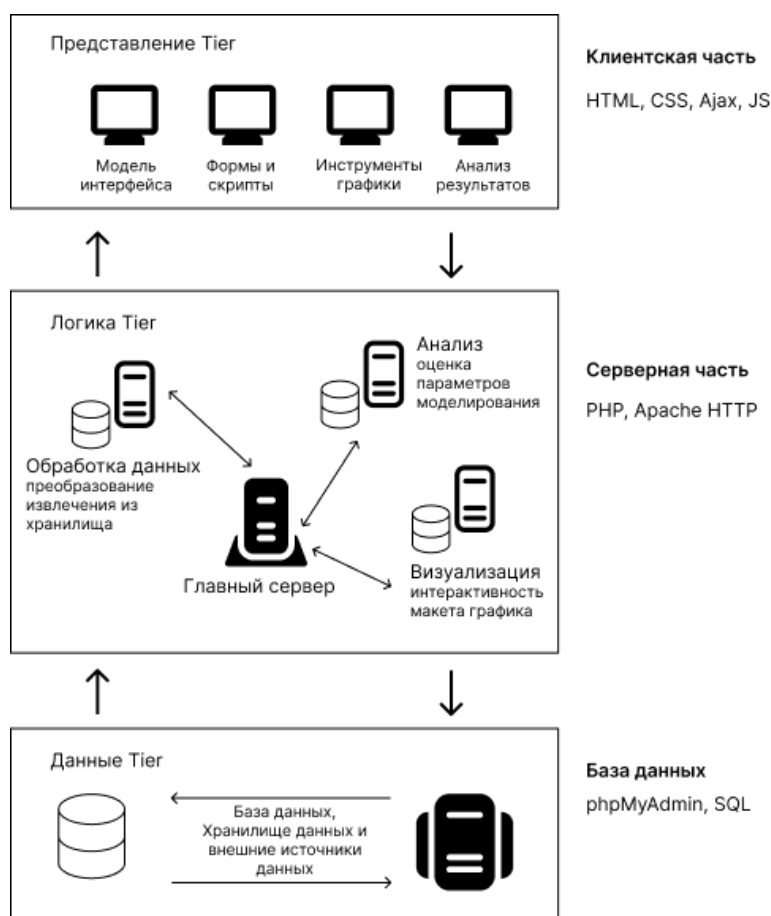


Рисунок 2 – Архитектура программного обеспечения

Пользователь интернет-сайта может просматривать только HTML-код страниц; он не видит, как работают PHP и MySQL. Работа с приложениями PHP и базами данных MySQL

выполняется сервером, на котором установлен интерпретатор PHP и СУБД MySQL. После выполнения любого PHP-кода сервер возвращает HTML-страницу пользователю сайта. пользователь не может просмотреть сам PHP-код.

Для работы с базами данных MySQL вам нужен интерпретатор PHP. Без этого сервер не сможет работать с базами данных MySQL.

На приведенной ниже диаграмме показано, как пользователь взаимодействует с базами данных и приложениями PHP (Рисунок 3).

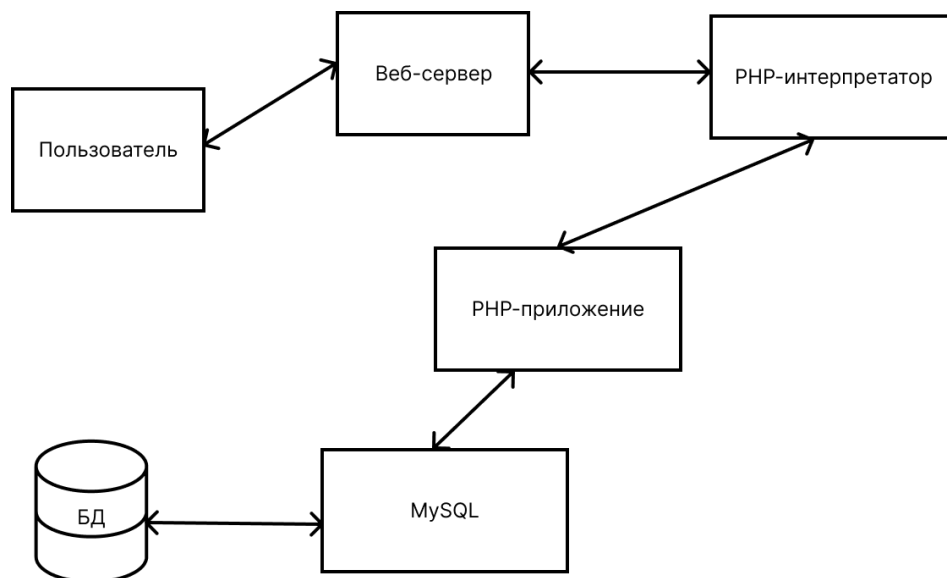


Рисунок 3 – Структурная схема программного обеспечения веб-сервера

Концептуальная модель представляет собой упорядоченное и осмысленное описание моделируемой системы, выполненное на неформальном языке. В процессе создания такого описания имитационной модели определяются ключевые элементы системы, их характеристики и взаимодействия, используя язык, понятный для разработчиков. Для визуализации и структурирования можно применять таблицы, графики, диаграммы и другие средства. Неформальное описание важно как для разработчиков (например, для проверки точности модели или внесения изменений), так и для эффективного взаимодействия с экспертами из других областей.

Для анализа и проектирования объектно-ориентированных программных систем применяются языки визуального моделирования, среди которых наиболее широко используется UML. При разработке спецификации программного обеспечения с помощью UML создаются несколько моделей: логическая, модель использования, реализации, процессов и развертывания [8].

Модель использования отражает функции программного обеспечения с позиции пользователя, тогда как логическая модель фокусируется на основных концепциях программной системы, таких как классы, интерфейсы и другие элементы, обеспечивающие её функциональность.





Рисунок 4 – Схема вариантов использования

Сотрудник научного отдела, который обобщает все данные о научно-исследовательской деятельности в университете. Ниже предоставлены определения использования:

- Вход
- Редактирования доступов
- Редактирования пользователей
- Авторизация и Регистрация
- Редактирования преподавателей
- Отчет и статистика

В системе выделяются четыре основные роли: гость, учитель, руководитель отдела и администратор (рисунок 4).

ER-диаграмма представляет собой модель данных, которая описывает концептуальные схемы предметной области. Это графическое представление предметной области с использованием стандартизированного набора символов. На основе ER-модели создается логическая модель, предназначенная для работы с выбранной СУБД.

Логическая модель отражает процесс проектирования информационной системы, опираясь на модель данных, но не учитывает специфику конкретной СУБД или физических ограничений.

Диаграмма сущностей и связей (ERD или ER-диаграмма) — это вид блок-схемы, отображающий, как различные сущности (например, люди, объекты или концепции) связаны в рамках системы. ER-диаграммы чаще всего применяются для проектирования и анализа реляционных баз данных в сферах образования, научных исследований, разработки программного обеспечения и бизнес-информационных систем.

Основой ER-диаграмм является стандартный набор символов, таких как прямоугольники для сущностей, ромбы для связей, овалы для атрибутов и линии для соединения элементов. Эти диаграммы по структуре напоминают грамматические конструкции, где сущности играют роль существительных, а связи — глаголов.

ER-диаграммы тесно связаны с диаграммами структуры данных (DSD), которые фокусируются на внутренних элементах сущностей, а не на связях между ними. Кроме того, ER-диаграммы нередко используют вместе с диаграммами потоков данных (DFD), которые показывают движение информации внутри системы или процесса (рисунок 4).

Данная база данных является информационной основой создаваемого сервиса, который будет использоваться для автоматизации процесса составления расписания в университете с помощью мобильного приложения. Она позволяет рассматривать запросы на добавление, удаление и выборку информации. Добавить сведения о предметах и кафедрах, а также данные о должностях и преподавателях может структурное подразделение "Деканаты", добавляя информацию о дисциплинах и курсах - структурное подразделение «Отдел кадров». Благодаря использованию структуры и интерфейса базы данных появилась возможность перераспределить работу, связанную с документооборотом, между различными подразделениями. Данные, которые находятся на сервере, имеют высокую степень надежности, что способствует ускорению процесса документооборота в подразделениях и повышает качество их работы.

### **Заключение.**

Целью данного исследования являлась разработка и внедрение электронной системы документооборота в образовательных учреждениях для повышения эффективности управленческих процессов и содействия цифровой трансформации образовательной среды. В ходе работы был проведён комплексный анализ существующих решений, разработана архитектура электронного портала, выбраны подходящие инструменты разработки и реализован проект пользовательского интерфейса.

Научная новизна исследования заключается в разработке интегрированной электронной системы документооборота, ориентированной на специфику образовательных организаций, с включением элементов аналитики и мониторинга продуктивности сотрудников. Данный подход позволяет не только автоматизировать процессы обработки документов, но и предоставляет новые возможности анализа данных о производительности труда и управленческой эффективности, которые недостаточно полно реализованы в существующих аналогах.

Оценка экономической эффективности автоматизации документооборота должна учитывать множество факторов и анализировать большие объёмы информации. Корректный выбор и качественное внедрение системы способствуют сокращению временных затрат на рутинные операции, увеличивая производительность труда и объём выполняемых задач без дополнительной нагрузки на сотрудников.

Несмотря на очевидные преимущества электронного документооборота, точная количественная оценка его экономического эффекта затруднена из-за многочисленных переменных факторов. Однако основные преимущества включают ускорение обмена

информацией, автоматизацию процессов, повышение прозрачности и исполнительской дисциплины, улучшение контроля и устранение необходимости в дополнительных ресурсах для хранения архивных документов.

Будущие исследования могут быть ориентированы на разработку методик количественной оценки экономической эффективности внедрения электронных систем документооборота, а также на совершенствование аналитических инструментов мониторинга продуктивности сотрудников.

#### Список литературы

1. EMC – Enterprise Content Management. (2018). *Sistema upravleniya korporativnymi informacionnymi resursami EMC – Enterprise Content Management: upravlenie cifrovymi dokumentami i drugimi tipami kontenta, a takzhe ih hranenie, obrabotka i dostavka v ramkah organizacii*.

2. Selezneva, N. A. (2019). *Kachestvo vysshego obrazovaniya kak ob"ekt sistemnogo issledovaniya: lekciya-doklad* (6-e izd., stereotip.). Issledovatel'skij centr problem kachestva podgotovki specialistov.

3. Obshchaya klassifikaciya sistem dokumentooborota. (n.d.). *Sistemy elektronnoho dokumentooborota, problemy, perspektivy* [Elektronnyj resurs]. Retrieved from <http://www.eos.ru/>.

4. Semenov, S. P., & Tatarincev, Ya. B. (2019). Avtomatizirovannaya sistema kontrolya dannyh po kadrovomu sostavu vuza i kontingentu studentov. *Sistemy upravleniya i informacionnye tekhnologii, Nauchno-tekhnicheskij zhurnal*, 3.2(37), 279–282.

5. Smirnova, G. N. (2022). *Elektronnye sistemy upravleniya dokumentooborotom: uchebnoe posobie*. Moskovskij mezhdunarodnyj institut ekonometriki, informatiki, finansov i prava.

6. Bakunova, O. M., Anokhin, E. V., Poluiko, A. F., Aleksandrovich, E. N., Antonov, E. D., Sitnik, M. Yu., Grechko, I. S., & Kabakov, D. M. (2018). Primenenie elektronnoho dokumentooborota v programme 1С. *International Journal of Innovative Technologies in Economy*, 4(16), 64-66.

7. Anatskaya, A. G. (2019). *Zashchita elektronnoho dokumentooborota: uchebnoe posobie*. Omsk: SibADI.

8. Selezneva, N. A. (2020). *Kachestvo vysshego obrazovaniya kak ob"ekt sistemnogo issledovaniya: lekciya-doklad* (6-e izd., stereotip.). Issledovatel'skij centr problem kachestva podgotovki specialistov.

#### БІЛІМ БЕРУДЕГІ ҚҰЖАТ БАСҚАРУ ЖҮЙЕСІН ДАМЫТУ

**Аңдатпа.** Бұл мақалада білім беру ұйымдарындағы дәстүрлі құжат айналымынан электронды құжат айналымына көшу процесі қарастырылған. Білім беру ұйымының автоматтандырылған ақпараттық жүйесіндегі функциялар мен процестерге талдау жүргізілді. Электрондық құжат айналымын енгізудің негізгі тәсілдері жүйелік көзқарас тұрғысынан қарастырылады. Практикалық тәжірибеге сүйене отырып, авторлар мамандардың біліктілігін арттыру жүйесінде оқытуды ұйымдастыру және өткізу тұжырымдамасын ұсынады.

Білім беру ұйымдарындағы құжат айналымын автоматтандыру кіріс және шығыс хат-хабарларды есепке алу, ішкі құжаттарды өңдеу, әкімшілік және шарттық құжаттарды келісу, бұйрықтардың орындалуын бақылау сияқты процестерде ерекше өзекті болып табылады. Электрондық құжат айналымы жүйелерін (ЭҚҚЖ) енгізу анықтамалықтарды бірыңғай деректер қоймасына біріктіруді және құжаттарды өңдеу процестерін мекеменің басқа бизнес-процестерімен синхрондауды талап етеді.

**Түйін сөздер:** іс жүргізу, электрондық құжат айналымы, білім беру ұйымын басқару, Мемлекеттік басқарудың тиімділігі, ақпараттық жүйелер, ақпараттық технологиялар, электрондық басқару жүйелері.

## DEVELOPMENT OF A DOCUMENT MANAGEMENT SYSTEM IN EDUCATION

**Abstract.** This article considers the process of transition from traditional to electronic document flow in educational institutions. The analysis of functions and processes within the framework of the automated information system of an educational organization is carried out. Key approaches to the implementation of electronic document management from the point of view of the system approach are considered. Based on practical experience, the authors propose the concept of organizing and conducting training in the system of professional development of specialists.

Automation of document management in educational institutions is especially relevant in such processes as accounting of incoming and outgoing correspondence, processing of internal documents, coordination of administrative and contractual documents, as well as control of execution of orders. Implementation of electronic document management systems (EDMS) requires integration of directories into a single data repository and synchronization of document processing with other business processes of the institution.

**Keywords:** office management, electronic document management, management of educational organization, efficiency of public administration, information systems, information technology, electronic management systems.

### Авторлар туралы мәлімет

Айтолқын Саметова	Магистр, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан, aberikkaliyeva@gmail.com
Рашидинов Дамир Рашидинович	PhD кандидаты, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан, damir.dmr88@gmail.com
Найзабаева Ляззат Кыдыргалиевна	Актерлік өнер Профессор техника ғылымдарының докторы, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан, l.naizabayeva@edu.iitu.kz
Сулейменова Мадина Утегеновна	Магистр технических наук, ассистент профессора каф ИС, madekin940@gmail.com

### Сведение об авторах

Айтолқын Саметова	магистр, Международный университет информационных технологий, г. Алматы, Казахстан, aberikkaliyeva@gmail.com
Рашидинов Дамир Рашидинович	докторант PhD, Международный университет информационных технологий, г. Алматы, Казахстан, damir.dmr88@gmail.com
Найзабаева Ляззат Кыдыргалиевна	и.о. профессора Доктор технических наук, Международный университет информационных технологий, г. Алматы, Казахстан, l.naizabayeva@edu.iitu.kz
Сулейменова Мадина Утегеновна	Техника ғылымдарының магистрі, ақпараттық жүйелер кафедрасының ассистенті, madekin940@gmail.com

### Information about the authors

Aitolkyn Sametova	Master, International University of Information Technology, Almaty, Kazakhstan, aberikkaliyeva@gmail.com
-------------------	--

Rashidinov Damir Rashidinovich	PhD candidate, International University of Information Technology, Almaty, Kazakhstan, damir.dmr88@gmail.com
Nayzabaeva Lyazzat Kydyrgalieвна	Acting Professor Doctor of Technical Sciences, International University of Information Technology, Almaty, Kazakhstan, l.naizabayeva@edu.iitu.kz
Suleimenova Madina Utegenovna	Master of Engineering Sciences, Assistant Professor of the Department of Information Systems, madekin940@gmail.com

UDC 629.7.08

IRSTI 81.75.29

[https://doi.org/10.53364/24138614\\_2025\\_36\\_1\\_16](https://doi.org/10.53364/24138614_2025_36_1_16)A. Koshekov<sup>1\*</sup>, D. Toleubekov<sup>1</sup>, I. Nautan<sup>1</sup>, Y. Kurbanov<sup>1</sup>, R. Togambayev<sup>2</sup><sup>1</sup>Civil Aviation Academy, Almaty, Kazakhstan<sup>2</sup>Transport and Telecommunication Institute, Riga, LatviaE-mail: [a.k.koshekov@gmail.com](mailto:a.k.koshekov@gmail.com)\*

## INNOVATIVE TRAINING COMPLEX WITH ENGINEERING SUPPORT: A NEW LEVEL OF TECHNICAL EXPLOITATION OF AVIATION

**Abstract.** This research explores the creation of a novel training complex that includes an engineering support system specifically designed for the technical management of military and specialized aviation equipment. The importance of this study stems from the need to improve the training efficiency of engineering and technical staff, given the growing technical complexity of aviation equipment and the increasing demands for aviation safety. The main aim of this research is to develop an adaptive educational platform that integrates digital twins of aviation systems, fault diagnosis algorithms, and intelligent methods to tailor the educational process. Within this study, the architecture of a comprehensive engineering support system was devised, which includes data collection, digital modeling, and analytics. An adaptive learning algorithm has been introduced, which takes into account the trainee's skill level, experience, and progress in mastering the material, thereby automatically adjusting the curriculum. The use of virtual simulators and simulation models enabled the development of a flexible training system that closely mirrors the actual operating conditions of aviation equipment. To evaluate the proposed system's effectiveness, experimental testing was carried out, comparing the training results of two groups of engineering and technical personnel: the experimental group (using adaptive training with digital models) and the control group (using traditional training methods). The findings showed that the new training complex reduced response time to emergencies by 22%, lowered the number of errors by 30%, and improved the accuracy of procedure execution by 18%. The adaptive training system developed is highly flexible, allowing for customization and integration with modern predictive diagnostics methods for aircraft malfunctions. Implementing such a complex in the training of specialists will enhance aviation safety, reliability, and the economic efficiency of technical aircraft operations.

**Keywords:** aviation equipment, engineering support, digital twins, adaptive learning, training complex, fault diagnostics.

### Introduction.

Modern aviation, both military and civil, is a complex technical system that requires strict compliance with the rules and standards of operation, maintenance, and repair developed by international organizations in aviation security. The main principle in aviation is to reduce the time aircraft spend on the ground [1]; this factor directly depends on the level of training of service engineering personnel and ensuring high-quality technical operations. In connection with the above, the development of training complexes with integrated engineering support systems is an urgent task to increase the reliability and combat readiness of aviation equipment.

To ensure the reliability of intelligent simulators, imitation of the real operating conditions of aircrafts is used by implementing modelling technologies, digital twins, and artificial

intelligence. The introduction of such a simulator will improve the efficiency of training engineering and technical personnel in the aviation industry through the use of multiscenario modes, which will allow the simulation of all possible circumstances that may arise during operation and repair.

Today, simulators for training flight personnel have become widespread in the aviation industry, having received good reviews from both personnel and regulatory authorities in the aviation sector, and the potential of simulator complexes for engineering and technical personnel has not been fully realized [2]. Simultaneously, ensuring a high level of technical maintenance will improve aviation safety and ensure the long-term operation of the equipment. Therefore, the use of such a simulator complex with engineering support tools can significantly increase the effectiveness of training in the educational environment and provide an expansion of training scenarios for retraining and advanced training of engineering personnel without significant financial investment from stakeholders.

The scientific novelty of this study lies in the development of a training complex that combines realistic modelling of the processes of maintenance and operation of aircraft equipment, intelligent fault diagnostic systems, and automated training tools for engineering and technical personnel. If we compare the training complex proposed in this study with those that currently exist, we can highlight the following competitive advantages: the presence of a digital platform that provides access to regulatory documentation, analytical data on the progress of trainees, and the availability of fault-finding algorithms.

The purpose of this study is to develop a training complex with an engineering support system for the technical operation of aircraft for various applications. To achieve this goal, researchers have set the following number of tasks:

1. Develop an integrated engineering support system architecture that includes digital models of aircraft and fault-diagnostic algorithms.
2. Develop an algorithm for the adaptive training of engineering and technical personnel considering aircraft specifications
3. Assess the effectiveness of the developed system and traditional methods for training aviation specialists.

According to the authors [3], one of the promising areas for the development of such systems is the use of digital twin technology for aviation equipment, which allows for the simulation of operational processes and modelling of all types of malfunctions. However, modern literature focuses on civil aviation, and the adaptation of such systems for military and special equipment requires significant work on the modification and consideration of the specifics of operation of such equipment [4].

Another direction that has been actively developed in recent years is the use of automated maintenance and repair management systems integrated with intelligent algorithms to predict unit failures [5] and automatic diagnostic processes [6] of aviation equipment. The use of such integrated methods increases the efficiency of fault detection and generates automatic recommendations for their elimination.

Another problematic issue is the insufficient material and technical base of educational institutions in the field of aviation specialist training, which leads to a lack of practical experience among graduates [7]. The introduction of such simulator complexes into the training process will improve the level of training of engineering and technical personnel and reduce the likelihood of errors under real working conditions [8].

Therefore, the development of a simulator complex with an engineering support system is associated with challenges in the operation and maintenance of aviation equipment [9]. The introduction of such a complex process of training engineering and technical specialists will ensure an increase in the level of aviation safety, reliability, and economic efficiency of aviation

equipment, which in turn will have a positive effect on the combat capability and operational readiness of aircraft [10].

### Materials and methods.

To achieve the project objectives, researchers have developed an integrated system architecture. Figure 1 illustrates the architecture of the integrated system.

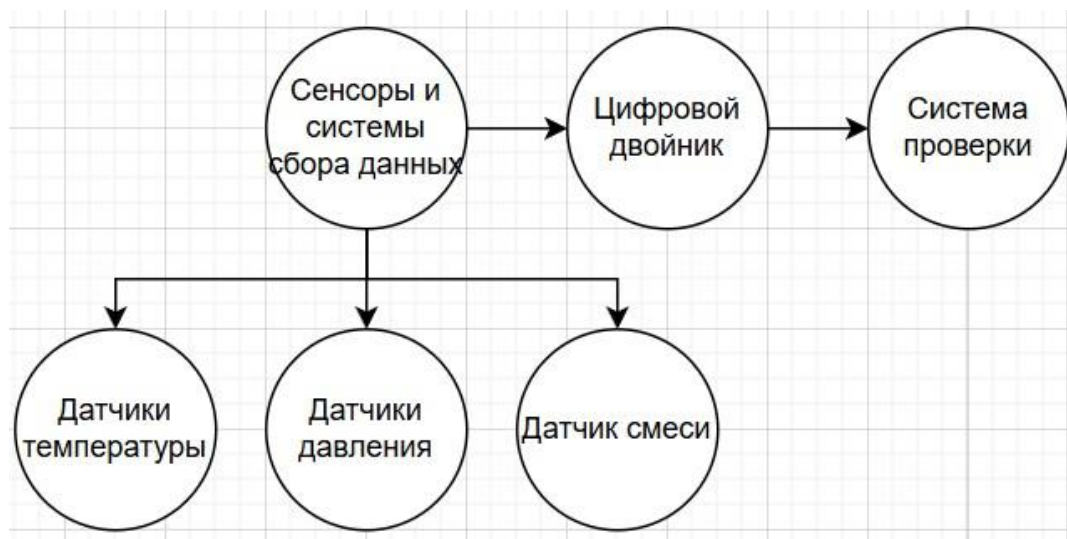


Figure 1 – Structural diagram of the integrated system architecture

The architecture shown in Figure 1 includes a data collection level (sensors and sensors) and temperature, pressure, and mixture sensors. The next level is the digital twin level, which is a simulation model of aviation equipment that includes an object (aircraft) and a tool (deicing machine). The last is the analytical level, which is necessary to assess the correctness of a trainee's actions. The anti-icing preparation of an aircraft is one of the important stages in preparation for departures, because upon reaching certain altitudes, the aircraft encounters extremely low temperatures, down to -50 -60 degrees. To improve the efficiency of training engineering and technical personnel, the use of the adaptive learning algorithm shown in Figure 2 was proposed.



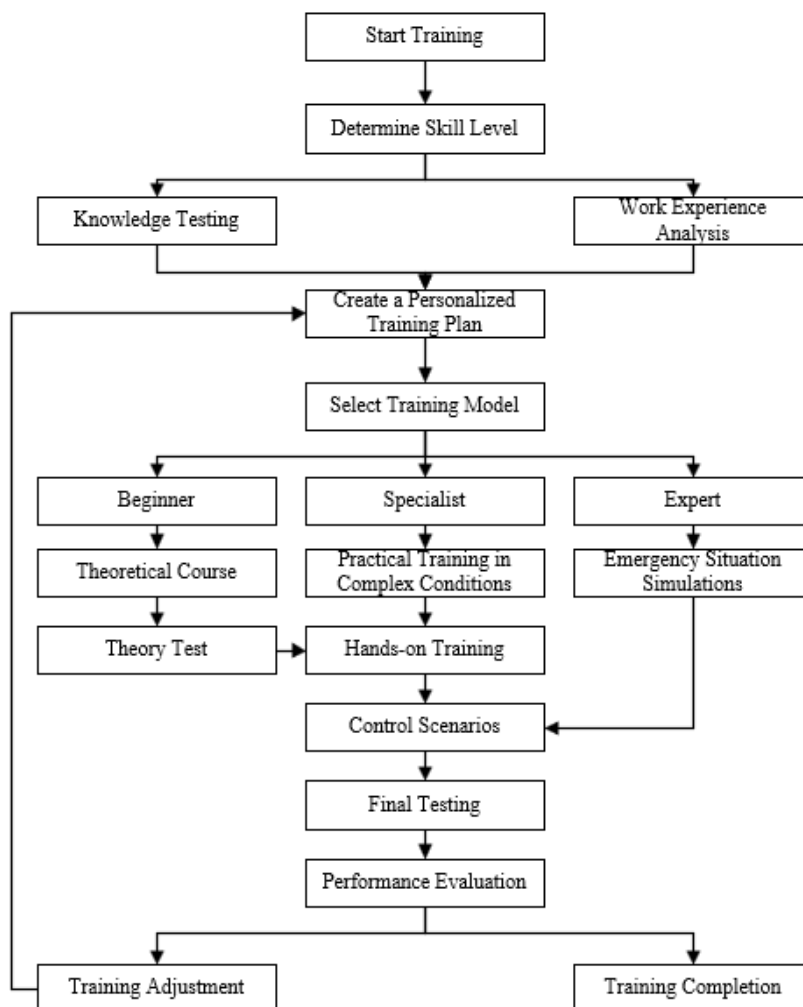


Figure 2 – Block diagram of the adaptive learning algorithm

The block diagram of the adaptive learning algorithm shown in the figure includes five main stages:

Stage 1. Determine the trainee's training level conducting an entrance test to assess the theoretical knowledge of anti-icing treatment, including regulations, standards, and methods of treatment, as well as practical skills (experience with equipment, reagents, and heating systems). Analysis of the presence of experience in this type of service (beginner: up to 1 year, specialist: from 1 to 5 years, expert: from 5 years). The system created a personal training plan based on the initial data.

Stage 2. Selection of the training model. Based on the results obtained from testing the trainee, the system selected an appropriate training model. The developed system includes three training models: beginner (studying the theoretical course and consolidating the material in basic practical exercises), specialists (practical exercises in more difficult weather conditions), and experts (simulation of emergency situations and training of non-standard cases).

Stage 3. Theoretical training. Access of trainees to electronic training materials developed by international organisations in the field of aviation security, study of interactive cases (incident analysis and study of the most common errors), and testing on the studied material.

Stage 4. Practical training with virtual modelling elements: Virtual simulator with integrated digital twins of aircraft (practising basic algorithms for anti-icing treatment and simulating various weather conditions). Testing professional skills in the control scenarios.

Stage 5. Evaluation of the effectiveness and adaptation of the training. Final testing. Evaluation of errors and analysis of error data. Correction of the training plan (study of additional modules and in-depth training on the simulator).

To evaluate the effectiveness of the developed simulator complex and traditional training methods, a comparative analysis was used within the framework of the project, within which a sample consisting of two groups of 14 people was formed, in accordance with the requirements for training specialists. These groups underwent an entrance test to determine their levels of knowledge. After completing the theoretical course, an exit test was conducted to assess knowledge, a practical exam with modelling of various anti-icing training situations. The Student's t-test was used to compare the results between the two groups.

### Results.

For the developed adaptive learning algorithm, software was developed that includes an interface written on the PyQt technology stack for graphical representation. Adaptive learning logic has been implemented, including determining the level of knowledge of the student by passing a theoretical test, after which a personal learning plan is generated, monitoring the student's progress, and adjusting the learning plan. The software used is shown in Figure 3.

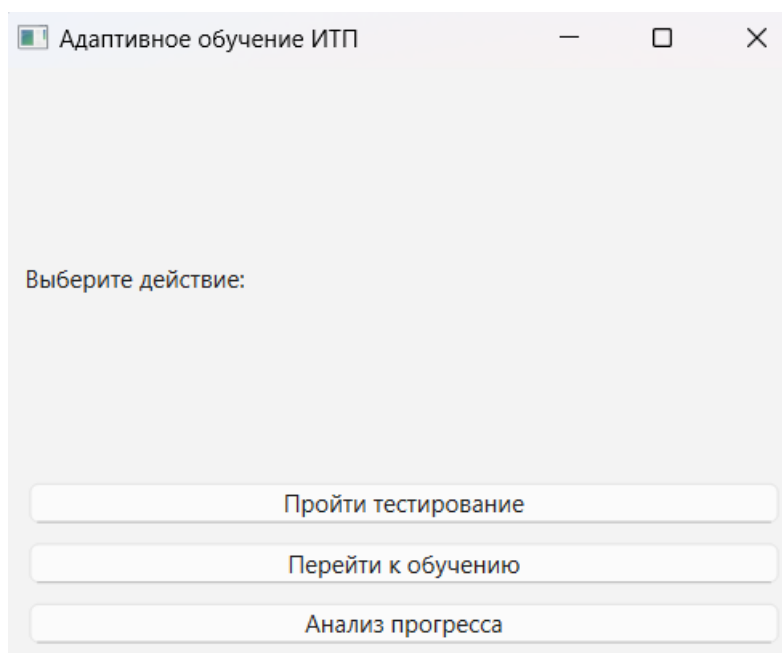


Figure 3 – Appearance of the adaptive learning program

The software shown in the figure includes a front-end interface with three buttons. When pressed, or student goes to a specific subprogram, such as a test base, training materials, and an analytical block. After completing the theoretical part, the student goes to the practical part; the external program for completing the practical part is shown in Figure 4.



engineering and technical personnel. The inclusion of digital technologies in the process of training specialists will ensure an increase in the level of aviation personnel training, which, in turn, directly affects the safety and reliability of the operation of aviation equipment. Further developments in this project are related to the implementation of augmented reality technologies and machine-learning algorithms.

### Financing

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19680080). «Development of training complex with a system of engineering support for the technical operation of military and special aviation transport equipment».

### References

1. Wu, C., & Caves, R. E. (2004). Modelling and optimization of aircraft turnaround time at an airport. *Transportation Planning and Technology*, 27(1), 47–66. <https://doi.org/10.1080/0308106042000184454>
2. Ross, G., & Gilbey, A. (2023). Extended reality (xR) flight simulators as an adjunct to traditional flight training methods: a scoping review. *CEAS Aeronautical Journal*, 14(4), 799–815. <https://doi.org/10.1007/s13272-023-00688-5>
3. Li, X., Niu, W., & Tian, H. (2024). Application of Digital Twin in Electric Vehicle Powertrain: A Review. *World Electric Vehicle Journal*, 15(5), 208. <https://doi.org/10.3390/wevj15050208>
4. Li, L., Wileman, A., Perinpanayagam, S., & Aslam, S. (2022). Digital Twin in Aerospace Industry: A Gentle Introduction. *IEEE Access*, 10, 9543–9562. <https://doi.org/10.1109/access.2021.3136458>
5. Bello, I., Abdullahi, U. A., Khan, A., Abdulhamid, S. M., Okesola, J. O., Chiroma, H., Gital, A. Y., & Jauro, F. (2020). Detecting ransomware attacks using intelligent algorithms: recent development and next direction from deep learning and big data perspectives. *Journal of Ambient Intelligence and Humanized Computing*, 12(9), 8699–8717. <https://doi.org/10.1007/s12652-020-02630-7>
6. Ali, R., & Cui, H. (2024). Unleashing the potential of AI in modern healthcare: Machine learning algorithms and intelligent medical robots. *Research on Intelligent Manufacturing and Assembly*, 3(1), 100–108. <https://doi.org/10.25082/rima.2024.01.002>
7. Phan, T. A., Vu, T. H. N., Le, T.-H., & Vo, N. T. N. (2024). Enhancing Educational Outcomes Through Strategic Guest Speaker Selection: A Comparative Study of Alumni and Industry Experts in University Settings. *Business and Professional Communication Quarterly*. <https://doi.org/10.1177/23294906241263035>
8. Alonso, F., Riera, J. V., Useche, S. A., Fernandez-Marin, M., & Faus, M. (2023). Effectiveness of Driving Simulators for Drivers' Training: A Systematic Review. *Applied Sciences*, 13(9), 5266. <https://doi.org/10.3390/app13095266>
9. Vianna, W. O. L., & Yoneyama, T. (2018). Predictive Maintenance Optimization for Aircraft Redundant Systems Subjected to Multiple Wear Profiles. *IEEE Systems Journal*, 12(2), 1170–1181. <https://doi.org/10.1109/jsyst.2017.2667232>
10. Alomar, I., & Yatskiv, I. (2023). DIGITALIZATION IN AIRCRAFT MAINTENANCE PROCESSES. *Aviation*, 27(2), 86–94. <https://doi.org/10.3846/aviation.2023.18923>

## ИННОВАЦИОННЫЙ УЧЕБНО-ТРЕНИРОВОЧНЫЙ КОМПЛЕКС С ИНЖЕНЕРНЫМ ОБЕСПЕЧЕНИЕМ: НОВЫЙ УРОВЕНЬ ТЕХНИЧЕСКОЙ ЭКСПЛУАТАЦИИ АВИАЦИИ

***Аннотация.** Это исследование посвящено созданию нового учебного комплекса, который включает в себя систему инженерного обеспечения, специально разработанную для технического управления военной и специализированной авиационной техникой. Важность данного исследования обусловлена необходимостью повышения эффективности подготовки инженерно-технического персонала, учитывая растущую техническую сложность авиационной техники и растущие требования к авиационной безопасности. Основной целью данного исследования является разработка адаптивной образовательной платформы, которая объединяет цифровые аналоги авиационных систем, алгоритмы диагностики неисправностей и интеллектуальные методы для адаптации учебного процесса. В рамках этого исследования была разработана архитектура комплексной системы инженерной поддержки, которая включает сбор данных, цифровое моделирование и аналитику. Был внедрен адаптивный алгоритм обучения, который учитывает уровень квалификации обучаемого, его опыт и прогресс в освоении материала, тем самым автоматически корректируя учебную программу. Использование виртуальных тренажеров и имитационных моделей позволило разработать гибкую систему обучения, которая точно отражает реальные условия эксплуатации авиационной техники. Для оценки эффективности предложенной системы было проведено экспериментальное тестирование, в ходе которого сравнивались результаты обучения двух групп инженерно-технического персонала: экспериментальной группы (с использованием адаптивного обучения с использованием цифровых моделей) и контрольной группы (с использованием традиционных методов обучения). Результаты показали, что новый учебный комплекс сократил время реагирования на чрезвычайные ситуации на 22%, снизил количество ошибок на 30% и повысил точность выполнения процедур на 18%. Разработанная адаптивная система обучения отличается высокой гибкостью и позволяет адаптировать ее к современным методам прогнозной диагностики неисправностей воздушных судов. Внедрение такого комплекса в подготовку специалистов повысит авиационную безопасность, надежность и экономическую эффективность технической эксплуатации воздушных судов.*

***Ключевые слова:** авиационное оборудование, инженерное обеспечение, цифровые двойники, адаптивное обучение, тренировочный комплекс, диагностика неисправностей.*

## ИНЖЕНЕРЛІК ҚАМТАМАСЫЗ ЕТІЛГЕН ИННОВАЦИЯЛЫҚ ОҚУ КЕШЕНІ: АВИАЦИЯНЫ ТЕХНИКАЛЫҚ ПАЙДАЛАНУДЫҢ ЖАҢА ДЕҢГЕЙІ

***Андатпа.** Бұл зерттеу әскери және мамандандырылған авиациялық техниканы техникалық басқару үшін арнайы әзірленген инженерлік қолдау жүйесін қамтитын жаңа оқу кешенін құруды зерттейді. Бұл зерттеудің маңыздылығы авиациялық техниканың техникалық күрделілігінің артуын және авиациялық қауіпсіздікке қойылатын талаптардың артуын ескере отырып, инженерлік-техникалық кадрларды даярлаудың тиімділігін арттыру қажеттілігінен туындайды. Бұл зерттеудің негізгі мақсаты-авиациялық жүйелердің цифрлық егіздерін, ақауларды диагностикалау алгоритмдерін және оқу процесін бейімдеудің интеллектуалды әдістерін біріктіретін адаптивті білім беру платформасын әзірлеу. Осы зерттеу аясында деректерді жинауды, цифрлық модельдеуді және аналитиканы қамтитын кешенді инженерлік қолдау жүйесінің архитектурасы әзірленді. Тыңдаушының біліктілік деңгейін, тәжірибесін және материалды меңгерудегі үлгерімін ескеретін, сол арқылы оқу жоспарын автоматты*

түрде реттейтін бейімделетін оқыту алгоритмі енгізілді. Виртуалды тренажерлар мен модельдеу модельдерін қолдану авиациялық техниканың нақты жұмыс жағдайларын дәл көрсететін икемді оқыту жүйесін жасауға мүмкіндік берді. Ұсынылған жүйенің тиімділігін бағалау үшін инженерлік-техникалық персоналдың екі тобының оқу нәтижелерін салыстыра отырып, эксперименттік тестілеу жүргізілді: эксперименттік топ (цифрлық модельдермен адаптивті оқытуды қолдана отырып) және бақылау тобы (дәстүрлі оқыту әдістерін қолдана отырып). Нәтижелер көрсеткендей, жаңа оқу кешені төтенше жағдайларды жою уақытын 22% - ға қысқартты, қателер санын 30% - ға азайтты және процедураларды орындау дәлдігін 18% - ға жақсартты. Әзірленген адаптивті оқыту жүйесі жоғары икемділікке ие, бұл әуе кемелерінің ақауларын диагностикалаудың заманауи болжамды әдістерімен теңшеуге және біріктіруге мүмкіндік береді. Мамандарды даярлауда мұндай кешенді енгізу авиациялық қауіпсіздікті, сенімділікті және әуе кемелерінің техникалық операцияларының экономикалық тиімділігін арттырады.

**Түйінді сөздер:** авиациялық техника, инженерлік қамтамасыз ету, цифрлық егіздер, адаптивті оқыту, оқу кешені, ақауларды диагностикалау.

#### Авторлар туралы мәлімет

Кошекков Кайратович	Абай	PhD, Азаматтық Авиация Академиясы «Авиациялық техника және технологиялар» қауымдастырылған профессоры, Алматы қ., Қазақстан E-mail: a.k.koshekov@gmail.com
Толубеков Дауренбекұлы	Даулетбек	Азаматтық Авиация Академиясы «Авиациялық техника және технологиялар» магистрант, Алматы қ., Қазақстан E-mail: d.tolubekov@agakaz.kz
Наутан Қуандыққызы	Ибтисэма	Азаматтық Авиация Академиясы «Авиациялық техника және технологиялар» магистрант, Алматы қ., Қазақстан E-mail: i.nautan@agakaz.kz
Курбанов Мухсатұлы	Якуб	Азаматтық Авиация Академиясы «Авиациялық техника және технологиялар» докторант, Алматы қ., Қазақстан E-mail: y.kurbanov@agakaz.kz
Тогамбаев Кумарбекович	Рустам	Көлік және коммуникация институты докторант, Рига қ., Латвия E-mail: r.togambayev@tsi.lv

#### Сведения об авторах

Кошекков Кайратович	Абай	PhD, ассоциированный профессор кафедры «Авиационная техника и технологии, Академия гражданской авиации, г. Алматы, Казахстан E-mail: a.k.koshekov@gmail.com
Толубеков Дауренбекұлы	Даулетбек	Магистрант, Академия гражданской авиации, г. Алматы, Казахстан E-mail: d.tolubekov@agakaz.kz
Наутан Қуандыққызы	Ибтисэма	Магистрант, Академия гражданской авиации, г. Алматы, Казахстан E-mail: i.nautan@agakaz.kz
Курбанов Мухсатұлы	Якуб	Докторант, Академия гражданской авиации, г. Алматы, Казахстан E-mail: y.kurbanov@agakaz.kz
Тогамбаев Кумарбекович	Рустам	Докторант, Институт транспорта и связи, Рига, Латвия E-mail: r.togambayev@tsi.lv

#### Information about the authors

Abay Koshekov	PhD, Associate Professor, Department of Aviation Engineering and Technology, Civil Aviation Academy, Almaty, Kazakhstan E-mail: a.k.koshekov@gmail.com
---------------	---

Dauletbek Toleubekov	Postgraduate student, Civil Aviation Academy, Almaty, Kazakhstan E-mail: d.toleubekov@agakaz.kz
Ibtisema Nautan	Postgraduate student, Civil Aviation Academy, Almaty, Kazakhstan E-mail: i.nautan@agakaz.kz
Yakub Kurbanov	Doctoral student, Civil Aviation Academy, Almaty, Kazakhstan E-mail: y.kurbanov@agakaz.kz
Rustam Togambayev	Doctoral student, Civil Aviation Academy, Almaty, Kazakhstan E-mail: r.togambayev@tsi.lv

**INFORMATION ABOUT THE AUTHORS**

1. **Sagdiyev Tulkun** – PhD, Tashkent State Transport University, Faculty of Aviation Transport Engineering, Tashkent, Uzbekistan;
2. **Maturazov Izzat** – Tashkent State Transport University, Faculty of Aviation Transport Engineering, PhD, Tashkent, Uzbekistan, E-mail: [maturazov\\_i@tstu.uz](mailto:maturazov_i@tstu.uz);
3. **Isakov Niyazbek** – Tashkent State Transport University, Faculty of Aviation Transport Engineering, Master's student, Tashkent, Uzbekistan, E-mail: [niyazbekalievich003@gmail.com](mailto:niyazbekalievich003@gmail.com);
4. **Nagimov Almas** – fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan, E-mail: [anm24012004@gmail.com](mailto:anm24012004@gmail.com);
5. **Beketova Gulzhanat** – PhD, Associate Professor of the "IT Engineering and Artificial Intelligence" Department at Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, E-mail: [Beketova2111@gmail.com](mailto:Beketova2111@gmail.com);
6. **Akhmet Arys** – fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan E-mail: [arysakhmet@gmail.com](mailto:arysakhmet@gmail.com);
7. **Gusmanov Altai** – fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan E-mail: [altair.gusmanov@mail.ru](mailto:altair.gusmanov@mail.ru);
8. **Akramkhanov Alikhan** – fourth-year student in the information systems specialty of the International University of Information Technology, Almaty, Kazakhstan E-mail: [aakramkhanov@bk.ru](mailto:aakramkhanov@bk.ru);
9. **Akimbay Shyryn** – 2nd year PhD student of Automation and IoT specialty at Al-Farabi Kazakh National University, Almaty, Kazakhstan E-mail: [akimbay.sh@gmail.com](mailto:akimbay.sh@gmail.com);
10. **Abdullayev Khagani Imran** – Phd, professor of the National Aviation Academy, Baku, The Republic of Azerbaijan, E-mail: [khagani61@gmail.com](mailto:khagani61@gmail.com)
11. **Allakhverdiyeva Solmaz Qulu** – Senior lecturer, National Aviation Academy, Baku, The Republic of Azerbaijan, E-mail: [sallahverdiyeva@naa.edu.az](mailto:sallahverdiyeva@naa.edu.az);
12. **Abzhapbarova Ainur** – Candidate of Technical Sciences, Professor, Academy of Civil Aviation, Almaty, Republic of Kazakhstan, [Ainur.abzhapbarova@mail.ru](mailto:Ainur.abzhapbarova@mail.ru);
13. **Garmash Olga** – Candidate of Technical Sciences, Associate Professor AGA, Academy of Civil Aviation, Almaty, Republic of Kazakhstan, e-mail: [olm78@mail.ru](mailto:olm78@mail.ru);
14. **Makogonova Viktoriya** – Mass.ter, methodologist, Academy of Civil Aviation, Almaty, Kazakhstan, E-t.mail: [cherry-98-23-04@mail.ru](mailto:cherry-98-23-04@mail.ru);
15. **Assilbekova Indira** – C.T.S., Professor, Academy of Civil Aviation, Almaty, Kazakhstan, E-mail: [a.indira@mail.ru](mailto:a.indira@mail.ru);
16. **Nazym Alimbekova** – Department of Computer and Software Engineering, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: [nazakhatovna@gmail.com](mailto:nazakhatovna@gmail.com);
17. **Hari Mohan Rai** – PhD, Associate Professor, School of Computing, Gachon University, Seong-nam-si, South Korea, E-mail: [drhmrai@gachon.ac.kr](mailto:drhmrai@gachon.ac.kr);
18. **Tursinbay Turymbetov** – Acting Associate Professor, PhD in Engineering, Humanities school, International University of Tourism and Hospitality, Turkistan, Kazakhstan; E-mail: [t.turimbetov@iuth.edu.kz](mailto:t.turimbetov@iuth.edu.kz);
19. **Ainur Zhumadillayeva** – Associate Professor, PhD in Engineering, Department of Computer and Software Engineering, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: [zhumadillayeva\\_ak@enu.kz](mailto:zhumadillayeva_ak@enu.kz);
20. **Kumarkanova Akbota** – Master, Lecturer, School of Digital Technology and Artificial Intelligence, NCJSC «D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan, Phone: 87474371007; E-mail: [akbota.vko@gmail.com](mailto:akbota.vko@gmail.com);



21. **Khasenova Zarina** – PhD, Dean, School of Digital Technology and Artificial Intelligence, NCJSC «D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan, Phone: 87715492087; E-mail: [zhasenova@mail.ru](mailto:zhasenova@mail.ru);
22. **Weiss Yuri** – Candidate of Technical Sciences, Associate Professor, School of Digital Technology and Artificial Intelligence, NCJSC «D. Serikbayev East Kazakhstan Technical University», Ust-Kamenogorsk, Kazakhstan, Phone: 87052502872; E-mail: [yuvais@edu.ektu.kz](mailto:yuvais@edu.ektu.kz);
23. **Koyschieva Dina** – Master's student in the specialty M094-Information Technology, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, E-mail: [d.koishiyeva@aes.kz](mailto:d.koishiyeva@aes.kz);
24. **Sydybaeva Madina** – Senior Lecturer, Department of Information Systems and Cybersecurity, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, E-mail: [m.sydybaeva@alt.edu.kz](mailto:m.sydybaeva@alt.edu.kz);
25. **Belginova Saule** – PhD, Associate Professor of Turan University, Almaty, Kazakhstan, E-mail: [sbelginova@gmail.com](mailto:sbelginova@gmail.com);
26. **Zhaksybaev Aidyn** – Master's degree in Technological Machines and Equipment, Shakarim University, Head of «OlzhaPROEKT» LLP, Ust-Kamenogorsk, Kazakhstan, E-mail: [olzha2010@mail.ru](mailto:olzha2010@mail.ru);
27. **Ersainova Zhansaya** – Teacher of " SITaIS ", D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: [Zhansaya.ersayynova@mail.ru](mailto:Zhansaya.ersayynova@mail.ru);
28. **Mekhdiev Turaj** – 2st year master's degree; specialty of «Information security», Eurasian National University named after L.N. Gumilyov, The Republic of Kazakhstan, ORCID: <https://orcid.org/0009-0004-6771-1584>, E-mail: [mehdiev.t@gmail.com](mailto:mehdiev.t@gmail.com);
29. **Shaikhanova Aigul** – PhD, Professor, Department of Information Security, L.N. Gumilyov Eurasian National University; e-mail: [shaikhanova\\_ak@enu.kz](mailto:shaikhanova_ak@enu.kz). ORCID: <https://orcid.org/0000-0001-6006-4813>, E-mail: [shaikhanova\\_ak@enu.kz](mailto:shaikhanova_ak@enu.kz);
30. **Bekeshova Gulvira** – Senior Lecturer at the Department of Information Security, L.N. Gumilyov Eurasian National University, ORCID: <https://orcid.org/0000-0002-1635-4693>, E-mail: [bekeshova\\_gb@enu.kz](mailto:bekeshova_gb@enu.kz);
31. **Iklasova Kajnizhamal** - PhD, Acting Professor of the Department of Information and Communication Technologies, Manash Kozybayev North Kazakhstan University, e-mail: [keiklasova@ku.edu.kz](mailto:keiklasova@ku.edu.kz), ORCID: <https://orcid.org/0000-0002-8330-4282>;
32. **Bakenova Kamila** - 2nd year doctoral student; specialty of «Information security», Eurasian National University named after L.N. Gumilyov, The Republic of Kazakhstan, ORCID: <https://orcid.org/0009-0004-2567-173X>, E-mail: [bakenova\\_ks@enu.kz](mailto:bakenova_ks@enu.kz);
33. **Nurgaliyeva Symbat** – Assistant professor of department of computer engineering, PhD, Astana IT University, Astana, Kazakhstan, E-mail: [symbat.nurgaliyeva@astanait.edu.kz](mailto:symbat.nurgaliyeva@astanait.edu.kz);
34. **Naiman Nurbek** – Master student of department of computer engineering, Astana IT university, Astana, Kazakhstan, E-mail: [242756@astanait.edu.kz](mailto:242756@astanait.edu.kz);
35. **Adikova Saltanat** - Associate Professor of the Department of Computer Modeling and Information Technology, PhD, Sarsen Amanzholov East Kazakhstan university, Ust-Kamenogorsk, Kazakhstan, E-mail: [ersal\\_7882@mail.ru](mailto:ersal_7882@mail.ru);
36. **Tursynkhan Tursunov** – Master's student, Department of Computer Engineering in Astana IT University, Astana, Kazakhstan, E-mail: [tursynkhan001@gmail.com](mailto:tursynkhan001@gmail.com);
37. **Dinara Kaibassova** – PhD, Associate Professor, Department of Computer Engineering in Astana IT University, Astana, Kazakhstan, E-mail: [Dinara.Kaibasova@astanait.edu.kz](mailto:Dinara.Kaibasova@astanait.edu.kz);
38. **Karymsakova Indira** – PhD, acting associate professor of the department of "IT technologies", Non-profit Joint-stock company "Shakarim University of Semey", Semey, Republic of Kazakhstan, e-mail: [karymsakova.indira@mail.ru](mailto:karymsakova.indira@mail.ru);

39. **Bekenova Dariga** – Master of mathematics, senior lecturer at the department of Information Technologies, “Turan-Astana” University, Astana, Republic of Kazakhstan, e-mail: [dariy6a@mail.ru](mailto:dariy6a@mail.ru);
40. **Kozhakhmetova Dinara** – PhD, Dean of the Higher School of Artificial Intelligence and Construction, Non-profit Joint-stock company "Shakarim University of Semey", Semey, Republic of Kazakhstan, e-mail: [dinara\\_kozhahmetova@mail.ru](mailto:dinara_kozhahmetova@mail.ru);
41. **Karmenova Marhaba** – PhD, associate professor of department "Computer modeling and information technology", Non-profit Joint-stock company “S. Amanzholov East Kazakhstan University”, Ust-Kamenogorsk, Republic of Kazakhstan, e-mail: [mmm058246@gmail.com](mailto:mmm058246@gmail.com);
42. **Ustinova Tatyana** – Master, senior lecturer of the department of "IT technologies", Non-profit Joint-stock company "Shakarim University of Semey", Semey, Republic of Kazakhstan, e-mail: [ustinova-t-a@yandex.ru](mailto:ustinova-t-a@yandex.ru);
43. **Batalova Madina** – doctoral student of "SDTaAI", D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: [esimkhan\\_kizi.m@mail.ru](mailto:esimkhan_kizi.m@mail.ru);
44. **Pawel Komada** – PhD, professor of Electronics and Information Technology at Faculty of Electrical Engineering and Computer Science of Lublin University of Technology, Poland, E-mail: [p.komada@pollub.pl](mailto:p.komada@pollub.pl);
45. **Alibekkyzy Karlygash** – Senior lecturer of "SDTaAI", Doctor of Philosophy PhD of D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: [karlygash.alibekkyzy@mail.ru](mailto:karlygash.alibekkyzy@mail.ru);
46. **Bugubayeva Alina** – Associate Professor "Digital Engineering and IT-Analytics", Doctor of Philosophy PhD, Karaganda University of Kazpotrebooyuz, Karaganda, Kazakhstan, E-mail: [alina\\_bugubayeva@mail.ru](mailto:alina_bugubayeva@mail.ru);
47. **Kunapianova Marzhan** – Lecturer of "SDTaAI", D. Serikbayev EKTU, Ust-Kamenogorsk, Kazakhstan, E-mail: [Kmarzhana92@mail.ru](mailto:Kmarzhana92@mail.ru);
48. **Aitolkyn Sametova** – Master, International University of Information Technology, Almaty, Kazakhstan, [aberikkaliyeva@gmail.com](mailto:aberikkaliyeva@gmail.com);
49. **Rashidinov Damir** – PhD candidate, International University of Information Technology, Almaty, Kazakhstan, [damir.dmr88@gmail.com](mailto:damir.dmr88@gmail.com);
50. **Nayzabaeva Lyazzat** – Acting Professor Doctor of Technical Sciences, International University of Information Technology, Almaty, Kazakhstan, [l.naizabayeva@edu.iitu.kz](mailto:l.naizabayeva@edu.iitu.kz);
51. **Suleimenova Madina** – Master of Engineering Sciences, Assistant Professor of the Department of Information Systems, [madekin940@gmail.com](mailto:madekin940@gmail.com);
52. **Abay Koshekov** – PhD, Associate Professor, Department of Aviation Engineering and Technology, Civil Aviation Academy, Almaty, Kazakhstan, E-mail: [a.k.koshekov@gmail.com](mailto:a.k.koshekov@gmail.com);
53. **Dauletbek Toleubekov** – Postgraduate student, Civil Aviation Academy, Almaty, Kazakhstan, E-mail: [d.toleubekov@agakaz.kz](mailto:d.toleubekov@agakaz.kz);
54. **Ibtisema Nautan** – Postgraduate student, Civil Aviation Academy, Almaty, Kazakhstan, E-mail: [i.nautan@agakaz.kz4](mailto:i.nautan@agakaz.kz4);
55. **Yakub Kurbanov** – Doctoral student, Civil Aviation Academy, Almaty, Kazakhstan, E-mail: [y.kurbanov@agakaz.kz](mailto:y.kurbanov@agakaz.kz);
56. **Rustam Togambayev** – Doctoral student, Civil Aviation Academy, Almaty, Kazakhstan, E-mail: [r.togambayev@tsi.lv](mailto:r.togambayev@tsi.lv)

**Азаматтық авиация академиясының Жаршысы**

Ғылыми журнал

2015 жылдан бастап шыға бастады

Қазақстан Республикасы Мәдениет және ақпарат министрлігінде тіркеліп, 2025 жылдың 31 қаңтарында мерзімді баспасөз басылымын, ақпараттық агенттікті және желілік басылымды қайта есепке қою туралы Куәлігі берілді (№ KZ89VPY00111238)

**Вестник Академии гражданской авиации**

Научный журнал

Издается с 2015 г.

Зарегистрирован Министерством культуры и информации Республики Казахстан  
Свидетельство № KZ89VPY00111238 от 31 января 2025 г.

**Bulletin of Civil aviation Academy**

Scientific journal

Published since 2015

It is registered by the Ministry of Culture and Information of the Republic of Kazakhstan.  
Certificate no. KZ89VPY00111238 dated January 31, 2025

---

Жауапты редактор: А. Т. Макеева

Ответственный редактор: А.Т. Макеева

Responsible editor: A.T. Makeeva

---

Басылымға 28.03.2025 ж. қол қойылды.

Формат 205\*290. Өлшемі - баспа табақшасы. Таралымы 50 дана.

Издание подписано 28.03.2025 г.

Формат 205\*290. Размер –печатная пластина. Тираж 50 экз.

The publication was signed on, March 28, 2025.

The format is 205\*290. The size is a printed plate. Edition of 50 copies

---

Азаматтық авиация академиясы. 050039, Алматы қаласы, Закарпатская көшесі,44  
Академия гражданской авиации. 050039, г. Алматы, ул. Закарпатская,44  
Academy of Civil Aviation. 44 Zakarpatskaya St., 050039, Almaty