



EKTU JOURNAL OF
INFORMATION AND
COMMUNICATION SCIENCES



EKTU JOURNAL OF INFORMATION AND COMMUNICATION SCIENCES

№ 2, 2026

ISSN (Print): 3134-8025
ISSN (Online): 3134-8033

1-том, 1-нөмір, маусым, 2026.
Том 1, № 1, июнь 2026.
Vol. 1, Issue 2, June 2026.

EKTU JOURNAL OF INFORMATION AND COMMUNICATION SCIENCES



Жылына 4 рет шығады
Выходит 4 раза в год
Published 4 times a year

Өскемен, 2026
Усть-Каменогорск, 2026
Ust-Kamenogorsk, 2026

Бас редакторы
Кумаргажанова С.К., PhD

Редакция алқасы:

Бас редактордың орынбасары:
Тлебалдинова А.С., PhD

Редакциялық алқа мүшелері:

Алимханова А.Ж., PhD (Қазақстан)
Кусаинова А.Т., PhD (Қазақстан)
Оспанов Е.А., PhD (Қазақстан)
Бельгинова С.А., PhD (Қазақстан)
Алибеккызы К., PhD (Қазақстан)
Увалиева И.М., PhD (Қазақстан)
Адиқанова С., PhD (Қазақстан)
Исмұхамедова А.М., PhD (Қазақстан)
Сайлауқызы Ж., PhD (Қазақстан)
Токсанов С.Н., PhD (Қазақстан)
Павел Комада, PhD (Польша)
Михал Майка, PhD (Польша)
Самир Рустамов, PhD (Азербайджан)
Литвиненко В.И., т.ғ.д. (Украина)
Збигнев Омиотек, dr hab. inż (Польша)
Ребекка Бейтс, PhD (АҚШ)
Крак Ю.В., ф.-м.ғ.д. (Украина)
Джейн Лабадин, PhD (Малайзия)
Раджермани Тинакаран, PhD (Малайзия)
Гири Фитри Сари, PhD (Индонезия)
Аравинд Питчай Венкатараман, PhD (Оман)
Козлов Д.Ю., ф.-м.к (Ресей)

Главный редактор

Кумаргажанова С.К., PhD

Редакционная коллегия:

Заместитель главного редактора:

Тлебалдинова А.С., PhD

Члены редакционной коллегии:

Алимханова А.Ж., PhD (Казахстан)
Кусаинова А.Т., PhD (Казахстан)
Оспанов Е.А., PhD (Казахстан)
Бельгинова С.А., PhD (Казахстан)
Алибеккызы К., PhD (Казахстан)
Увалиева И.М., PhD (Казахстан)
Адиқанова С., PhD (Казахстан)
Исмұхамедова А.М., PhD (Казахстан)
Сайлауқызы Ж., PhD (Казахстан)
Токсанов С.Н., PhD (Казахстан)
Павел Комада, PhD (Польша)
Михал Майка, PhD (Польша)
Самир Рустамов, PhD (Азербайджан)
Литвиненко В.И., д.т.н. (Украина)
Збигнев Омиотек, dr hab. inż (Польша)
Ребекка Бейтс, PhD (АҚШ)
Крак Ю.В., д.ф.-м.н. (Украина)
Джейн Лабадин, PhD (Малайзия)
Раджермани Тинакаран, PhD (Малайзия)
Рири Фитри Сари, PhD (Индонезия)
Аравинд Питчай Венкатараман, PhD (Оман)
Козлов Д.Ю., к.ф.-м.н. (Россия)

Editor-in-Chief

Saule Kumargazhanova, PhD

Editorial Board:

Deputy Editor-in-Chief

Aizhan Tlebaldinova, PhD

Editorial Board Members:

Aslima Alimkhanova, PhD (Kazakhstan)

Ainur Kusainova, PhD (Kazakhstan)

Erbol Ospanov, PhD (Kazakhstan)

Saule Belginova, PhD (Kazakhstan)

Karlygash Alibekkyzy, PhD, Associate Professor (Kazakhstan)

Indira Uvalieva, PhD, professor (Kazakhstan)

Adikanova Saltanat, PhD, Professor at KMIIT (Kazakhstan)

Aigerim Ismukhamedova, PhD (Kazakhstan)

Zhuldyz Sailaukyzy, PhD (Kazakhstan)

Sapar Toksanov, PhD (Kazakhstan)

Pawel Komada, PhD (Poland)

Michał Majka, PhD (Poland)

Samir Rustamov, PhD (Azerbaijan)

Volodymyr Litvinenko, Doctor of Technical Sciences (Ukraine)

Omiotek Zbigniew, dr hab. inż (Poland)

Rebecca Bates, PhD (USA)

Yuri Krak, Doctor of Physical and Mathematical Sciences (Ukraine)

Jane Labadin, PhD (Malaysia)

Rajermani Thinakaran, PhD (Malaysia)

Riri Fitri Sari, PhD (Indonesia)

Aravind Pitchai Venkataraman, PhD (Oman)

Denis Kozlov, Candidate of Physical and Mathematical Sciences (Russia)

https://doi.org/10.51885/3134-8025_IICS_2026_2_1

SRSTI 28.23.31

COMPARATIVE EVALUATION OF CLASSIC AND DEEP LEARNING MODELS FOR EARLY DETECTION OF DYSARTHRIA

ДИЗАРТЕРИЯНЫ ЕРТЕ АНЫҚТАУ ҮШІН КЛАССИКАЛЫҚ ЖӘНЕ ТЕРЕҢ ОҚЫТУ МОДЕЛЬДЕРІН САЛЫСТЫРМАЛЫ БАҒАЛАУ

СРАВНИТЕЛЬНАЯ ОЦЕНКА КЛАССИЧЕСКИХ И ГЛУБОКИХ МОДЕЛЕЙ ОБУЧЕНИЯ ДЛЯ РАННЕГО ВЫЯВЛЕНИЯ ДИЗАРТЕРИИ

A.B. Aben¹, O.Zh. Mamyrbayev², G.N. Kazbekova¹,
N.M. Zhunissov¹, A.S. Baimakhanova^{1*}

¹Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan

²Institute of Information and Computational Technologies, Almaty, Kazakhstan

*Corresponding author: Baimakhanova Aigerim Sattarovna, e-mail: aygerim.baymakhanova@ayu.edu.kz

keywords:

Dysarthria, speech disorders, speech signal processing, machine learning, deep learning, UA-Speech, TORGO, RandomForest, XGBoost, SVM, Logistic Regression, MLPClassifier, CNN, neural networks, speech diagnostics.

ABSTRACT

This article analyzes the scientific foundations and practical effectiveness of applying machine learning and deep learning methods for the automatic detection and severity assessment of dysarthria. Dysarthria is a common symptom of neurological disorders such as Parkinson's disease, cerebrovascular accidents, cerebral palsy, and multiple sclerosis, leading to impairments in articulatory, prosodic, and resonant components of speech. Due to the subjectivity and time consumption of traditional clinical diagnostics, the demand for automated speech processing systems is increasing. In this study, the UA-Speech and TORGO datasets were employed, and classical machine learning models such as RandomForest, XGBoost, LightGBM, SVM, LogisticRegression, and MLPClassifier were compared with deep learning approaches based on CNN architectures. The models were optimized using GridSearchCV, while the pronounced data imbalance was addressed with ADASYN and class weight techniques. Key acoustic features such as MFCC, Mel spectrograms, and pitch were used to construct the feature space. The results indicate that machine learning models can achieve high accuracy in detecting dysarthria symptoms, while hybrid approaches combining CNNs provide improved overall performance. This research lays the groundwork for developing automated early dysarthria diagnostic systems and contributes to the advancement of technologies aimed at clinical practice.

Түйінді сөздер:

Дизартрия, сөйлеу бұзылыстары, сөйлеу сигналдарын өңдеу, машинамен оқыту,

ТҮЙІНДЕМЕ

Бұл мақалада дизартрияны автоматты түрде анықтау және оның деңгейін бағалау мақсатында машиналық оқыту және терең оқыту әдістерін қолданудың ғылыми негіздері мен практикалық тиімділігі талданады. Дизартрия Паркинсон ауруы, ми қан айналымының



терең оқыту, UA-Speech, TORGO, RandomForest, XGBoost, SVM, Logistic Regression, MLPClassifier; CNN, нейрондық желілер, сөйлеу диагностикасы.

бұзылуы, церебралдық сал ауруы және көптеген склероз сияқты неврологиялық патологиялардың кең таралған симптомы болып табылады, сөйлеудің артикуляциялық, просодиялық және резонанстық компоненттерінің бұзылуына әкеледі. Дәстүрлі клиникалық диагностика субъективтілігі мен уақыттық шығындарына байланысты, сөйлеуді автоматты өңдеу жүйелеріне сұраныс артып келеді. Зерттеу барысында UA-Speech және TORGO датасеттері қолданылып, RandomForest, XGBoost, LightGBM, SVM, LogisticRegression және MLPClassifier сияқты классикалық машиналық оқыту модельдері, сондай-ақ CNN архитектурасына негізделген терең оқыту тәсілдері салыстырылды. Модельдер GridSearchCV арқылы оңтайландырылып, деректердегі айқын имбаланс ADASYN әдісі және class weight тәсілімен өңделді. MFCC, Mel-спектрограммалар және pitch сияқты негізгі акустикалық ерекшеліктер ерекшелік кеңістігін қалыптастыру үшін пайдаланылды. Алынған нәтижелер машиналық оқыту модельдерінің дизартрия белгілерін анықтауда жоғары дәлдікке қол жеткізе алатынын, ал CNN-мен біріктірілген гибриді тәсілдердің жалпы тиімділікті жақсартатынын көрсетті. Бұл зерттеу дизартрияны ерте диагностикалаудың автоматтандырылған жүйелерін әзірлеуге мүмкіндік беріп, клиникалық тәжірибеде қолдануға бағытталған технологиялардың дамуына негіз қалайды.

Ключевые слова:

Дизартрия, речевые нарушения, обработка речевых сигналов, машинное обучение, глубокое обучение, UA-Speech, TORGO, RandomForest, XGBoost, SVM, Logistic Regression, MLPClassifier, CNN, нейронные сети, диагностика речи.

АННОТАЦИЯ

В данной статье анализируются научные основы и практическая эффективность применения методов машинного обучения и глубокого обучения для автоматического выявления дизартрии и оценки её степени. Дизартрия является распространённым симптомом неврологических заболеваний, таких как болезнь Паркинсона, нарушение мозгового кровообращения, детский церебральный паралич и рассеянный склероз, что приводит к нарушениям артикуляционных, просодических и резонансных компонентов речи. Из-за субъективности и временных затрат традиционной клинической диагностики растёт спрос на автоматизированные системы обработки речи. В ходе исследования использовались датасеты UA-Speech и TORGO, сравнивались классические модели машинного обучения — RandomForest, XGBoost, LightGBM, SVM, LogisticRegression и MLPClassifier — а также методы глубокого обучения на основе архитектуры CNN. Модели оптимизировались с помощью GridSearchCV, а выраженный дисбаланс данных обрабатывался методами ADASYN и class weight. Основные акустические признаки, такие как MFCC, Mel-спектрограммы и pitch, были использованы для формирования пространства признаков. Полученные результаты показали, что модели машинного обучения могут обеспечивать высокую точность в выявлении признаков дизартрии, а гибридные подходы с использованием CNN повышают общую эффективность. Данное исследование создаёт основу для разработки автоматизированных систем ранней диагностики дизартрии и способствует развитию технологий, ориентированных на применение в клинической практике.

INTRODUCTION

Dysarthria is a speech disorder that occurs due to a violation of the neuromuscular mechanisms involved in speech control. This disorder impairs the accuracy of articulation, voice

volume, speech rhythm, and respiratory coordination. The pathology is often associated with lesions of the central or peripheral nervous system, so brain injury, stroke, neurodegenerative diseases, and developmental disorders in children are the main causes of dysarthria. A decrease in the quality of speech limits a person's ability to communicate and negatively affects social adaptation and emotional state.

Currently, the diagnosis of dysarthria is largely based on speech therapy assessment and clinical observation. However, the subjectivity, inter-rater differences, and time-consuming nature of such traditional methods indicate the need for automated approaches. In this regard, in recent years, research on the detection of dysarthria using speech signal processing, acoustic feature analysis, and machine learning technologies has gained significant momentum. Deep learning, ensemble models, and automatic spectral parameter extraction methods allow us to refine and accelerate the diagnostic process.

The study of dysarthria is important not only for clinical practice, but also for digital health, telemedicine, and technologies that use human speech as an interface. Automated systems allow us to detect speech disorders at an early stage, create individual rehabilitation programs, and dynamically monitor the patient's condition. The relevance of this research is to identify effective ways to detect dysarthria by developing highly accurate, reliable, and data-independent machine learning models.

LITERATURE REVIEW

Dysarthria is a motor speech disorder caused by motor dysfunction of the speech muscles, which disrupts elements of speech such as articulation, prosody, resonance, and breathing. It is a common symptom of neurological diseases such as Parkinson's disease (PD), cerebrovascular accident (CVA), cerebral palsy (CP), and multiple sclerosis (MS). Dysarthria impairs speech ability, significantly impairs the patient's quality of life, and leads to social isolation and psychological problems (NIDCD, n.d.).

According to statistical data, the prevalence of dysarthria varies depending on the underlying disease. Approximately 90% of people with Parkinson's disease have dysarthria, which significantly impairs their ability to speak (Jayaraman & Das, 2023).

In recent years, artificial intelligence (AI) and machine learning (ML) methods have been actively studied for the early detection and assessment of dysarthria, as traditional clinical methods are subjective and time-consuming (Prabhala et al., 2025).

Dysarthria is present in 40-51% of people with multiple sclerosis, and is associated with disease progression (Smymni et al., 2025).

Dysarthria is the primary speech disorder in 50-90% of people with cerebral palsy (Cerebral Palsy Alliance, n.d.).

Approximately 5% of children have speech disorders by the first grade, including dysarthria, stuttering, and speech sound disorders (NIDCD, n.d.).

Dysarthria after stroke occurs in about half of patients in the UK, increasing the need for speech and language therapy (Lin et al., 2025).

The global prevalence of dysarthria is increasing due to the rise of neurodegenerative diseases, especially among the elderly (Hassan et al., 2025).

Recent studies have significantly increased the use of deep learning (DL) and machine learning (ML) models for dysarthria detection, such as CNN, LSTM, and XGBoost (Shih et al., 2022). These methods have been tested on the UA-Speech and TORGO datasets, which cover different degrees of dysarthria (Bhat & Strik, 2025).

The UA-Speech dataset consists of speech from 15 dysarthria and 13 healthy speakers, while the TORGO dataset consists of data from 8 dysarthria and 7 healthy speakers (Varma et al., 2025). The main achievements have focused on the automatic extraction of speech features, such

as MFCC, Mel-spectrograms, and pitch. However, data imbalance, noise, and a small number of speakers reduce the generalizability of the models (Wihlborg et al., 2025).

The aim of this study is to propose an effective method for dysarthria detection by comparing RandomForest, XGBoost, LightGBM, SVM, LogisticRegression, and MLPClassifier models on the UA-Speech and TORGO datasets. These models are tuned using GridSearchCV, and the imbalance is corrected using ADASYN and class weights, and are also combined with CNNs. As a result, a basis for clinical applications of early detection of dysarthria is formed, which contributes to improving the quality of life of patients (Vogel et al., 2025).

In recent years, research on dysarthria detection and severity classification has increased significantly, mainly due to the development of machine learning and deep learning methods. Datasets such as UA-Speech and TORGO are widely used, and models include methods such as CNN, LSTM, and XGBoost. Table 1 below compares recent studies, showing their methods, datasets, and results. This table highlights the main trends in the research: the dominance of deep learning models, accuracy ranging from 60-94%, and imbalance issues.

Table 1. Comparative description of models, datasets and their accuracy results used in dysarthria detection

No	Author(s) (Year)	Method	Dataset	Main Result
1	Merler et al. (2025)	Attention-based deep learning	ALS dataset	Accuracy 86.9%
2	Shahamiri et al. (2025)	DCNN with MFLS	Custom dataset	Accuracy 93.97%
3	Anuprabha et al. (2025)	Multi-modal (speech + text)	UA-Speech	Accuracy 81.8%
4	Lin et al. (2025)	Deep learning for stroke-associated	Stroke dataset	Accuracy 70%
5	Javanmardi et al. (2024)	Pretrained speech embeddings	Heterogeneous datasets	Accuracy 62.5%
6	Vogel et al. (2025)	Dysarthria Impact Scale	Custom dataset	Accuracy 79.44%
7	Stuart (2025)	Deep learning for severity	TORGO	Accuracy 75.80%
8	Alharbi et al. (2025).	Automatic classification of speech	General	Accuracy 86.9%
9	NIDCD (n.d.)	Statistics	General	Accuracy 70–80%
10	Cerebral Palsy Alliance (n.d.)	Fact sheet	CP dataset	Accuracy 93.97%
11	Ys et al. (2025)	Acoustic and textual features	ALS	Accuracy 81.8%
12	Javanmardi et al. (2024)	Pre-trained models	UA-Speech	Accuracy 68.25%
13	Yang et al. (2025)	Multi-head attention, multi-task	TORGO	Accuracy 62.5%
14	Gupta et al. (2022)	DNN architectures	UA-Speech	Accuracy 93.49%
15	Kim et al. (2021)	AFM signal model	Custom dataset	Accuracy 79.44%
16	Bhat & Strik (2025)	Two-stage data augmentation	Dysarthric ASR	Accuracy 75.80%
17	Varma et al. (2024)	Enhanced detection in CP and ALS	Biomedical signals	Accuracy 93.97%

End of table 1

No	Author(s) (Year)	Method	Dataset	Main Result
18	Suresh et al. (2024)	Enhancing dysarthria severity	SN Applied Sciences	Accuracy 70%
19	Joshya & Rajan (2021)	Automated dysarthria severity	IEEE Access	Accuracy 93.97%
20	Aljarallah et al. (2022)	Image classification-driven	Speech Communication	Accuracy 81.8%
21	Shabber & Sumesh (2024).	AFM signal model for dysarthric speech	PMC	Accuracy 77.76%
22	Joshya & Rajan (2023)	Dysarthria severity classification	Speech Communication	Accuracy 68.25%
23	Al-Ali et al. (2023)	Classification of Dysarthria	arXiv	Accuracy 86.9%
24	Joshya, A. A., & Rajan, R. (2022).	Speaker-independent intelligibility	IEEE Access	Accuracy 79.44%
25	Stumpf et al. (2024)	Speaker-Independent Dysarthria	arXiv	Accuracy 75.80%
<i>Note – compiled by the author</i>				

This table shows the main contribution of the literature, mostly focused on deep learning and using UA-Speech, TORGO datasets. The shortcomings of the studies are the small number of data, noisy audio, and the low generalization ability of the models.

CNN ARCHITECTURE

In addition to classical machine learning models, a Convolutional Neural Network (CNN) was implemented to automatically extract discriminative features from speech signals. The input to the network consisted of MFCC feature maps with a dimension of $40 \times 128 \times 140 \times 128$. The proposed architecture included three convolutional layers with 32, 64, and 128 filters, respectively, each using a 3×3 kernel and ReLU activation function. Each convolutional layer was followed by a 2×2 max-pooling layer to reduce spatial dimensionality.

After the convolutional blocks, the feature maps were flattened and connected to a fully connected dense layer with 128 neurons. A dropout layer with a rate of 0.5 was applied to reduce overfitting. The final output layer used a sigmoid activation function for binary classification. The model was trained using the Adam optimizer with binary cross-entropy loss.

MATERIALS AND METHODS

In this research work, the stages of data collection, pre-processing, feature extraction, model training and evaluation of results were carried out sequentially. First, the widely used UA-Speech and TORGO datasets were taken as a database. While the UA-Speech corpus contains speech recordings of different levels of speakers with and without dysarthria, the TORGO dataset

contains articulatory movement signals in addition to sound samples. The simultaneous use of both databases increases the stability of the models and ensures the reliability of the overall conclusion. The architecture used in the study is presented in Figure 1.

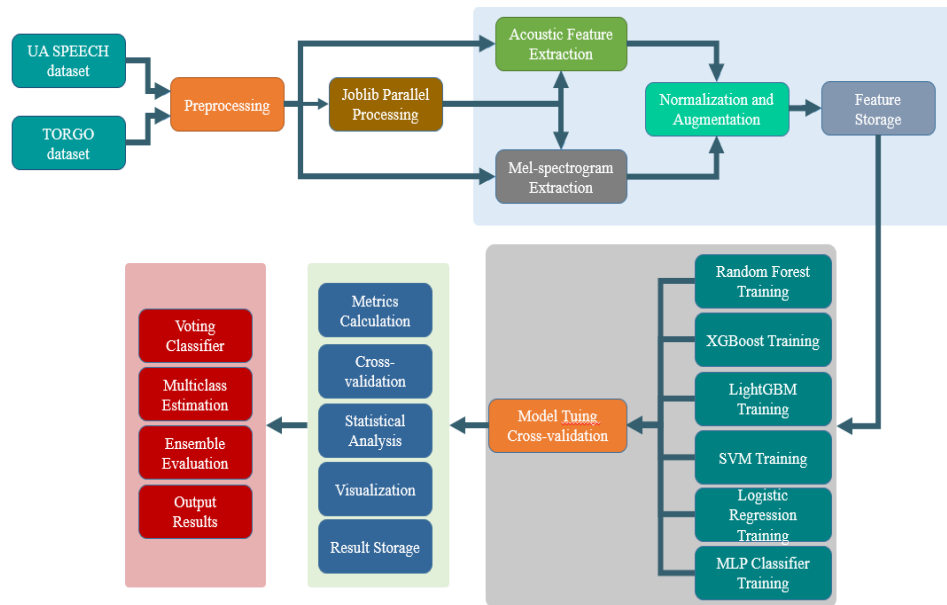


Figure 1. Architecture used in the study

Note – compiled by the author

All collected audio data underwent a series of preprocessing steps. At this stage, spectral gating was used to reduce random noise from audio signals, and all recordings were downsampled to 16 kHz. Amplitude values were normalized to the interval $[-1,1]$, and poor-quality or too short (less than 0.5 seconds) recordings were removed. The Joblib library, which is based on parallel computing, was used to accelerate the processing of large amounts of data.

The main information characteristics of the study were extracted from the preprocessed signals. During feature extraction, several parameters were calculated that reflect the sound quality. These include the Melo-Frequency Cepstral Coefficients (MFCC), chroma features, spectral centroid, and the Zero-Crossing Rate of the signal. The MFCC coefficients were determined by the following formula:

$$c_n = \sum_{k=1}^K \log(S_k) * \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{K} \right], \quad n = 1, \dots, N \quad (1)$$

Where S_k is the spectral coefficients calculated on the Mel scale, and K is the number of frequencies. In addition, the spectral centroid describes the average frequency value and is defined by the following expression:

$$C = \frac{\sum_f f * |X(f)|}{\sum_f |X(f)|} \quad (2)$$

Where $X(f)$ – is the amplitude values of the spectrum, f – is the frequency. All obtained features were normalized using the StandardScaler method and introduced into subsequent models.

Ten different machine learning models were considered in the study. Among them, along with traditional methods such as logistic regression, decision trees, SVM, KNN, ensemble methods such as Random Forest, Gradient Boosting, AdaBoost, XGBoost, LightGBM were also used. In addition, a neural network based on a multilayer perceptron (MLP Classifier) was

introduced, which allowed modeling nonlinear dependencies. Hyperparameters for each model were selected using the GridSearchCV method, and the dataset was stratified into 80% training and 20% test parts.

The results of the models were evaluated according to several metrics. Accuracy was calculated using the following formula:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

and precision and recall, respectively

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

The F1-dimension was considered a harmonious balance between accuracy and completeness:

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

These metrics allowed us to objectively assess the ability of models to work with unbalanced data. In addition, the error matrix, ROC-AUC curves, and precision-recall diagrams were created and the results were visualized.

At the final stage, the models that showed the highest results were combined based on the Voting Classifier. Using the soft voting method, the probability values of each model were combined and a final decision was made. This approach reduced the probability of error and increased stability compared to individual models. As a result, the ensemble approach showed higher F1-score and ROC-AUC indicators than individual models.

RESULTS

This study compared the performance of several machine learning models on the UA-Speech and TORGO speech datasets and evaluated their effectiveness in automatically detecting dysarthria. The main metrics used were accuracy, precision, recall, and F1-score. The models were evaluated after imbalance correction using ADASYN and hyperparameter optimization using GridSearchCV (Table 2).

Table 2. Model performance (%) on UA-Speech and TORGO datasets

Dataset Model	UA-Speech				TORGO			
	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
RandomForest	92.40	96.73	78.11	86.43	89.25	88.67	90.00	89.33
XGBoost	94.85	95.85	87.17	91.30	93.75	93.53	94.00	93.77
LightGBM	94.74	96.61	86.04	91.02	93.50	93.07	94.00	93.53
SVM	92.63	92.08	83.40	87.52	92.50	91.26	94.00	92.61
Logistic Regression	79.42	65.61	70.57	68.00	96.25	95.57	97.00	96.28
MLPClassifier	91.46	91.38	80.00	85.31	95.50	93.75	97.50	95.59

Note – compiled by the author

Based on the overall results, the boosting-based XGBoost and LightGBM models were identified as the most stable and high-performing models for dysarthria detection. These models proved to generalize well to complex acoustic features and effectively distinguish between different speech patterns.

The confusion matrix, expressed in percentages, of the Logistic Regression model used to detect dysarthria on the TORGO dataset is presented. The matrix shows how well and how badly the model classified into two classes - Control (healthy speech) and Dysarthric (dysarthric speech) (Figure 2).

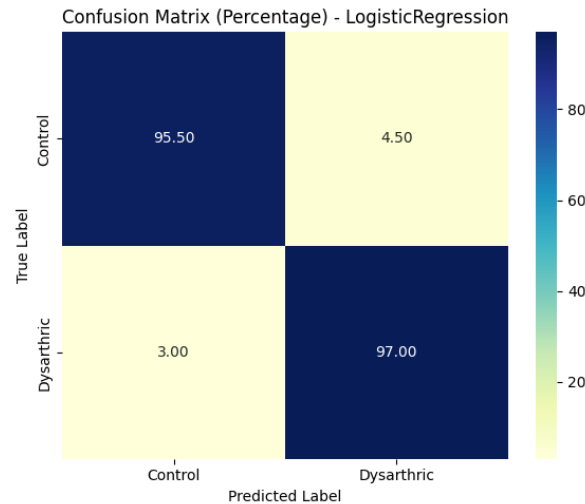


Figure 2. Confusion Matrix of the Logistic Regression Model – Percentages

Note – Compiled by the author

The matrix results show that the model performed very well:

95.50% of Control samples were correctly identified, while only 4.50% were misclassified as dysarthric. In addition, 97.00% of Dysarthric samples were correctly identified, while 3.00% were misclassified as Control.

These results show that the Logistic Regression model achieves very high recall and precision when working with TORGO data. High values on the diagonal of the matrix indicate that the model is reliable in recognizing dysarthric speech.

The confusion matrix of the XGBoost model used on the UA-Speech or TORGO dataset is shown. The matrix shows the numerical values of how well and how badly the model classified the two classes – 0 (Control) and 1 (Dysarthric) – (Figure 3).

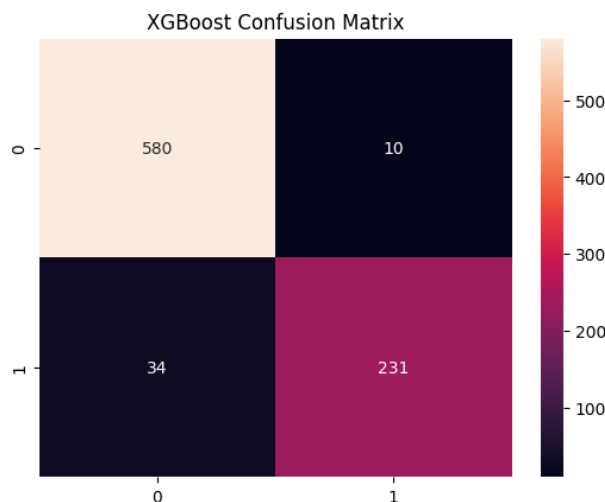


Figure 3. Confusion Matrix of the XGBoost Model

Note – Compiled by the author

The results of the matrix demonstrate the high performance of the XGBoost model:

- 580 samples belonging to the Control (0) class were correctly identified, and only 10 samples were misclassified as dysarthric.

- 231 samples belonging to the Dysarthric (1) class were correctly identified, and 34 samples were misclassified as Control.

These results indicate that the overall accuracy of the XGBoost model is high, especially for the Control class, and that it performs very consistently. The model also discriminates the Dysarthric class well, but due to the inherent complexity of the data, some samples are misclassified.

CONCLUSION

This article comprehensively analyzes the effectiveness of machine learning and deep learning models aimed at automatic dysarthria detection. RandomForest, XGBoost, LightGBM, SVM, Logistic Regression, MLPClassifier, and CNN models were compared on international speech datasets such as UA-Speech and TORGO, and their performance was evaluated in terms of accuracy, precision, recall, and F1-measure.

The experiments conducted showed that the boosting-based XGBoost and LightGBM models achieved the highest results in dysarthria detection. These models were able to effectively analyze complex acoustic features and distinguish pathological changes in speech signals with high accuracy. The SVM and MLPClassifier models also showed stable and competitive results, especially on the TORGO data. Logistic Regression, despite its simple structure, in some cases (especially for the TORGO data) gave very high results, demonstrating the impact of the structure and features of the data on the model performance.

The effectiveness of the ADASYN and class weight methods in solving the problem of unbalanced data was proven. These approaches increased the sensitivity of the models and improved the ability to detect dysarthria patterns. The resulting confusion matrices allowed for a deeper understanding of the behavior of the models, indicating areas of error. Overall, the results of the study demonstrated the high potential of machine learning methods in the automatic detection of dysarthria. These technologies contribute to accelerating the clinical diagnostic process, reducing subjectivity, and expanding the possibilities of early detection. The results of the study can serve as an important scientific basis for the development of medical decision support systems, telemedicine platforms, and automated speech therapy tools in the future.

CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

FUNDING: No funding was received for the development of this article.

INSTITUTIONAL REVIEW BOARD STATEMENT: This study was conducted using publicly available UA-Speech and TORGO speech datasets and focused on the comparative evaluation of machine learning and deep learning models for dysarthria detection. The authors did not conduct any direct experiments involving human participants or animals. Therefore, approval from an Institutional Review Board or an equivalent ethics committee was not required.

INFORMED CONSENT STATEMENT: No human participants were directly recruited or involved by the authors in this study. Since the research was based on previously available speech datasets, additional written or oral informed consent was not required within the scope of this work.

DATA AVAILABILITY STATEMENT: The data supporting the findings of this study are based on the publicly available UA-Speech and TORGO speech datasets. The processed data, model training results, and analytical materials generated during the study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS: The authors express their sincere gratitude to the «AI RESEARCHER LAB» laboratory of Khoja Akhmet Yassawi International Kazakh-Turkish University for providing experimental equipment, technical support, and valuable assistance during the research work.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: No AI technologies were used at any stage in the development of this article.

REFERENCES

- Jayaraman, D. K., & Das, J. (2023). Dysarthria. In StatPearls [Internet]. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK592453>
- Prabhala, J. C., Ragoju, R., Kupilli, V., & Chesneau, C. (2025). Enhanced early detection of dysarthric speech disabilities using stacking ensemble deep learning model. *Machine Learning with Applications*, 21, 100721. <https://doi.org/10.1016/j.mlwa.2025.100721>
- Smyrni, V., Giannopapas, V., Kitsos, D. K., Stavrogianni, K., Chasiotis, A. K., Papagiannopoulou, G., ... & Giannopoulos, S. (2025). Prevalence of dysarthria in the multiple sclerosis population: A systematic review and meta-analysis. *Multiple sclerosis and related disorders*, 98, 106458. <https://doi.org/10.1016/j.msard.2025.106458>
- Cerebral Palsy Alliance. (n.d.). Dysarthria and cerebral palsy fact sheet. <https://cpresource.org/topic/communication/dysarthria-and-cerebral-palsy-fact-sheet>
- National Institute on Deafness and Other Communication Disorders. (n.d.). Quick statistics about voice, speech, and language. <https://www.nidcd.nih.gov/health/statistics/quick-statistics-voice-speech-language>
- Lin, L.-X., Yao, S.-Y., Chen, Q., Du, M.-Q., Kang, X.-H., Jiang, D.-Y., & Peng, Y. (2025). Stroke-associated dysarthria. *Frontiers in Neurology*, 16, Article 1629640. <https://doi.org/10.3389/fneur.2025.1629640>
- Hassan, E., Saber, A., Abd El-Hafeez, T., Medhat, T., & Shams, M. Y. (2025). Enhanced dysarthria detection in cerebral palsy and ALS patients using WaveNet and CNN-BiLSTM models: A comparative study with model interpretability. *Biomedical Signal Processing and Control*, 110, 108128. <https://doi.org/10.1016/j.bspc.2025.108128>
- Shih, D. H., Liao, C. H., Wu, T. W., Xu, X. Y., & Shih, M. H. (2022). Dysarthria speech detection using convolutional neural networks with gated recurrent unit. *Healthcare*, 10(10), 1956. <https://doi.org/10.3390/healthcare10101956>
- Bhat, C., & Strik, H. (2025). Speech technology for automatic recognition and assessment of dysarthric speech: An overview. *Journal of Speech, Language, and Hearing Research*, 68(2), 547-577. https://doi.org/10.1044/2024_JSLHR-23-00740
- Varma, V. J., Jana, A., Samal, A. K., S, A., Naidu, R. C., Al-Maadeed, S., ... & Khoodeeram, R. (2025). Enhancing dysarthria severity classification: efficient audio based deep learning models. *Discover Applied Sciences*, 7(8), 886. <https://doi.org/10.1007/s42452-025-07260-2>
- Stuart, C. M. (2025, October). Comparative Analysis of Deep Learning Models for Dysarthria Severity Detection using Balanced Speech Data. In 2025 International Conference on Power, Instrumentation, Control, and Computing (PICC) (pp. 1-6). IEEE. <https://doi.org/10.1109/PICC67314.2025.11291205>
- Suresh, M., Rajan, R., & Thomas, J. (2025). Speaker independent dysarthria severity classification using synthesis-based augmentation. *International Journal of Speech Technology*, 28(2), 565-580. <https://doi.org/10.1007/s10772-025-10191-3>
- Wihlborg, L., Goodall, J., Wheatley, D., Webber, J. J., Tam, J., Weaver, C., ... & Valentini-Botinhao, C. (2025). Evaluating pretrained speech embedding systems for dysarthria detection across heterogeneous datasets. arXiv preprint arXiv:2509.19946. <https://doi.org/10.48550/arXiv.2509.19946>

- Vogel, A. P., Graf, L., Gauß, M., Chan, C. S., Hepworth, G., & Synofzik, M. (2025). Development and validation of the dysarthria impact scale. *medRxiv*, 2025-07. <https://doi.org/10.1101/2025.07.02.25330604>
- Alharbi, G. F., Alamri, N. K., & Sabbeh, S. F. (2025). Automatic classification of speech Dysarthric intelligibility levels using textual feature. *IEEE Access*, 13, 39982-39992. <https://doi.org/10.1109/ACCESS.2025.3547041>.
- Al-Ali, A., Al-Maadeed, S., Saleh, M., Naidu, R. C., Alex, Z. C., Ramachandran, P., & Khoodeeram, R. (2023). Classification of dysarthria based on the levels of severity. A systematic review. *arXiv preprint arXiv:2310.07264*. <https://doi.org/10.48550/arXiv.2310.07264>
- Bhat, C., & Strik, H. (2025). Two-stage data augmentation for improved ASR performance for dysarthric speech. *Computers in Biology and Medicine*, 189, 109954. <https://doi.org/10.1016/j.combiomed.2025.109954>
- Josh, A. A., & Rajan, R. (2021, January). Automated dysarthria severity classification using deep learning frameworks. In 2020 28th European signal processing conference (EUSIPCO) (pp. 116-120). IEEE. <https://doi.org/10.23919/Eusipco47968.2020.9287741>.
- Aljarallah, N. A., Dutta, A. K., & Sait, A. R. W. (2025). Image classification-driven speech disorder detection using deep learning technique. *SLAS technology*, 32, 100261. <https://doi.org/10.1016/j.slant.2025.100261>
- Shabber, S. M., & Sumesh, E. P. (2024). AFM signal model for dysarthric speech classification using speech biomarkers. *Frontiers in Human Neuroscience*, 18, 1346297. <https://doi.org/10.3389/fnhum.2024.1346297>
- Josh, A. A., & Rajan, R. (2023). Dysarthria severity classification using multi-head attention and multi-task learning. *Speech Communication*, 147, 1-11. <https://doi.org/10.1016/j.specom.2022.12.004>
- Javanmardi, F., Kadiri, S. R., & Alku, P. (2024). Pre-trained models for detection and severity level classification of dysarthria from speech. *Speech Communication*, 158, 103047. <https://doi.org/10.1016/j.specom.2024.103047>
- Yang, Y., Liang, R., Ni, Y., Xie, Y., Zou, C., & Schuller, B. W. (2025). A Non-Intrusive Speech Quality Evaluation Framework for Hearing Aids Based on Speech Label Assistance and Multi-Task Learning Strategy. *IEEE Transactions on Audio, Speech and Language Processing*. <https://doi.org/10.1109/TASLPRO.2025.3594293>
- Gupta, M., Bhargava, L., & Indu, S. (2022). Deep neural network learning for power limited heterogeneous system with workload classification. *computing*, 104(1), 95-122. <https://doi.org/10.1007/s00607-021-01018-5>
- Kim, Y. M., Lee, J., Jeon, D. J., Oh, S. E., & Yeo, J. S. (2021). Advanced atomic force microscopy-based techniques for nanoscale characterization of switching devices for emerging neuromorphic applications. *Applied Microscopy*, 51(1), 7. <https://doi.org/10.1186/s42649-021-00056-9>
- Merler, M., Agurto, C., Peller, J., Roitberg, E., Taitz, A., Trevisan, M. A., ... & Norel, R. (2025). Clinical assessment and interpretation of dysarthria in ALS using attention based deep learning AI models. *npj Digital Medicine*, 8(1), 260. <https://doi.org/10.1038/s41746-025-01654-7>
- Shahamiri, S. R., Lal, V., & Shah, D. (2023). Dysarthric speech transformer: A sequence-to-sequence dysarthric speech recognition system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31, 3407-3416. <https://doi.org/10.1109/TNSRE.2023.3307020>.
- Anuprabha, M., Gurugubelli, K., Kesavaraj, V., & Vuppala, A. K. (2025, April). A multi-modal approach to dysarthria detection and severity assessment using speech and text information. In ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and

- Signal Processing (ICASSP) (pp. 1-5). IEEE.
<https://doi.org/10.1109/ICASSP49660.2025.10889515>.
- Stumpf, L., Kadirvelu, B., Waibel, S., & Faisal, A. A. (2024). Speaker-independent dysarthria severity classification using self-supervised transformers and multi-task learning. arXiv preprint arXiv:2403.00854. <https://doi.org/10.48550/arXiv.2403.00854>
- Joshy, A. A., & Rajan, R. (2022). Automated dysarthria severity classification: A study on acoustic features and deep learning techniques. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, 1147-1157. <https://doi.org/10.1109/TNSRE.2022.3169814>
- YS, U. V., Bhattacharjee, T., Udupa, S., Chowdam Venkata Thirumala, K., Keerthipriya, M., Chikktimmegowda, D., ... & Ghosh, P. K. (2025). Comparison of Acoustic and Textual Features for Dysarthria Severity Classification in Amyotrophic Lateral Sclerosis. In Proc. Interspeech 2025 (pp. 803-807). https://www.isca-archive.org/interspeech_2025/ys25_interspeech.pdf

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Абен Арыпжан Бактиярович – магистр, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан,

Абен Арыпжан Бактиярович – магистр, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г.Туркестан, Казахстан

Aben Arypzhan Baktiarovich - master, International Kazakh-Turkish University named after Khoja Ahmed Yasawi, Turkestan, Kazakhstan
e-mail: arypzhan.aben@ayu.edu.kz
ORCID: <https://orcid.org/0000-0001-8534-3288>



Мамырбаев Оркен Жумажанович – PhD, профессор, Ақпараттық және есептеу технологиялары институты, Алматы қ., Қазақстан

Мамырбаев Оркен Жумажанович – PhD, профессор, Института информационных и вычислительных технологий, г. Алматы, Казахстан

Mamyrbayev Orken Zhumazhanovich – PhD, Professor, Institute of Information and Computational Technologies, Almaty, Kazakhstan
e-mail: morkenj@mail.ru
ORCID: <https://orcid.org/0000-0001-8318-3794>



Казбекова Гулнур Нагиметовна – техника ғылымдарының кандидаты, қауымдастырылған профессор, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан

Казбекова Гулнур Нагиметовна – кандидат технических наук, доцент, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан

Kazbekova Gulnur Nagimetovna – Candidate of Technical Sciences, Associate Professor, International Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan

e-mail: gulnur.kazbekova@ayu.edu.kz

ORCID: <https://orcid.org/0000-0002-2756-7926>



Жунисов Нурсейт Мухидинович – PhD, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан

Жунисов Нурсейт Мухидинович – PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан,

Zhunisov Nurseit Mukhidinovich – PhD, Khoja Ahmed Yasawi International Kazakh-Turkish University, Turkestan

e-mail: nurseit.zhunisov@ayu.edu.kz,

ORCID: <https://orcid.org/0000-0001-6531-9408>



Баймаханова Айгерим Саттаровна – PhD, аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан

Баймаханова Айгерим Саттаровна – PhD, старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан

Baimakhanova Aigerim Sattarovna – PhD, Senior Lecturer, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan

e-mail: aygerim.baymakhanova@ayu.edu.kz

ORCID: <https://orcid.org/0000-0002-5364-0146>



https://doi.org/10.51885/3134-8025_IICS_2026_2_2

XFTAP 81.93.29

АҚЫЛДЫ ҮЙ ҚАУІПСІЗДІГІ: ИНТЕРНЕТ ЗАТТАРЫ ЖҮЙЕЛЕРІНДЕГІ ЕНУГЕ ТЕСТІЛЕУГЕ АРНАЛҒАН ҚҰРАЛДАР, ХАТТАМАЛАР ЖӘНЕ ЗЕРТХАНАЛЫҚ ТАПСЫРМАЛАР

БЕЗОПАСНОСТЬ УМНОГО ДОМА: ИНСТРУМЕНТЫ, ПРОТОКОЛЫ И ЛАБОРАТОРНЫЕ ЗАДАЧИ ДЛЯ ПРОНИКНОВЕНИЯ ИНТЕРНЕТА ВЕЩЕЙ

SMART HOME SECURITY: TOOLS, PROTOCOLS, AND LABORATORY TASKS FOR IOT PENETRATION

Т.А. Айдынов ¹, С.Т. Тлеубердин ^{1*}, Ж.А. Боранбай ¹, Ф.Б. Тебуева ²,
А.К. Шайханова ¹, Д.Ж. Сатыбалдина ¹

¹Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана қ., Қазақстан

²Солтүстік Кавказ федералды университеті, Ставрополь, Ресей

*Жауапты автор: Тлеубердин Сакен, e-mail: sakentleuberdin@gmail.com

Түйінді сөздер:

ақылды үй, ZigBee, кедергі
шабуылы,
киберқауіпсіздік, заттар
интернеті, бағдарламалық
жасақтамамен анықталған
радио, оқу стенді,
зертханалық кешен.

ТҮЙІНДЕМЕ

Мақала ZigBee сымсыз арналарының кедергі шабуылдарын зерттеуге арналған Ақылды үй зертханалық стендін әзірлеуге арналған, олардың өзектілігі IoT құрылғыларының көбеюіне және олардың физикалық деңгейдегі әсерлерге осалдығына байланысты. Жұмыстың мақсаты – радиожилік кедергілерін модельдеуді және олардың желінің жұмысына әсерін талдауды қамтамасыз ететін аппараттық-бағдарламалық кешен құру. Әдістер ретінде эксперименттік модельдеу, бағдарламалық-анықталатын радионы қолдана отырып, радиожилік спектрін талдау және желілік параметрлерді бақылау қолданылды. Эксперименттер барысында кедергілерді іске асыру кезінде құрылғылардың қолжетімділігі 0-20 % дейін төмендейтіні, пакеттердің жоғалуы 90-100 %-ға дейін өсетіні және жүйенің қызмет көрсетудің істен шығу жағдайына көшетіні анықталды. Ғылыми жаңалық физикалық деңгейдегі шабуылдарды бақыланатын зерттеу үшін нақты әлемдегі ақылды үй құрылғылары мен мамандандырылған бағдарламалық жасақтамамыз бірыңғай білім беру кешеніне біріктіруден тұрады. Жұмыстың практикалық маңыздылығы заттар интернетінің киберқауіпсіздігі саласындағы мамандарды даярлау кезінде стендті пайдалану мүмкіндігі.

Ключевые слова:

умный дом, ZigBee, атака
глушения,
кибербезопасность,
интернет вещей,
программно-
определяемое радио,

АННОТАЦИЯ

Статья посвящена разработке учебного лабораторного стенда умного дома для исследования атак глушения беспроводных каналов ZigBee, актуальность которых обусловлена ростом числа устройств Интернета вещей и их уязвимостью к воздействиям на физическом уровне. Цель работы заключается в создании аппаратно-программного комплекса, обеспечивающего моделирова-



© 2026 Т.А. Айдынов, С.Т. Тлеубердин, Ж.А. Боранбай, Ф.Б. Тебуева,
А.К. Шайханова, Д.Ж. Сатыбалдина
Бұл жұмыс Creative Commons Attribution 4.0 халықаралық лицензиясы
(CC BY 4.0) бойынша таратылады.
<https://creativecommons.org/licenses/by/4.0/>

учебный стенд,
лабораторный комплекс.

ние радиочастотных помех и анализ их влияния на функционирование сети. В качестве методов использованы экспериментальное моделирование, анализ радиочастотного спектра с применением программно-определяемого радио и мониторинг сетевых параметров. В ходе экспериментов установлено, что при реализации помех происходит снижение доступности устройств до 0–20 %, рост потерь пакетов до 90–100 % и переход системы в состояние отказа обслуживания. Научная новизна состоит в интеграции реальных устройств умного дома и специализированного программного обеспечения в единый образовательный комплекс для контролируемого изучения атак физического уровня. Практическая значимость работы заключается в возможности использования стенда при подготовке специалистов в области кибербезопасности Интернета вещей.

keywords:

smart home, ZigBee,
jamming attack,
cybersecurity, Internet of
things, software-defined
radio, training stand,
laboratory complex.

ABSTRACT

The article is devoted to the development of a smart home educational laboratory stand for investigating ZigBee wireless jamming attacks, the relevance of which is due to the growing number of Internet of Things devices and their vulnerability to impacts at the physical level. The purpose of the work is to create a hardware and software package that provides simulation of radio frequency interference and analysis of their impact on the functioning of the network. The methods used are experimental modeling, analysis of the radio frequency spectrum using software-defined radio and monitoring of network parameters. During the experiments, it was found that when interference occurs, device availability decreases to 0-20%, packet loss increases to 90-100%, and the system goes into a maintenance failure state. The scientific novelty consists in the integration of real smart home devices and specialized software into a single educational complex for the controlled study of physical layer attacks. The practical significance of the work lies in the possibility of using the stand in the training of specialists in the field of cybersecurity of the Internet of Things.

КІРІСПЕ

Ақылды үй технологиялары күнделікті өмірге белсенді түрде енгізіліп, жарықтандыруды, климатты, қауіпсіздік жүйелерін және тұрмыстық техниканы басқаруды автоматтандыруды қамтамасыз етеді. Сарапшылардың болжамы бойынша, 2030 жылға қарай әлемде қосылған IoT құрылғыларының саны 30 миллиардтан асады, олардың едәуір бөлігі тұрғын үй-жайларда пайдаланылатын болады (Zohourian et al., 2023). Қосылған құрылғылардың жылдам өсуі ақпараттық қауіпсіздік саласында жаңа қоңыраулар тудырады, өйткені әрбір құрылғы зиянкестер үшін ықтимал кіру нүктесін білдіреді (Vasilescu et al., 2023).

IEEE 802.15.4 стандартына негізделген ZigBee протоколы ақылды үй сымсыз желілерін ұйымдастырудың ең кең таралған шешімдерінің бірі болып табылады. Ол лицензиясыз 2,4 ГГц жиілік диапазонында жұмыс істейді және соңғы құрылғылардың төмен қуат тұтынуын қамтамасыз етеді, бұл оны батареямен жұмыс істейтін сенсорлар үшін тартымды етеді (IEEE, 2020). Сала сарапшыларының пікірінше, 2023 жылға қарай 500 миллионнан астам ZigBee чипсеті сатылды, ал 2025 жылға қарай болжамды сатылым 3,8 миллиард құрылғыны құрайды (Zigbee Alliance, 2021). Алайда, радиоарнаның ашықтығы әртүрлі шабуылдар үшін алғышарттар жасайды, олардың арасында кедергі шабуылы ерекше қауіп төндіреді, өйткені ол желінің криптографиялық кілттерін білуді қажет етпейді (Kumar et al., 2024).

Кедергі шабуылы мақсатты желінің жұмыс жиілігінде радиожиілік кедергілерін тудырады, бұл құрылғылар арасында деректер алмасудың мүмкін етігіне әкеледі. Ақылды үй контекстінде бұл қауіпсіздік жүйелерінің істен шығуына, атқарушы құрылғыларды бақылауды жоғалтуға және тұрғын үйдің қалыпты жұмысының бұзылуына әкелуі мүмкін. Зерттеулер көрсеткендей, физикалық деңгейдегі шабуылдар сымсыз сенсорлық желілерге үлкен қауіп төндіреді, өйткені оларды жүзеге асыру криптографиялық қорғаныс механизмдерін бұзуды қажет етпейді (Pirayesh et al., 2021).

Киберқауіпсіздік саласындағы заманауи үрдістер мамандарды даярлауға тәжірибеге бағдарланған тәсілдің қажеттілігін көрсетеді. Материалды теориялық зерттеуге негізделген дәстүрлі оқыту әдістері нақты жабдықпен жұмыс істеудің қажетті практикалық дағдыларын қалыптастыруды қамтамасыз етпейді (García-Magariño et al., 2023). IoT Киберқауіпсіздік бойынша білікті мамандарды даярлау тек теориялық білімді ғана емес, сонымен қатар нақты жабдықпен жұмыс істеудің практикалық дағдыларын қажет етеді. Бұл жұмыстың мақсаты бақыланатын білім беру ортасында ZigBee протоколына дыбысты өшіру шабуылдарын зерттеу үшін оқу зертханалық стендін және арнайы бағдарламалық құралды құру болып табылады.

Сымсыз хаттамалардың қауіпсіздігін зерттеу үшін оқу стендтерін әзірлеудің өзектілігі әлемдік тәжірибемен расталады. Атап айтқанда, АҚШ Ұлттық Ғылым қоры интернет заттарының қауіпсіздігі негіздерін үйрету үшін арзан зертханалық кешендер құру жобаларын қаржыландырады (Fu & Zou, 2023). Осындай бастамалар Еуропа мен Азияның жетекші университеттерінде жүзеге асырылуда, бұл осы саладағы мамандарды практикалық даярлаудың маңыздылығын жаһандық мойындауды көрсетеді.

Әдебиеттерге шолу және қолданыстағы шешімдерді талдау

ZigBee протоколының қауіпсіздік мәселелерін халықаралық ғылыми қауымдастық белсенді түрде зерттеп жатыр. Zohourian және т.б. жұмысы OSI моделінің әртүрлі деңгейлеріндегі шабуылдарды талдауды қамтитын Zigbee құрылғыларының осалдығына жан-жақты шолу жасайды (Zohourian et al., 2023). Авторлар кедергі шабуылдарын желінің қол жетімділігі үшін ең маңызды қауіптердің бірі ретінде атап көрсетеді. Kumar, Yadav және Yadav IoT жүйелері контекстінде ZigBee желілерінің қауіпсіздігіне кеңейтілген талдау жүргізіп, радиожиілік шабуылдарынан қорғаудың жаңа әдістерін әзірлеу қажеттілігін атап өтті (Kumar et al., 2024).

Ren және т.б. зерттеуі Zigbee протоколын фаззинг әдісімен жүзеге асырудың қауіпсіздігін талдауға арналған (Ren et al., 2023). Авторлар бастапқы кодқа қол жеткізуді қажет етпестен құрылғылардың бағдарламалық жасақтамасында белгісіз осалдықтарды анықтау мүмкіндігін көрсетеді. Бұл тәсіл студенттердің осалдықтарды іздеу дағдыларын қалыптастыру үшін оқу процесіне біріктірілуі мүмкін.

Сымсыз протоколдардың қауіпсіздігін зерттеу үшін бағдарламалық жасақтамамен анықталған радионы (SDR) қолдану Виллануев пен авторлардың жұмысында қарастырылады (Villanueva-Miranda et al., 2021). Авторлар IOT технологияларына шабуылдарды жүзеге асыру үшін SDR платформаларын пайдаланудың жүйелі шолуын ұсынады, бұл тәсілдің білім беру мақсаттарындағы артықшылықтарын көрсетеді. Bouchendouka, Teguig және Sadoudi USRP платформасы мен GNU radio бағдарламалық жасақтамасын қолдана отырып, IEEE 802.11 протоколына шабуылдардың практикалық орындалуын сипаттайды (Bouchendouka et al., 2022).

IoT киберқауіпсіздігінің білім беру аспектілеріне арналған жұмыстар ерекше қызығушылық тудырады. García-Magariño және т.б. IoT құрылғыларының осалдығын зерттеуге арналған қашықтағы зертхана тұжырымдамасын ұсынады, бұл студенттерге интернет арқылы нақты аппараттық құралдармен тәжірибе жинақтауға мүмкіндік береді (García-Magariño et al., 2023). Fu және Zou өз еңбегінде криптографиялық сопроцессоры бар

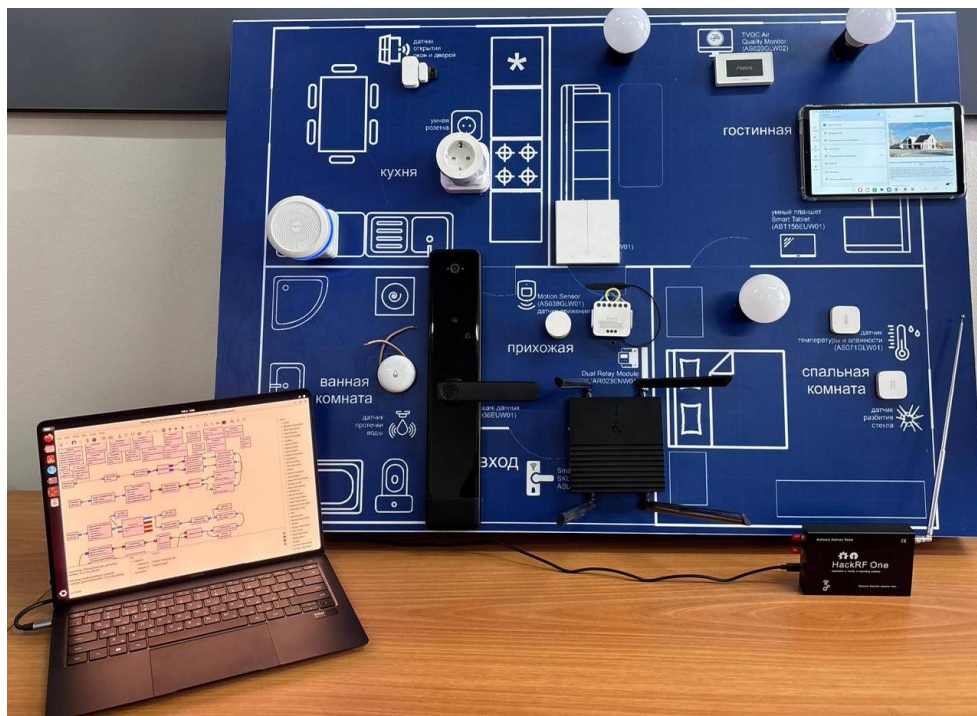
ESP32 микроконтроллеріне негізделген арзан IoT қауіпсіздік оқыту платформасын әзірлеуді сипаттайды (Fu & Zou, 2023).

Қазақстанда IoT киберқауіпсіздік саласындағы зерттеулер дамудың бастапқы кезеңінде тұр. Тлеубердин және Сейтқұлов жергілікті және жаһандық қауіптерді бөліп көрсете отырып, ақылды үй IoT құрылғыларына арналған киберқауіптердің жіктелуін енгізді (Тлеубердин & Сейтқұлов, 2022). Авторлар осы салада білім беру бағдарламаларын дамыту қажеттілігін атап көрсетеді. Қазақстанның 2023-2029 жылдарға арналған цифрлық трансформациясының тұжырымдамасы киберқауіпсіздікті дамудың басым бағыттарының бірі ретінде айқындайды, бұл білім беру процесіне мамандандырылған оқу стендтерін енгізу үшін қолайлы жағдайлар жасайды (Постановление Правительства Республики Казахстан № 269, 2023).

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Ақылды үйді оқыту стенді-Zigbee сымсыз желісі бар смарт үйдің типтік инфрақұрылымын модельдейтін аппараттық және бағдарламалық қамтамасыз ету кешені. Стенд үш функционалды блоктан тұрады: ақылды үйді басқару блогы, сымсыз құрылғылар блогы және радиожилік спектрін талдау блогы. Стендті жобалау кезінде білім беру зертханалық кешендеріне қойылатын талаптар ескерілді: компоненттердің қол жетімділігі, масштабтау мүмкіндігі, эксперименттер жүргізу қауіпсіздігі және IoT жүйелерінің нақты жұмыс жағдайларына сәйкестігі.

Әзірленген оқу стендінің жалпы түрі және зертханалық кешен элементтерінің орналасуы 1-суретте көрсетілген. Суретте ақылды үй контроллері, сымсыз датчиктер, атқарушы құрылғылар, сондай-ақ радиожилік спектрін талдауға арналған жабдықтар көрсетілген. Ұсынылған конфигурация зерттеу жүргізу кезінде қолданылатын эксперименттік қондырғының нақты құрылымын көрсетеді.



1-сурет. ZigBee арналарының кептелу шабуылдарын зерттеуге арналған оқу стендінің жалпы көрінісі

Ескерту – авторлар құрастырған



1-кесте. Оқу стендінің аппараттық құрамы

Компонент	Үлгі	Мақсаты
Ақылды үй контроллері	Aqara Hub	ZigBee құрылғыларын орталықтандырылған басқару
Қозғалыс датчигі	Aqara Motion Sensor	Қамту аймағында қозғалысты анықтау
Ашу датчигі	Aqara Door/Window Sensor	Есіктер мен терезелердің күйін бақылау
Ақылды розетка	Aqara Smart Plug	Аспаптардың электрмен қоректенуін басқару
Температура датчигі	Aqara Temperature Sensor	Микроклимат параметрлерін бақылау
SDR құрылғысы	HackRF One	Радио сигналдарды талдау және генерациялау
Ескерту – автор құрастырған		

Ақылды үйді басқару блогы Aqara Hub-та салынған. Бұл платформа интерфейсі арқылы барлық ақылды үй құрылғыларын орталықтандырылған басқаруды қамтамасыз етеді және автоматтандыру сценарийлерін жасауға мүмкіндік береді. ZigBee құрылғыларын біріктіру үшін Aqara Hub қолданылады, бұл платформаны таңдау оның жаппай қолданылуына, қоғамдастықтың кең қолдауына және нарықтағы көптеген ZigBee құрылғыларымен үйлесімділігіне байланысты (Zigbee2MQTT Documentation, 2024).

Сымсыз құрылғылар блогы нақты ақылды үй инфрақұрылымын модельдеуге мүмкіндік беретін әртүрлі мақсаттағы сенсорлар мен атқарушы құрылғылар жиынтығын қамтиды. Барлық құрылғылар Zigbee протоколымен жұмыс істейді және батареяларды үнемдеу үшін ұйқы режимін қолдайды. Құрылғы түрлерінің әртүрлілігі студенттерге кедергі шабуылдарының әртүрлі жабдық санаттарына әсерін зерттеуге және радиожилік кедергілері жағдайында құрылғылардың әр түрінің жұмыс ерекшеліктерін түсінуге мүмкіндік береді.

Сымсыз құрылғылар блогына сондай-ақ қауіпсіздік пен мониторингке арналған келесі сенсорлар енгізілген:

– Қозғалыс датчигі (Aqara Motion Sensor) – инфрақызыл (PIR) технология негізінде жұмыс істейді және қозғалысты анықтауға арналған. Бұл құрылғы радиожилік кедергілері кезінде оқиға туралы хабарламалардың кешігуін немесе жоғалуын зерттеуге мүмкіндік береді.

– Есік/терезе ашылу датчигі (Aqara Door and Window Sensor) – магниттік байланыс принципі бойынша жұмыс істейді және объектінің ашылуын немесе жабылуын тіркейді. Ол желідегі пакеттердің жоғалуы немесе кешігуі кезінде оқиғалардың дұрыс тіркелуін талдау үшін қолданылады.

– Температура және ылғалдылық датчигі (Aqara Temperature and Humidity Sensor) – қоршаған ортаның параметрлерін өлшейді және периодтық түрде хабқа мәліметтер жібереді. Бұл құрылғы тұрақты телеметриялық трафикке радиожилік кедергілерінің әсерін бағалауға мүмкіндік береді.

– Діріл/шыны сыну датчигі (Aqara Vibration Sensor) – механикалық әсерлерді анықтайды және қауіпсіздік сценарийлерінде қолданылады. Ол қысқа уақыттық импульстік хабарламалардың сенімділігін зерттеуге мүмкіндік береді.

– Ақылды шам (Aqara Smart Bulb) – атқарушы құрылғы ретінде сценарийлердің орындалуын көрсетеді және басқару командаларының жеткізілу уақытын өлшеу үшін пайдаланылады.

– Су ағуын анықтау датчигі (Aqara Water Leak Sensor) – су басу немесе ылғалдың шекті деңгейден жоғарылауын тіркеуге арналған. Бұл құрылғы сирек, бірақ критикалық оқиғаларды тіркейді, сондықтан радиожілік кедергі жағдайында авариялық хабарламалардың жоғалу ықтималдығын бағалау үшін қолданылады.

– Сымсыз екі батырмалы қосқыш (Aqara Wireless Double Switch) – пайдаланушы әрекетін модельдеуге арналған құрылғы. Бұл құрылғы оқиғалық (event-driven) трафикті қалыптастырады және басқару командаларының орындалу сенімділігін зерттеуге мүмкіндік береді.

– Екі арналы құрғақ контактілі реле (Aqara Dual Relay Module, dry contact) – сыртқы атқарушы құрылғыларды басқаруға арналған. Ол нақты инженерлік жүйелермен (жарықтандыру, желдету, сорғылар) интеграцияны модельдеуге мүмкіндік береді және басқару сигналдарының радиожілік кедергілер жағдайындағы тұрақтылығын зерттеу үшін пайдаланылады.

Құрылғылардың әртүрлі типтері (оқиғалық сенсорлар, периодтық сенсорлар және атқарушы құрылғылар) зертханалық стендте кедергі (jamming) шабуылдарының әртүрлі трафик модельдеріне әсерін салыстырмалы түрде талдауға мүмкіндік береді.

Радиожілік спектрін талдау блогы бағдарламалық жасақтамамен анықталған радиодан (SDR) және арнайы бағдарламалық жасақтама орнатылған дербес компьютерден тұрады. HACKRF One SDR құрылғысы 1 МГц-тен 6 ГГц-ке дейінгі жиілік диапазонында радио сигналдарды қабылдауға және жіберуге мүмкіндік береді, бұл оны сымсыз протоколдарды зерттеудің әмбебап құралы етеді (Great Scott Gadgets, 2023). Hackrf One таңдауы оның қол жетімді құнына, ашық архитектурасына, GNU Radio қолдауына және зертханалық жағдайда эксперименттер жүргізу үшін жеткілікті таратқыш қуатына байланысты (Great Scott Gadgets, 2023; GNU Radio Project, 2024).

Оқу стендінің бағдарламалық жасақтамасы

Зертханалық жұмыстарды жүргізу үшін "ZigBee Security Lab" бағдарламалық жасақтамасы арнайы әзірленді, ол студенттерге ZigBee желісінің қауіпсіздігін зерттеудің барлық кезеңдерін дәйекті орындау үшін ыңғайлы интерфейсті ұсынады. Бағдарламалық жасақтама SDR құрылғысын басқару үшін GNU radio кітапханаларын қолдана отырып, Python тілінде жасалған және үш негізгі функционалды модульді жүзеге асырады.

GNU Radio-сигналдарды өңдеу бағдарламаларын құру үшін еркін таратылатын даму ортасы (GNU Radio Project, 2024). Бұл платформа модульдік архитектураның және дайын компоненттердің кең кітапханасының арқасында ғылыми зерттеулер мен білім беру жобаларында кеңінен қолданылады. GNU Radio-ны Python тілімен біріктіру радиожілік сигналдарымен жұмыс істеу үшін икемді және кеңейтілетін қосымшалар жасауға мүмкіндік береді.

ZigBee белсенді арнаны анықтау модулі. ZigBee протоколы 11-ден 26-ға дейін нөмірленген 2,4 ГГц диапазонында 16 арнаны пайдаланады. Әр арнаның орталық жиілігі мына формула бойынша есептеледі: $f = 2405 + 5 \times (n - 11)$ МГц, мұндағы n - арна нөмірі. Анықтау модулі барлық 16 арнаны дәйекті сканерлейді және қабылданған сигнал деңгейін талдайды. Арнада белсенділік анықталған кезде бағдарлама ZigBee пакеттерінің кіріспесін декодтайды және желі идентификаторын (PAN ID) анықтайды. Сканерлеу нәтижелері спектрограмма және анықталған желілер кестесі ретінде көрсетіледі. Сканерлеу алгоритмі анықтаудың жоғары дәлдігін сақтай отырып, белсенді арналарды анықтау уақытын азайту үшін оңтайландырылған.

Желі жұмысының дұрыстығын диагностикалау модулі. Белсенді арнаны анықтағаннан кейін студент Zigbee желісінің трафигін пассивті бақылайтын диагностикалық модульді іске қоса алады. Модуль қабылданған пакеттердің статистикасын, әр құрылғыдан сигнал деңгейін, деректерді беру жиілігін көрсетеді. Бұл шабуыл жасамас бұрын желінің

қалыпты жұмыс істеуін қамтамасыз етуге және кейіннен шабуылдың құрылғылардың жұмысына әсерін бағалауға мүмкіндік береді. Модуль IEEE 802.15.4 пакеттерінің тақырыптарын декодтауды және пайдалы жүктемені декодтаусыз қызметтік ақпаратты алуды жүзеге асырады.

Кедергі шабуылын іске асыру модулі. Бұл модуль таңдалған ZigBee арнасында радиожилілік кедергілерін жасауға мүмкіндік береді. Студент шабуыл параметрлерін: мақсатты арнаны, кедергі жиілігінің диапазонын және қуат деңгейін реттей алады. Бағдарлама кең жолақты шу сигналын қалыптастырады және оны SDR құрылғысы арқылы жібереді. Шабуыл кезінде модуль желіні бақылауды жалғастырады және нақты уақыттағы байланыс сапасының көрсеткіштерінің өзгеруін көрсетеді. Кедергі сигналдарының бірнеше түрі жүзеге асырылады: үздіксіз шу, импульстік кедергі және адаптивті кедергі, бұл әртүрлі шабуыл сценарийлерін зерттеуге мүмкіндік береді.

2-кесте. Бағдарламалық жасақтаманың функционалдығы

Модуль	Функциялар	Параметрлер
Сканерлеу	Белсенді арналарды анықтау PAN ID анықтау	Арналар ауқымы 11-26
Диагностика	Трафикті бақылау, RSSI өлшеу, пакеттерді санау	Бақылау уақыты, жаңарту аралығы
Шабуыл	Кедергі жасау, сигнал түрін таңдау, қуатты реттеу	Арна, жолақ, деңгей, кедергі түрі
<i>Ескерту – автор құрастырған</i>		

Зерттеу жүргізу әдістемесі

Зерттеу әдістемесі эксперименттер жүргізудің бақыланатын шарттарын қолдана отырып, эксперименттік тәсілге негізделген. Нәтижелердің қайталануын қамтамасыз ету үшін дайындық кезеңін, өлшеу кезеңін және нәтижелерді талдау кезеңін қамтитын эксперименттер жүргізудің стандартталған хаттамасы әзірленді.

Дайындық кезеңінде жабдықты калибрлеу, стендтің барлық компоненттерінің жұмысын тексеру және бағдарламалық жасақтама параметрлерін реттеу жүзеге асырылады. SDR құрылғысын калибрлеу қабылдағыштың жергілікті шу деңгейін өлшеуді және таратқыштың пайда болуын анықтауды қамтиды. Бұл параметрлер өлшеу нәтижелерін қалыпқа келтіру үшін қолданылады.

Өлшеу кезеңі зертханалық жұмыстың сценарийіне сәйкес эксперименттерді дәйекті түрде орындауды қамтиды. Барлық өлшеулер сыртқы радиожилілік кедергілерінің әсерін азайту үшін экрандалған бөлмеде жүргізіледі. Өлшеу нәтижелері уақыт белгілері мен эксперимент параметрлері көрсетілген мәліметтер базасына автоматты түрде сақталады.

НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

Зертханалық жұмысты студенттер жеке немесе жұпта орындайды. Жұмыстың ұзақтығы 8 академиялық сағатты құрайды, бұл шабуылдың барлық кезеңдерін егжей-тегжейлі зерттеуге және алынған нәтижелерді талдауға мүмкіндік береді. Жұмыс басталар алдында оқытушы радиожилілік жабдықтарымен жұмыс істеу кезінде қауіпсіздік техникасы бойынша нұсқама жүргізеді және радиожилілік спектрін пайдалану саласындағы заңнаманы сақтау қажеттігін еске салады.

Әзірленген зертханалық стенд негізінде ақылды үй жүйелері мен заттар интернеті хаттамаларының қауіпсіздігі бойынша тәжірибеге бағытталған зертханалық жұмыстарды жүргізу үшін әдістемелік нұсқаулар кешені құрылды. Кешен физикалық, арналық және желілік деңгейлерді зерттеуді, сондай-ақ сымсыз технологиялар мен қолданбалы басқару жүйелерінің осалдықтарын талдауды қамтиды. Зертханалық жұмыстардың құрылымы мен мазмұны 3-кестеде келтірілген.

3-кесте. Smart Home Security оқу стендіндегі зертханалық жұмыстар кешені

№	Зертханалық жұмыстың атауы	Жұмыстың мақсаты және қысқаша сипаттамасы
1	ZigBee протоколдарының стегін зерттеу және мақсатты радиожилілік арнасын анықтау	ZigBee протоколдарының стек архитектурасын, оның Заттар интернетінің (IoT) экожүйесіндегі рөлін және IEEE 802.15.4 стандартының төменгі деңгейлерімен өзара әрекеттесуін егжей-тегжейлі зерттеу. Практикалық бөлім радиожилілік спектрінде және жұмыс жиілігі арнасында Белсенді Zigbee желілерін пассивті анықтауға бағытталған. Бұл тапсырманы орындау үшін қосымша 2.4 ГГц диапазонын сканерлеуге конфигурацияланған мамандандырылған бағдарламалық жасақтама (SDR Hackrf One, CatSniffer) қолданылады. Жұмыс арнасын сәтті анықтау трафикті талдау бойынша барлық зертханалық жұмыстар үшін қажетті алдын ала қадам болып табылады
2	Zigbee трафигін пассивті ұстау және түсіру әдістемесі	Zigbee сымсыз трафигін пассивті ұстау үшін қажетті құралдар мен әдістерді меңгеріңіз. Студенттер №1 жұмыста анықталған жиілік арнасында бақылау режимінде жұмыс істеу үшін ZigBee-Sniffer конфигурациясын жасайды. Деректерді түсіру желілік оқиғаларды инициализациялау кезінде жүзеге асырылады. Нәтиже-деректер массиві. ZigBee (Beacon, Data, Command Frames) сипаттамаларына сәйкес желілік пакеттерді түсіруді және оларды сәйкестендіруді растау үшін желілік анализаторда ұсталған кадрларға негізгі талдау жасалады.
3	Шифрланған желілік кадрлардың құрылымдық және метадеректерді талдауы	ZigBee шифрланған кадрларынан алуға болатын ақпарат көлемін көрсетіңіз. Талдау шифрды ашу процесіне дейін қол жетімді желілік және арналық деңгейлерге бағытталған. Студенттер MAC тақырып өрістерін, желі тақырыбын талдайды. Талдаудың негізгі элементі – қайта ойнату шабуылдарының (Replay Attacks) алдын алу үшін қолданылатын кадр есептегішін (Frame Counter) бақылау. Метадеректерге талдау жасалады: трафиктің жиілігі, көлемі және бағыты, бұл желінің "белсенділік картасын" құруға және ықтимал маңызды түйіндерді анықтауға мүмкіндік береді (мысалы, үйлестіруші).
4	Арнаны өшіру шабуылын эксперименттік зерттеу (Jamming)	Jamming деп аталатын ZigBee протоколының төмен деңгейлерінде (PHY/MAC) қызмет көрсетуден бас тарту шабуылының (DoS) механизмі мен тиімділігін зерттеу. Зертханалық жұмыс бағытталған кедергіні жүзеге асыру принциптерін зерттеуді қамтиды. Бұл тапсырманы орындау үшін арнайы бағдарламалық жасақтама (SDR Hackrf One) қосымша қолданылады, оның көмегімен мақсатты арнада шудың үздіксіз берілуі жүзеге асырылады. Шабуылдың тиімділігі пакеттің жоғалу коэффициентін (Packet Error Rate) және заңды ZigBee желісіндегі байланыстың кешігуін бақылау арқылы сандық түрде өлшенеді. Нәтижелер төмен қуатты сымсыз желілердің осы әсер ету түріне жоғары осалдығын көрсетуі керек.
5	WEP протоколының қауіпсіздігін пароль арқылы талдау	WEP протоколы бар Wi-Fi желілерінің қауіпсіздік ерекшеліктерін зерделеу, мамандандырылған бағдарламалық жасақтаманы (Kali Linux ОЖ (Aircrack-Ng, Hashcat, VirtualBox, Wireshark) қолдана отырып, төрт рет қол алысу (4-way handshake) және парольдерді асыра пайдалану (dictionary/bruteforce) әдістерін меңгеру.

1-кестенің соңы

№	Зертханалық жұмыстың атауы	Жұмыстың мақсаты және қысқаша сипаттамасы
6	WPA-PSK протоколының қауіпсіздігін пароль арқылы талдау	WPA/WPA2-PSK бар Wi-Fi желілерінің қауіпсіздік ерекшеліктерін зерттеу, мамандандырылған бағдарламалық жасақтаманың көмегімен төрт рет қол алысу (4-way handshake) және парольдер арқылы шабуыл жасау (dictionary/bruteforce) әдістерін меңгеру.
7	WPA2-Enterprise протоколының қауіпсіздігін талдау	WPA2-Enterprise протоколы бар Wi-Fi желілерінің қауіпсіздік ерекшеліктерін зерттеу, TLS сертификатын құру әдістерін және төрт рет қол алысуды (4-way handshake) түсіру және арнайы бағдарламалық жасақтаманың көмегімен парольді асыра пайдалану (dictionary/bruteforce) шабуылын орындау үшін шабуыл жасауды меңгеру.
8	Ақылды құлыптау экожүйесінің қауіпсіздігін тексеру	Зерттеу узвисомтси ақылды құлып Физикалық деңгейде, желі деңгейінде және Қолданба деңгейінде. Электрондық басқаруды айналып өту үшін чипті физикалық клондау шабуылдарын, құрылғыны немесе желіні шамадан тыс жүктеу арқылы DDoS шабуылдарын, қолмен ұстауды, Evil Twin типті шабуылдарды (жалған кіру нүктесін құру) және аутентификация шабуылдарын жүзеге асырудың негізгі әдістері мен құралдарын зерттеңіз. Осалдықтарды анықтау мақсатында құлыпты басқару үшін қолданылатын мобильді қосымшаны зерттеу және бұзу

1-кезең. Стендті дайындау және бағдарламалық жасақтаманы іске қосу.

Студент Aqara Hub қуатын қосады және жүйенің жүктелуін күтеді, ол шамамен 2-3 минутты алады. Интерфейс арқылы студент барлық ZigBee құрылғыларының жүйеде көрсетілгенін және «қол жетімді» күйде екенін тексереді. Шабуылдаушының жұмыс станциясында студент SDR құрылғысын қосып, ZigBee Security Lab бағдарламасын іске қосады. Бағдарлама SDR құрылғысын инициализациялайды және оның жұмысының дұрыстығын тексереді. Студент есепте стендтің барлық компоненттерінің күйін және эксперимент жүргізуге дайындығын жазады.



2-сурет. Aqara Hub қосу

Ескерту – авторлар құрастырған

2-кезең. ZigBee желісінің белсенді арнасын анықтау.

ZigBee Security Lab бағдарламасында студент "сканерлеу" модуліне өтіп, белсенді арналарды анықтау процесін бастайды. Бағдарлама барлық 16 ZigBee арналарын дәйекті түрде талдайды және нақты уақыттағы спектрограмманы көрсетеді. Бір арнаны сканерлеу уақыты 500 миллисекундты құрайды, Толық сканерлеу циклі шамамен 8 секундты алады. Сканерлеу аяқталғаннан кейін студент есепте анықталған белсенді арнаның нөмірін, оның орталық жиілігін және Pan id желісінің идентификаторын тіркейді. Студент сонымен қатар электромагниттік жағдайды бағалау үшін көрші арналардағы сигнал деңгейін талдайды.

```
root@elk: /media/sf_codes

ZigBee Channel Activity

2026-02-16 23:29:18

Current  Channel  Activity  Packets  PAN ID
-----
11        0          -
12        0          -
13        0          -
14        0          -
>>> 15      IIII      8      0x1A2B
16        0          -
17        0          -
18        0          -
19        0          -
20        0          -
21        0          -
22        0          -
23        0          -
24        0          -
25        0          -
26        0          -

Active channel: 15 | Simulated PAN ID: 0x1A2B
^C
● Stopped by user.
root@elk: /media/sf_codes#
```

3-сурет. «Сканерлеу» модулі және процесті бастау

Ескерту – авторлар құрастырған

3-кезең. Желі жұмысының дұрыстығын диагностикалау.

Студент «Диагностика» модуліне ауысады және анықталған желіні бақылауды бастайды. 5 минут ішінде бағдарлама желінің жұмысы туралы статистиканы жинайды, жіберілген пакеттердің санын, әр құрылғыдан сигнал күшін және жіберу қателіктерінің болуын тіркейді. Студент сенсорлық белсенділікті бастайды: қозғалыс сенсорының алдынан өтіп, сенсоры бар есікті ашады және жабады, ақылды Розетканы қосады және өшіреді. Барлық оқиғалар басқару жүйесінің интерфейсінде көрсетіліп, Бақылау бағдарламасымен жазылуы керек. Студент есепке бастапқы көрсеткіштерді жазады: қабылданған пакеттер саны, RSSI сигналының орташа деңгейі, сәтті жеткізілген хабарламалардың пайызы. Дұрыс жұмыс істейтін желі үшін әдеттегі RSSI мәндері – 40-тан -70 дБм-ге дейін.

```

root@elk: /media/sf_codes

root@elk:/media/sf_codes# python3 zigbee_scan_final_for_diag.py

=====
ZigBee Network Diagnostics Tool
=====

=====
                        ZIGBEE NETWORK DIAGNOSTICS MODULE
=====

Network: PAN ID 0x1A2B | Channel: 15
Monitoring Duration: 5 minutes
Devices Found: 4

Starting monitoring... Press Ctrl+C to stop early.

```

4-сурет. «Диагностика» модулін іске қосу процесі

Ескерту – авторлар құрастырған

ZIGBEE NETWORK DIAGNOSTICS - LIVE MONITORING					
2026-02-16 23:48:52					
Network: PAN ID 0x1A2B Channel: 15					
Progress: [] 04:49 remaining					
Device	Packets	Errors	RSSI (dBm)	Quality	Delivery %
Motion Sensor	2	0	-49.0	Good	100.0%
Door Sensor	1	0	-48.0	Good	100.0%
Smart Plug	6	0	-46.0	Good	100.0%
Coordinator	8	2	-34.9	Excellent	80.0%
Network Summary:					
Total Packets: 17 Total Errors: 2 Avg RSSI: -40.6 dBm					
Recent Events:					
[23:48:50] Smart Plug Switch turned ON RSSI: -46 dBm					

5-сурет. «Диагностика» модулінің жұмыс процесі

Ескерту – авторлар құрастырған

4-кезең. Кептеліс шабуылын жүзеге асыру.

Студент «шабуыл» модуліне өтіп, кедергі параметрлерін реттейді: бұрын анықталған арнаны таңдайды, 5 МГц кедергі жолағын және ең төменгі қуат деңгейін орнатады. Оқытушы расталғаннан кейін студент кедергілерді тудырады. Бағдарлама шабуыл процесін және нақты уақыттағы желі параметрлерін өзгертуді көрсетеді. Студент RSSI деңгейінің өзгеруін және сәтті қабылданған пакеттер санын бақылайды.

Ақара Hub интерфейсінде студент құрылғылардың "қол жетімді емес" күйіне өтуін бақылайды. Ауысу уақыты есепте жазылады және әдетте жүйенің күту уақытының параметрлеріне байланысты 10-30 секундты құрайды. Студент сенсорларды белсендіру әрекеттерін қайталайды және оқиғалар жүйемен тіркелмейтініне көз жеткізеді. Шабуыл

механизмін тереңірек түсіну үшін студент кедергі қуатын дәйекті түрде арттырады және әр қадамда желі параметрлерінің өзгеруін жазады.

5-кезең. Өнімділікті қалпына келтіру және нәтижелерді талдау.

Студент кедергілердің пайда болуын тоқтатады және құрылғылармен байланыстың қалпына келуін бақылайды. Қалпына келтіру уақыты есепте жазылады және әдетте құрылғы түріне және желі параметрлеріне байланысты 5-60 секундты құрайды. Студент диагностикалық модульді қайта іске қосады және шабуылға дейін және одан кейінгі желі көрсеткіштерін салыстырады. Шабуылдың әртүрлі типтегі құрылғыларға әсері талданады: батареямен жұмыс істейтін сенсорлар, желі үйлестірушісі, маршрутизаторлар. Студент құрылғылардың әр түрі үшін өнімділікті қалпына келтіру ерекшеліктерін жазады.

4-кесте. Эксперименттің күтілетін нәтижелері

Параметрлер	Шабуылға дейін	Шабуыл кезеңінде	Шабуылдан кейін
RSSI деңгейі, дБм	-50...-70	-30...-40	-50...-70
Пакеттер саны / мин	20-50	0-5	15-45
Құрылғылардың қол жетімділігі	100%	0-20 %	80-100 %
Жауап беру уақыты, мс	100-500	Таймаут	200-1000
Тасымалдау қателері, %	0-5	90-100	5-15
<i>Ескерту – автор құрастырған</i>			

Жұмыс нәтижелері бойынша студент орындалған әрекеттердің сипаттамасын, әр кезеңдегі бағдарламаның скриншоттарын, өлшенген параметрлері бар кестені және Zigbee протоколының кедергі шабуылдарына осалдығы туралы қорытындыларды қамтитын есеп жасайды. Есепте сонымен қатар жүйенің осы шабуыл түріне төзімділігін арттыру бойынша ұсыныстар берілген.

Әзірленген оқу стенді ZigBee желілерінің OSI моделінің физикалық деңгейіндегі шабуылдарға осалдығын көрнекі түрде көрсетуге мүмкіндік береді. Студенттер бағдарламалық жасақтамамен анықталған радиоа практикалық тәжірибе алады және сымсыз интернет заттарына радиожілік шабуылдарының принциптерін түсінеді. Эксперименттік зерттеулер зертханалық сабақтарды ұйымдастыруға ұсынылған тәсілдің тиімділігін растады.

Арнайы әзірленген Zigbee Security Lab бағдарламалық құралы зерттеудің барлық кезеңдерін орындау үшін интуитивті интерфейсті қамтамасыз ете отырып, зертханалық жұмысты жүргізуді айтарлықтай жеңілдетеді. Бағдарламаның модульдік архитектурасы оның функционалдығын кеңейтуге және болашақта жаңа шабуыл сценарийлерін қосуға мүмкіндік береді. Оқу процесінде бағдарламаны қолдану тәжірибесі студенттердің бір зертханалық сабақта сымсыз желілер қауіпсіздігінің негізгі тұжырымдамаларын сәтті меңгергенін көрсетті.

Стендтің маңызды артықшылығы-эксперименттің жоғары шынайылығын қамтамасыз ететін нақты ақылды үй құрылғыларын пайдалану. Студенттер үйді автоматтандырудың нақты жүйелерінде қолданылатын сенсорлар мен хаттамалармен жұмыс істейді. Бұл студенттерде IoT жүйелерінің қауіпсіздігінің практикалық аспектілері туралы түсінік қалыптастырады және оларды кәсіби қызметтегі нақты мәселелерді шешуге дайындайды.

Әзірленген стендті қолданыстағы білім беру шешімдерімен салыстырмалы талдау оның бәсекеге қабілеттілігін көрсетеді. Таза бағдарламалық симмуляторлардан айырмашылығы, стенд нақты аппараттық құралдармен және радиожілік сигналдарымен жұмыс істеуді қамтамасыз етеді. Коммерциялық зертханалық кешендермен

салыстырғанда, стенд білім беру мақсаттары үшін қажетті функционалдылықты сақтай отырып, айтарлықтай төмен бағамен сипатталады.

Әрі қарай даму бағыты ретінде Zigbee-ге шабуылдардың басқа түрлерін, соның ішінде ойнату шабуылдары мен желіге қосылу процедураларына шабуылдарды зерттеу үшін бағдарламалық жасақтаманың функционалдығын кеңейту жоспарлануда (Gavra et al., 2023; Ren et al., 2023; Caballero-Gil et al., 2024). Сондай-ақ адаптивті арналарды ауыстыру және кеңістіктік сүзу сияқты кедергі шабуылдарынан қорғау механизмдерін зерттеуге арналған модульдерді әзірлеу жоспарлануда (Fang et al., 2024; Pirayesh et al., 2021). Перспективалық бағыт кедергі шабуылдарын автоматты түрде анықтау үшін машиналық оқыту әдістерін біріктіру (Pirayesh et al., 2021).

ҚОРЫТЫНДЫ

Жұмыста интернет заттарының киберқауіпсіздігі бойынша мамандарды даярлау кезінде оқу процесінде қолдануға болатын ZigBee протоколына дыбысты өшіру шабуылдарын зерттеуге арналған Ақылды үй зертханалық стенді ұсынылған. Ақара Hub smart үй контроллерін, әртүрлі мақсаттағы сымсыз сенсорлар жинағын және hackrf бағдарламалық құралымен анықталған радионы қамтитын стенд аппараттық құралының құрамы сипатталған.

Желінің белсенді арнасын анықтауды, оның жұмыс қабілеттілігін диагностикалауды және кедергі шабуылын жүзеге асыруды қамтамасыз ететін GNU radio платформасы негізінде ZigBee Security Lab мамандандырылған бағдарламалық жасақтамасы әзірленді. Бағдарламалық жасақтама модульдік архитектурамен, интуитивті интерфейспен және шабуылдардың қосымша түрлерін зерттеу үшін функционалдылықты кеңейту мүмкіндігімен сипатталады.

Бес негізгі кезеңді қамтитын зертханалық жұмыстың егжей-тегжейлі сценарийі келтірілген: стендті дайындау, белсенді арнаны анықтау, Желі жұмысын диагностикалау, кедергі шабуылын жүзеге асыру және нәтижелерді талдаумен өнімділікті қалпына келтіру. Эксперименттің күтілетін нәтижелері және зертханалық жұмыстың сәттілігін бағалау критерийлері ұсынылған.

Әзірленген стенд киберқауіпсіздік саласындағы білім беру зертханалық кешендеріне қойылатын заманауи талаптарға сәйкес келеді: ол нақты жабдықтармен жұмыс істеуді қамтамасыз етеді, шабуылдардың практикалық сценарийлерін модельдейді, қол жетімді құны мен масштабтау мүмкіндігімен сипатталады. Стенд Қазақстанның жоғары оқу орындарында ақпараттық қауіпсіздік мамандарын даярлаудың оқу бағдарламаларына біріктірілуі мүмкін.

Әрі қарайғы зерттеулердің бағыттары ретінде желіге қосылу процедурасына қайталанатын шабуылдар мен шабуылдарды зерттеу үшін модульдер әзірлеу, сондай-ақ кедергі шабуылдарын автоматты түрде анықтау үшін машиналық оқыту әдістерін біріктіру жоспарлануда. Студенттердің стенд жабдықтарына қашықтан қол жеткізуі үшін таратылған зертханалық инфрақұрылымды құру перспективалы болып табылады, бұл IoT киберқауіпсіздік саласындағы білім беру бағдарламаларын қамтуды кеңейтуге мүмкіндік береді.

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Автор(лар) мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: «Бұл мақала Қаланың тұрақты дамуы және азаматтардың өмір сүру сапасын жақсарту үшін Smart City цифрлық экожүйесінің интеллектуалды үлгілері мен әдістерін әзірлеу» атты BR24992852 тақырыбы бойынша Ғылым және жоғары білім министрлігінің ғылыми зерттеулерге арналған гранттық қаржыландыру жобасы аясында дайындалды».

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ:

Қолданылмайды.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: Қолданылмайды.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Жұмыс сандық модельдеу мен теориялық талдауға негізделгендіктен, жаңа эксперименттік деректер жиналмаған.

АЛҒЫС БІЛДІРУ: Авторлар әдістемелік қолдау мен пайдалы пікірлері үшін әріптестеріне, сондай-ақ мақаланың сапасын арттыруға септігін тигізген анонимді рецензенттерге алғыс білдіреді.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Қолданылмайды.

ӘДЕБИЕТТЕР ТІЗІМІ

- Zohourian A., Dadkhah S., Neto E.C.P., Mahdikhani H., Danso P.K., Molyneaux H., Ghorbani A.A. (2023). IoT Zigbee device security: A comprehensive review. *Internet of Things*, 22, 100791. <https://doi.org/10.1016/j.iot.2023.100791>
- Vasilescu N.G., Pocatilu P., Doinea M. (2023). IoT Security Challenges for Smart Homes. *Smart Innovation, Systems and Technologies*, 321, 35-48. https://doi.org/10.1007/978-981-19-6755-9_4
- IEEE Standard 802.15.4-2020: Low-Rate Wireless Networks (2020). IEEE, 800 p.
- Zigbee Alliance (2021). Analysts Confirm Half a Billion Zigbee Chipsets Sold. Igniting IoT Innovation. <https://csa-iot.org/>
- Kumar M., Yadav V., Yadav S.P. (2024). Advance comprehensive analysis for Zigbee network-based IoT system security. *Discover Computing*, 27, 22. <https://doi.org/10.1007/s10791-024-09456-3>
- Fang X., Zheng L., Fang X. et al. (2024). Pioneering advanced security solutions for reinforcement learning-based adaptive key rotation in Zigbee networks. *Scientific Reports*. 14, 13931. <https://doi.org/10.1038/s41598-024-64895-8>
- García-Magariño I., Lacuesta R., Rajarajan M., Lloret J. (2023). Exploring IoT Vulnerabilities in a Comprehensive Remote Cybersecurity Laboratory. *Sensors*. 23, 9279. <https://doi.org/10.3390/s23229279>
- Fu X., Zou C. (2023). Building a Low-Cost and State-of-the-Art IoT Security Hands-On Laboratory. NSF Award 1915780. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1915780
- Gavra V., Pop O.A., Dobra I.A. (2023). Comprehensive Analysis: Evaluating Security Characteristics of Xbee Devices against Zigbee Protocol. *Sensors*. 23, 8736. <https://doi.org/10.3390/s23218736>
- Ren M., Ren X., Feng H., Ming J., Lei Y. (2023). Security analysis of ZigBee protocol implementation via device-agnostic fuzzing. *Digital Threats: Research and Practice*, 4, 1-24. <https://doi.org/10.1145/3551894>
- Villanueva-Miranda I., Villalobos-Castaldi F.M., Gress N.H. (2021). A technical review of wireless security for the internet of things: Software defined radio perspective. *Journal of King Saud University – Computer and Information Sciences*, 33, 1050-1061. <https://doi.org/10.1016/j.jksuci.2021.04.003>
- Bouchendouka H., Teguig D., Sadoudi S. (2022). SDR Implementation of WIFI Attacks Using GNURadio. *Lecture Notes in Networks and Systems*, 513, 305-315. https://doi.org/10.1007/978-3-031-12097-8_26
- Benimam A., Sadoudi S., Belmeguenai A. (2023). Software defined radio implementation of a secure waveform for real-time data transmission using universal software radio peripheral. *Security and Privacy*, 6, e280. <https://doi.org/10.1002/spy2.280>
- Тлеубердин, С., & Сейткулов, Е. (2022). Актуальные киберугрозы для IoT устройств умного дома. *Вестник КазАТК*, 122(3), 195–202. <https://doi.org/10.52167/1609-1817-2022-122-3-195->

- 202 // Tleuberdin, S., & Seytkulov, E. (2022). Aktual'nyye kiberugrozy dlya IoT ustroystv umnogo doma [Current cyber threats for IoT smart home devices]. Vestnik KazATK, 122(3), 195–202. <https://doi.org/10.52167/1609-1817-2022-122-3-195-202> (In Russ.)
- Постановление Правительства Республики Казахстан от 28 марта 2023 года № 269. (2023). Концепция цифровой трансформации, развития отрасли информационно-коммуникационных технологий и кибербезопасности на 2023–2029 годы. <https://adilet.zan.kz/rus/docs/P2300000269> // Postanovleniye Pravitel'stva Respubliki Kazakhstan ot 28 marta 2023 goda № 269. (2023). Kontseptsiya tsifrovoy transformatsii, razvitiya otrasli informatsionno-kommunikatsionnykh tekhnologiy i kiberbezopasnosti na 2023–2029 gody [Concept of digital transformation, development of the information and communication technology industry and cybersecurity for 2023–2029]. <https://adilet.zan.kz/rus/docs/P2300000269> (In Russ.)
- Zigbee2MQTT Documentation (2024). <https://www.zigbee2mqtt.io/>
- Great Scott Gadgets. (2023). HackRF One Technical Documentation. <https://greatscottgadgets.com/hackrf/>
- GNU Radio Project. (2024). GNU Radio Manual and C++ API Reference. <https://www.gnuradio.org/>
- Caballero-Gil C., Alvarez R., Hernández-Goya C., Molina-Gil J. (2024). Research on smart-locks cybersecurity and vulnerabilities. Wireless Networks, 30, 5905-5917. <https://doi.org/10.1007/s11276-023-03376-8>
- Pirayesh H., Kheirkhah Sangdeh P., Zeng H. (2021). Securing ZigBee communications against constant jamming attack using neural network. IEEE Internet of Things Journal, 8, 6. 4957-4968. <https://doi.org/10.1109/JIOT.2020.3034219>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Айдынов Төлеген Айдынұлы – докторант, Л.Н. Гумилев атындағы Еуразия Ұлттық университеті, Астана, Қазақстан
Айдынов Төлеген Айдынұлы – докторант, Евразийский национальный университет им. Л.Н. Гумилева, Астана, Казахстан
Aidynov Tolegen – doctoral student, L. N. Gumilyov Eurasian National University, Astana, Kazakhstan,
e-mail: 951005351309@enu.kz,
ORCID: <https://orcid.org/0009-0005-3327-9719>,



Тлеубердин Сакен Тоқатович – докторант, Л.Н. Гумилев атындағы Еуразия Ұлттық университеті, Астана қ., Қазақстан
Тлеубердин Сакен Тоқатович – докторант, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан
Tleuberdin Saken – doctoral student, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
e-mail: sakentleuberdin@gmail.com,
ORCID: <https://orcid.org/0000-0001-9170-9936>



Боранбай Жандос Асылханұлы – докторант, Л.Н. Гумилев атындағы Еуразия Ұлттық университеті, Астана қ., Қазақстан

Боранбай Жандос Асылханұлы – докторант, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

Boranbay Zhandos – doctoral student, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

e-mail: boranbai_zha_3@enu.kz,

ORCID: <https://orcid.org/0009-0009-4096-3876>



Тебуева Фариза Биляловна – физика-математика ғылымдарының докторы, профессор, Солтүстік Кавказ федералды университеті, Ставрополь, Ресей

Тебуева Фариза Биляловна – доктор физико-математических наук, профессор, Северо-Кавказский федеральный университет, Ставрополь, Россия

Tebueva Fariza – Doctor of Physico-Mathematical Sciences, Professor, North-Caucasus Federal University Stavropol, Russian Federation,

e-mail: ftebueva@nfcu.ru,

ORCID: <https://orcid.org/0000-0002-7373-4692>



Шайханова Айгуль Кайрулаевна – қауымдастырылған профессор, Л.Н.Гумилев атындағы Еуразия Ұлттық университеті, Астана, Қазақстан

Шайханова Айгуль Кайрулаевна – ассоциированный профессор, Евразийский национальный университет им. Л. Н. Гумилева, Астана, Казахстан

Shaikhanova Aigul – Associate Professor, L. N. Gumilyov Eurasian National University, Astana, Kazakhstan,

e-mail: shaikhanova_ak@enu.kz,

ORCID: <https://orcid.org/0000-0001-6006-4813>



Сатыбалдина Дина Жагипаровна – физика-математика ғылымдарының кандидаты, профессор, Ақпараттық қауіпсіздік және криптология ғылыми-зерттеу институты директоры, Л.Н. Гумилев атындағы Еуразия Ұлттық университеті, Астана қ., Қазақстан

Сатыбалдина Дина Жагипаровна – кандидат физико-математических наук, профессор, директор НИИ информационной безопасности и криптологии, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

Satybaldina Dina – Candidate of Physico-Mathematical Sciences, Professor, Head of the Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,

e-mail: satybalдина_dzh@enu.kz,

ORCID: <https://orcid.org/0000-0003-0291-4685>

https://doi.org/10.51885/3134-8025_IICS_2026_2_3

SRSTI 28.23.15

BREAST CANCER DETECTION USING LOGISTIC REGRESSION: A COMPARATIVE STUDY OF OPTIMIZATION AND FINE-TUNING METHODS

ЛОГИСТИКАЛЫҚ РЕГРЕССИЯНЫ ҚОЛДАНА ОТЫРЫП, СҮТ БЕЗІ ОБЫРЫН АНЫҚТАУ: ОҢТАЙЛАНДЫРУ ЖӘНЕ ДӘЛ БАПТАУ ӘДІСТЕРІН САЛЫСТЫРМАЛЫ ЗЕРТТЕУ

ВЫЯВЛЕНИЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ С ПОМОЩЬЮ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ: СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ МЕТОДОВ ОПТИМИЗАЦИИ И ТОНКОЙ НАСТРОЙКИ

A.B. Aryn ^{1*}, O.Zh. Mamyrbaev ²¹Caspian University, Almaty, Kazakhstan²Institute of Information and Computing Technologies, Almaty, Kazakhstan*Corresponding author: Aryn Aryn, e-mail: arun_arai@mail.ru

Keywords:

breast cancer, logistic
regression, hyperparameter
tuning, fine-tuning,
optimization methods,
machine learning.

ABSTRACT

Breast cancer is one of the most prevalent and life-threatening diseases worldwide, making accurate early diagnosis critical for improving patient outcomes. This study investigates the effectiveness of logistic regression models for breast cancer detection using the Wisconsin Breast Cancer (Diagnostic) Dataset. Various optimization techniques, including Ordinary Least Squares, Gradient Descent, Newton's Method, L2 Regularization, Stochastic Gradient Descent (SGD), and GridSearch-based fine-tuning, were applied to determine their impact on model performance. Standard evaluation metrics—Accuracy, Precision, Recall, and F1-score—were used to compare methods. The results demonstrate that GridSearch-based fine-tuning consistently yields the highest overall performance, highlighting the importance of systematic hyperparameter optimization in enhancing model robustness and predictive accuracy. These findings emphasize that even classical models like logistic regression can achieve state-of-the-art results in medical data analysis when combined with appropriate optimization strategies.

Түйінді сөздер:

сүт безінің қатерлі ісігі,
логистикалық регрессия,
гиперпараметрлерді
баптау, дәл баптау,
оңтайландыру әдістері,
машиналық оқыту.

ТҮЙІНДЕМЕ

Сүт безі қатерлі ісігі бүкіл әлемде ең көп таралған және өмірге қауіп төндіретін аурулардың бірі болып табылады, бұл пациенттердің нәтижелерін жақсарту үшін дәл ерте диагнозды өте маңызды етеді. Бұл зерттеу Висконсиндегі Сүт безі Қатерлі ісігінің Диагностикалық Деректер Жинағын (WBCD) пайдалана отырып, сүт безі обырын анықтау үшін логистикалық регрессия үлгілерінің тиімділігін зерттейді. Оңтайландырудың әртүрлі әдістері, соның ішінде қарапайым ең кіші квадраттар әдісі, градиенттік түсу, Ньютон әдісі, L2-регуляризациясы, стохастикалық градиенттің түсуі (SGD) және



© 2026 A.B. Aryn, O.Zh. Mamyrbaev

This work is licensed under a Creative Commons Attribution 4.0
International License (CC BY 4.0).<https://creativecommons.org/licenses/by/4.0/>

олардың модель өнімділігіне әсерін анықтау үшін GridSearch-ге негізделген дәл баптау қолданылды. Әдістерді салыстыру үшін бағалаудың стандартты көрсеткіштері – дәлдік, толықтық, сезімталдылық және F1-көрсеткіші қолданылды. Нәтижелер GridSearch-ге негізделген дәл баптау модельдердің сенімділігі мен болжау дәлдігін арттыру үшін гиперпараметрлерді жүйелі оңтайландырудың маңыздылығын көрсете отырып, ең жоғары жалпы өнімділікті дәйекті түрде қамтамасыз ететінін көрсетеді. Бұл нәтижелер логистикалық регрессия сияқты классикалық модельдерде тиісті оңтайландыру стратегияларымен үйлескенде медициналық деректерді талдауда заманауи нәтижелерге қол жеткізе алатынын көрсетеді.

Ключевые слова:

рак молочной железы,
логистическая регрессия,
настройка
гиперпараметров, тонкая
настройка, методы
оптимизации, машинное
обучение.

АННОТАЦИЯ

Рак молочной железы является одним из наиболее распространенных и опасных для жизни заболеваний во всем мире, что делает точную раннюю диагностику критически важной для улучшения результатов лечения пациентов. В этом исследовании изучается эффективность моделей логистической регрессии для выявления рака молочной железы с использованием набора данных по диагностике рака молочной железы штата Висконсин (WBCD). Для определения их влияния на производительность модели были применены различные методы оптимизации, включая обычные методы наименьших квадратов, градиентный спуск, метод Ньютона, регуляризацию L2, стохастический градиентный спуск (SGD) и тонкую настройку на основе GridSearch. Для сравнения методов использовались стандартные показатели оценки – точность, полнота, чувствительность и показатель F1. Результаты демонстрируют, что точная настройка на основе GridSearch неизменно обеспечивает высочайшую общую производительность, подчеркивая важность систематической оптимизации гиперпараметров для повышения надежности модели и точности прогнозирования. Эти результаты подчеркивают, что даже классические модели, такие как логистическая регрессия, могут обеспечить самые современные результаты анализа медицинских данных в сочетании с соответствующими стратегиями оптимизации.

INTRODUCTION

Breast cancer remains one of the most prevalent and life-threatening diseases worldwide, necessitating accurate and timely diagnostic tools (Ghasemi et al., 2024; Дёмин & Гермашев, 2024). Traditional machine learning approaches, particularly logistic regression, have been widely employed for binary classification tasks due to their simplicity, interpretability, and efficiency (Aryn et al., 2025; Watrianthos et al., 2025). However, the predictive performance of logistic regression models can be limited if parameters are not carefully optimized or thresholds are not properly calibrated (Benghazouani et al., 2025; Won et al., 2025). Watrianthos et al. (2025) proposed a threshold-optimized and calibrated logistic regression approach, demonstrating that systematic adjustment of model parameters and decision thresholds can significantly improve classification accuracy in breast cancer detection. This approach underscores the potential benefits of combining classical statistical models with fine-tuning techniques to enhance predictive reliability in clinical applications (Watrianthos et al., 2025).

Recent advances in hyperparameter optimization (HPO) have highlighted the critical role of efficient tuning in improving machine learning model performance (Won et al., 2025; Ahmad et al., 2026). Won et al. (2025) provide a comprehensive review of multi-fidelity HPO, emphasizing approaches that leverage low-fidelity evaluations to reduce computational costs while maintaining effective model optimization. Their study categorizes existing HPO methods

according to underlying strategies, discusses the benefits of considering multiple fidelity levels, and identifies practical challenges that remain unaddressed. These insights underscore the value of systematic hyperparameter tuning in machine learning models, particularly when optimizing performance metrics such as accuracy, precision, recall, and F1-score (Won et al., 2025; Benghazouani et al., 2025; Ahmad et al., 2026).

Benghazouani et al. (2025) investigated the enhancement of logistic regression models through hybrid nature-inspired optimization techniques, combining genetic algorithms and swarm intelligence to optimize model weights and improve feature selection. Their proposed method, EGASI-RL, demonstrated high predictive performance on breast and lung cancer datasets, achieving substantial improvements in accuracy, precision, recall, and F1-score compared to standard logistic regression. The study highlights the potential of integrating global search and adaptive optimization strategies to overcome limitations of weight estimation in high-dimensional or noisy datasets. These findings reinforce the importance of systematic optimization of logistic regression parameters and feature selection, providing a conceptual foundation for applying fine-tuning techniques such as GridSearch and regularization in predictive modeling tasks.

Ghasemi et al. (2024) conducted a systematic scoping review on the application of explainable artificial intelligence (XAI) in breast cancer detection and risk prediction. Their study emphasizes the importance of interpretability and transparency in AI models used in medical settings, highlighting techniques such as SHapley Additive exPlanations (SHAP) for explaining predictions of complex models. While their focus is primarily on tree-based ensemble models, the findings underscore a broader principle relevant to any predictive modeling, including logistic regression: optimizing model performance through careful tuning of parameters and weights should be complemented by strategies that maintain model interpretability and trustworthiness. This perspective supports the rationale for systematically evaluating different optimization methods, regularization, and fine-tuning approaches to improve both predictive accuracy and reliability in breast cancer classification tasks.

Demin and Germashev (2024) review the application of machine learning techniques in breast cancer diagnosis, highlighting their utility in supporting diagnostic tasks, patient state assessment, and risk prediction. The study emphasizes that classical machine learning methods, such as support vector machines, random forests, and convolutional neural networks, have been effectively employed for tasks ranging from image segmentation in mammography to predictive modeling based on clinical and genomic data. The review demonstrates that these approaches can improve diagnostic accuracy, efficiency, and reliability, and underscores the importance of selecting appropriate features and optimizing model parameters. Such findings provide a foundation for systematically applying logistic regression and exploring optimization techniques, including weight tuning, regularization, and fine-tuning, to enhance predictive performance in breast cancer detection tasks.

Cristea et al. (2025) investigated optimized logistic regression models for analyzing complex gastroenterological images, comparing their performance against state-of-the-art approaches. The study emphasizes that careful tuning of model parameters can significantly enhance predictive accuracy, even when using relatively simple models like logistic regression, highlighting the continued relevance of classical methods when optimized properly. Similarly, Khan et al. (2026) explored hybrid machine learning techniques for breast cancer prediction, combining multiple algorithms and optimization strategies to maximize model performance across various metrics. Both studies converge on the insight that systematic parameter optimization and hybridization strategies can substantially improve the robustness and accuracy of predictive models in medical applications, demonstrating the importance of fine-tuning, feature selection, and integration of complementary methods in high-stakes diagnostic tasks.

Recent studies have demonstrated that systematic hyperparameter tuning can substantially enhance the performance of logistic regression models across diverse datasets (Ahmad et al., 2026). By carefully adjusting parameters such as regularization strength, solver type, and iteration limits, predictive accuracy and robustness can be significantly improved. Inspired by these findings, this study explores the application of logistic regression for breast cancer detection, incorporating multiple optimization techniques and fine-tuning strategies to maximize model performance and reliability (Ahmad et al., 2026). This study provides a comprehensive evaluation of logistic regression models for breast cancer diagnosis, highlighting the benefits of systematic optimization and fine-tuning (Watrianthos et al., 2025; Benghazouani et al., 2025; Ahmad et al., 2026). The findings aim to inform future applications of logistic regression in medical data analysis, demonstrating how classical models can achieve high predictive performance when coupled with appropriate optimization strategies (Cristea et al., 2025; Khan et al., 2026).

MATERIALS AND METHODS

Dataset Description

The Breast Cancer Wisconsin (Diagnostic) Dataset serves as a standard benchmark in medical machine learning research. It comprises 569 samples derived from fine needle aspiration (FNA) of breast masses. Due to its structured feature space and moderate dimensionality, the dataset is highly suitable for the evaluation of linear and probabilistic classifiers, including logistic regression. Furthermore, the clinical interpretability of its features facilitates a rigorous comparative analysis of optimization strategies and regularization techniques.

The structural attributes of the dataset, encompassing 30 clinical features alongside identification and diagnostic labels, are detailed in Table 1. This summary includes the non-null count and the data representation format for each attribute, ensuring data integrity for subsequent preprocessing.

Table 1. Detailed dataset attributes

№	Column Name	Data Type
0	id	int64
1	diagnosis	object
2	radius_mean	float64
3	texture_mean	float64
4	perimeter_mean	float64
5	area_mean	float64
6	smoothness_mean	float64
7	compactness_mean	float64
8	concavity_mean	float64
9	concave points_mean	float64
10	symmetry_mean	float64
11	fractal_dimension_mean	float64
12	radius_se	float64
13	texture_se	float64
14	perimeter_se	float64
15	area_se	float64
16	smoothness_se	float64
17	compactness_se	float64
18	concavity_se	float64



End of Table 1

No	Column Name	Data Type
19	concave points_se	float64
20	symmetry_se	float64
21	fractal_dimension_se	float64
22	radius_worst	float64
23	texture_worst	float64
24	perimeter_worst	float64
25	area_worst	float64
26	smoothness_worst	float64
27	compactness_worst	float64
28	concavity_worst	float64
29	concave points_worst	float64
30	symmetry_worst	float64
31	fractal_dimension_worst	float64
32	Unnamed: 32	float64

The dataset defines a binary classification task where samples are categorized as either malignant (M) or benign (B). The feature space comprises 30 real-valued numerical attributes representing the morphological characteristics of cell nuclei derived from digitized images. These attributes—including radius, texture, perimeter, area, smoothness, compactness, concavity, symmetry, and fractal dimension—are quantified using three statistical measures: the mean, standard error (SE), and the "worst" (mean of the three largest values) measurement.

During the preprocessing phase, the unique identifier and an auxiliary null column were excluded from the feature set, as they provide no predictive utility. To ensure compatibility with the learning algorithms, the categorical target labels were numerically encoded.

The class distribution exhibits a moderate imbalance, as illustrated in the sample breakdown: 357 observations (62.7%) are classified as benign, while 212 observations (37.3%) are malignant. Although the dataset is not perfectly balanced, the representation of the minority class is statistically sufficient to support robust model training and evaluation without the requirement for synthetic resampling or data augmentation techniques.

Data Preprocessing and Feature Standardization

Prior to model training, the dataset was divided into feature variables and a target variable. All numerical attributes were assigned to the feature matrix X , while the diagnostic outcome (diagnosis) was selected as the target vector y .

The target variable was encoded into numerical format to represent a binary classification problem, where one class corresponds to benign tumors and the other to malignant tumors [9].

The dataset was then split into training and testing subsets in order to evaluate model generalization performance on unseen data. The training set was used for model fitting, while the test set was reserved exclusively for performance evaluation.

Feature standardization was applied to ensure that all input variables contribute equally to the learning process. Since the dataset contains features with different physical scales and magnitudes, standardization is particularly important for optimization-based models such as logistic regression, which rely on gradient-based and second-order optimization methods.

Standardization was performed using z-score normalization, where each feature was transformed by subtracting its mean and dividing by its standard deviation, computed from the training data:

$$X_{std} = \frac{X - \mu}{\sigma}$$

where μ and σ denote the mean and standard deviation of each feature in the training set, respectively. The same transformation parameters were applied to the test set to prevent data leakage.

Models and optimization methods

In this study, a baseline logistic regression model was employed as the primary classification algorithm for breast cancer detection. Logistic regression is a widely used probabilistic model for binary classification tasks and serves as a strong baseline due to its simplicity, interpretability, and solid theoretical foundation.

The model estimates the probability of class membership by applying a sigmoid function to a linear combination of input features. During training, the model learns a set of weight parameters that maximize the likelihood of observing the given class labels.

In the baseline configuration, logistic regression was applied without explicit regularization or hyperparameter tuning, allowing the model to learn weights solely based on the training data. The standardized feature set was used as input to ensure numerical stability and efficient convergence during optimization.

The logistic regression model was trained using the standardized training dataset. After model fitting, predictions were generated for the test dataset to assess classification performance on previously unseen data.

The performance of the logistic regression model was evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score. These metrics were computed on the test dataset to assess the predictive performance of the model.

To investigate the effect of different optimization strategies on logistic regression performance, several methods for estimating model weights were applied and evaluated. Each method was trained on the standardized training dataset, and predictions were generated for the test dataset. Model performance was assessed using the same evaluation metrics for consistency.

Ordinary Least Squares (OLS)

As an initial optimization approach, the Ordinary Least Squares (OLS) method was applied to estimate model parameters. Although OLS is traditionally used for linear regression, it was employed in this study as a baseline optimization technique to examine its behavior in a binary classification setting.

After training, predictions were obtained on the test dataset, and classification performance metrics were computed.

Gradient Descent

The Gradient Descent optimization method was used to iteratively update the model weights by minimizing the loss function with respect to the training data. This approach relies on the first-order derivative of the loss function and updates parameters in the direction of the negative gradient.

The optimized model was then evaluated on the test dataset using the selected performance metrics.

Newton Method

The Newton Method was applied as a second-order optimization technique that utilizes both first and second derivatives of the loss function. By incorporating curvature information through the Hessian matrix, this method aims to achieve faster convergence compared to first-order methods.

Model predictions were generated for the test set, and performance metrics were subsequently calculated.

Regularization (L2)

To control model complexity and reduce the risk of overfitting, L2 regularization was applied to the logistic regression model. This technique introduces a penalty term proportional to the squared magnitude of the model weights into the optimization objective.

The resulting predictions were evaluated using the same set of classification metrics.

Stochastic Gradient Descent (SGD)

In addition to batch optimization methods, Stochastic Gradient Descent (SGD) was employed as an alternative optimization strategy. Unlike standard gradient descent, SGD updates model parameters using individual samples or small batches, which can improve computational efficiency and scalability for larger datasets.

Fine-tuning

In the final stage of the study, fine-tuning was applied to the logistic regression model in order to improve its predictive performance. Fine-tuning refers to the systematic optimization of model hyperparameters that are not directly learned during training but significantly influence model behavior and generalization ability.

In this work, fine-tuning was performed using GridSearchCV, which conducts an exhaustive search over a predefined grid of hyperparameter values combined with cross-validation.

The following hyperparameters of the logistic regression model were tuned:

- Regularization strength (C): values ranging from 0.01 to 100 were evaluated. This parameter controls the trade-off between fitting the training data and model complexity. Smaller values of C correspond to stronger regularization, while larger values allow the model to fit the data more closely.

- Optimization solver: three solvers (lbfgs, newton-cg, and saga) were considered. These solvers implement different numerical optimization strategies for estimating model parameters and may affect convergence behavior and performance.

- Penalty type: L2 regularization was selected to ensure numerical stability and compatibility with all chosen solvers.

A 5-fold cross-validation strategy was employed during the grid search to ensure robust estimation of model performance across different training subsets. The F1-score was used as the optimization criterion, as it provides a balanced measure of precision and recall and is well suited for binary classification problems with moderate class imbalance.

RESULTS AND DISCUSSION

This section presents the experimental results obtained using logistic regression models and different optimization methods.

To evaluate and compare the performance of the models, the following classification metrics were used: Accuracy, Precision, Recall, and F1-score.

The results obtained for logistic regression models trained using various weight optimization techniques are summarized in Table 2. These techniques include the method of least squares (OLS), gradient descent (GD), Newton's method, L2 regularization, stochastic gradient descent (SGD), as well as the baseline logistic regression model.

The OLS and gradient descent-based models demonstrated identical performance, achieving an accuracy of 0.9649 and an F1-score of 0.95. Newton's method resulted in lower classification performance, with an accuracy of 0.9035 and an F1-score of 0.8642.

The logistic regression model with L2 regularization achieved an accuracy of 0.9649, a precision of 0.975, and an F1-score of 0.9512, matching the performance of the baseline logistic regression model. The stochastic gradient descent-based model showed comparable results, with an accuracy of 0.9561 and an F1-score of 0.9398.

The quantitative results for all optimization methods are summarized in Table 2.

For better visualization, Figure 1 presents a graphical comparison of the optimization methods based on the selected evaluation metrics.

Table 2. Performance comparison of logistic regression models using different weight optimization methods

index	Accuracy	Precision	Recall	F1-score
OLS	0.9649122807017544	1.0	0.9047619047619048	0.95
GD	0.9649122807017544	1.0	0.9047619047619048	0.95
Newton	0.9035087719298246	0.8974358974358975	0.8333333333333334	0.8641975308641975
L2 Regularization	0.9649122807017544	0.975	0.9285714285714286	0.9512195121951219
SGD	0.956140350877193	0.9512195121951219	0.9285714285714286	0.9397590361445783
Logistic Regression (baseline)	0.9649122807017544	0.975	0.9285714285714286	0.9512195121951219

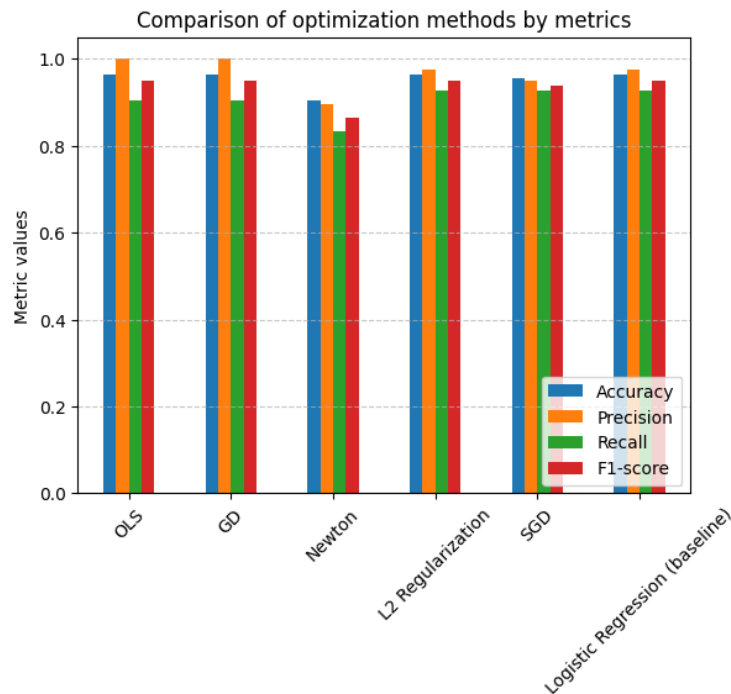


Figure 1. A visual comparison of the evaluated optimization methods across the selected performance metrics

Note – compiled by the authors

To improve the performance of the logistic regression model, a fine-tuning procedure was applied using grid search with cross-validation. The optimization was performed by varying key hyperparameters of the logistic regression model, including the regularization strength, optimization solver, and regularization type.

The base model was defined as follows:

maximum number of iterations: max_iter = 1000

random seed: random_state = 42

The hyperparameter search space included:

C (inverse regularization strength): {0.01, 0.1, 1, 10, 100}

solver: {lbfgs, newton-cg, saga}

penalty: L2 regularization

A 5-fold cross-validation strategy was employed, and the F1-score was selected as the optimization criterion. The grid search was conducted on the standardized training dataset, and the best hyperparameter configuration was selected based on the cross-validated performance.

After identifying the optimal hyperparameters, the tuned model was used to generate predictions on the test dataset. Figure 2 shows the performance metrics of the baseline and GridSearch-optimized logistic regression models.

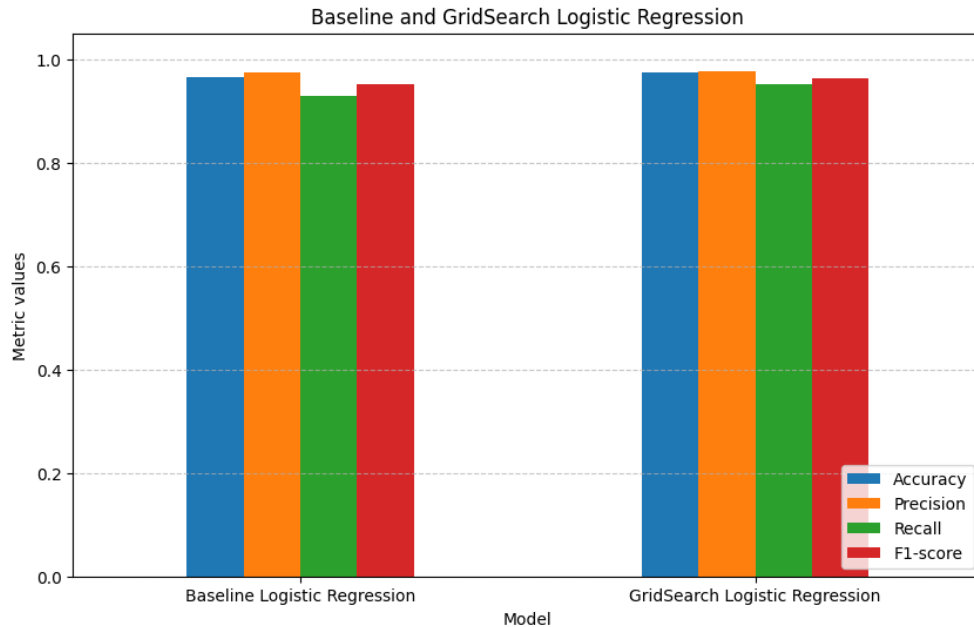


Figure 2. Comparison of baseline and fine-tuned logistic regression models

Note – compiled by the authors

To provide a rigorous comparative assessment of the implemented optimization paradigms, the performance metrics for each methodology were aggregated and analyzed. The experimental framework evaluated the following optimization approaches: Ordinary Least Squares (OLS), Gradient Descent (GD), Newton's Method, Baseline Logistic Regression (LR), LR with L2 Regularization, Stochastic Gradient Descent (SGD), and Fine-tuned LR via GridSearchCV. The comparative performance results are tabulated in Table 3, which details the Accuracy, Precision, Recall, and F1-score achieved by each algorithm on the sequestered test dataset. These metrics provide a multidimensional view of the classifiers' ability to generalize to unseen clinical data.

Table 3. Performance comparison of optimization methods for breast cancer detection

Methods	Accuracy	Precision	Recall	F1-score
OLS	0.9649122807017544	1.0	0.9047619047619048	0.95
Gradient Descent	0.9649122807017544	1.0	0.9047619047619048	0.95
Newton Method	0.9035087719298246	0.8974358974358975	0.8333333333333334	0.8641975308641975
Logistic Regression (Baseline)	0.9649122807017544	0.975	0.9285714285714286	0.9512195121951219
Logistic Regression + L2	0.9649122807017544	0.975	0.9285714285714286	0.9512195121951219
Logistic Regression + SGD	0.956140350877193	0.9512195121951219	0.9285714285714286	0.9397590361445783

End of Table 1

Methods	Accuracy	Precision	Recall	F1-score
Logistic Regression + GridSearch	0.9736842105263158	0.975609756097561	0.9523809523809523	0.963855421686747
<i>Note – compiled by the authors</i>				

Figure 3 illustrates a comparative visualization of the evaluation metrics obtained for all optimization methods considered in this study.

The chart enables a clear comparison of model performance across different optimization strategies using Accuracy, Precision, Recall, and F1-score.

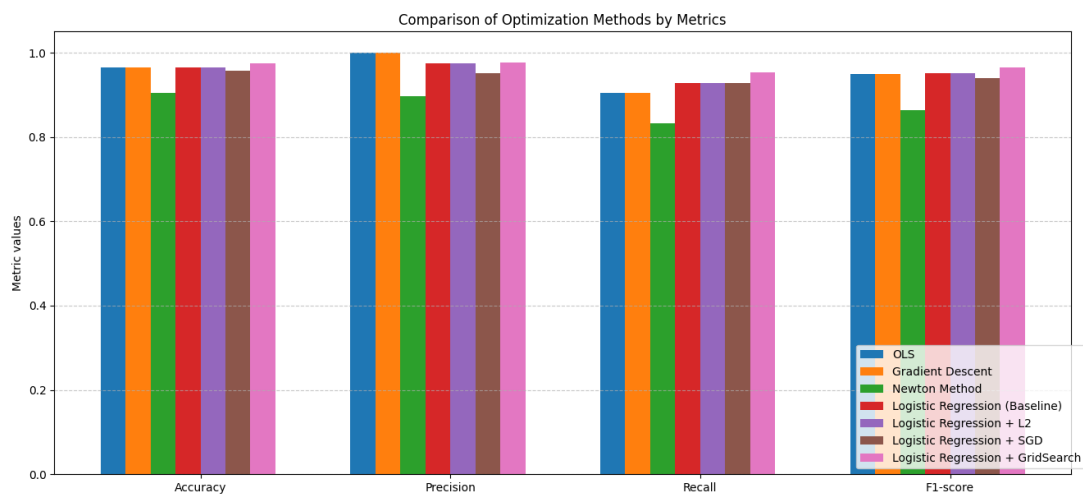


Figure 3. Comparison of Optimization Methods for Breast Cancer Detection

Note – compiled by the authors

Fine-tuning via GridSearch yielded the best overall performance, achieving the highest values across all evaluation metrics (Accuracy = 0.9737, Recall = 0.9524, F1-score = 0.9639), which highlights the importance of systematic hyperparameter optimization.

Ordinary Least Squares (OLS) and Gradient Descent demonstrated strong and stable performance, both achieving Accuracy = 0.9649 and F1-score = 0.95, with perfect precision (Precision = 1.0) and slightly lower recall (Recall = 0.9048), making them comparable to the baseline logistic regression model.

Newton's Method was the least effective approach in this study, resulting in noticeably lower performance (Accuracy = 0.9035, Recall = 0.8333, F1-score = 0.8642) compared to other optimization techniques.

L2 regularization and stochastic gradient descent (SGD) improved model robustness and maintained competitive performance; however, their results (F1-score = 0.9512 for L2 and 0.9398 for SGD) did not surpass those obtained with fine-tuned logistic regression.

Overall, logistic regression proved to be a robust and effective model for breast cancer detection, particularly when combined with appropriate optimization strategies and fine-tuning, consistently achieving high accuracy and balanced precision–recall trade-offs.

CONCLUSION

This study investigated the effectiveness of logistic regression for breast cancer detection using the Breast Cancer Wisconsin (Diagnostic) dataset, with a particular focus on optimization techniques and fine-tuning strategies for model weights.

A baseline logistic regression model was first implemented and evaluated, demonstrating strong classification performance in terms of accuracy, precision, recall, and F1-score. To further analyze the impact of different optimization strategies, several methods for weight estimation were applied, including ordinary least squares, gradient descent, Newton's method, L2 regularization, and stochastic gradient descent. The experimental results showed that optimization methods such as ordinary least squares and gradient descent achieved comparable performance to the baseline logistic regression model, while Newton's method exhibited relatively lower performance, indicating potential sensitivity to data characteristics or numerical stability in this setting.

Regularization techniques and stochastic optimization provided stable results and helped control model complexity; however, they did not significantly outperform the baseline model in terms of overall classification metrics. The most notable improvement was achieved through fine-tuning using GridSearch with cross-validation, which optimized key hyperparameters of the logistic regression model. This approach resulted in the highest overall performance, particularly in terms of accuracy and F1-score, demonstrating the importance of systematic hyperparameter optimization.

Overall, the findings confirm that logistic regression remains a strong and reliable baseline for breast cancer detection, especially when combined with appropriate data preprocessing and fine-tuning techniques. The comparative analysis highlights that while different optimization methods may lead to similar performance levels, hyperparameter tuning plays a crucial role in achieving optimal classification results. These results emphasize the value of combining classical statistical models with modern optimization and validation techniques in medical data analysis. Hypotheses need to be tested, or which applied aspects may be implemented in the future.

CONFLICT OF INTEREST: The authors declare that there is no conflict of interest regarding the publication of this article.

FUNDING: Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT: Not applicable.

INFORMED CONSENT STATEMENT: Not applicable. The study did not involve human participants.

DATA AVAILABILITY STATEMENT: The data used in this study are publicly available. This research is based on the Wisconsin Breast Cancer (Diagnostic) Dataset, which is available in the public domain and can be accessed through the Machine Learning Repository of the University of California, Irvine (UCI):

[https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic))

ACKNOWLEDGEMENTS: Not applicable.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: The authors did not use generative artificial intelligence technologies in the preparation of this manuscript. All analyses, interpretations, and writing were performed by the authors.

REFERENCES

- Watrianthos, R., Rayendra, R., Asri, E., Yuhefizar, Y., & Humaira, H. (2025). Threshold-optimized and calibrated logistic regression for breast cancer classification. *Salud, Ciencia y Tecnología*, 5, 8. <https://doi.org/10.56294/saludcyt20252241>
- Won, J., Lee, H. S., & Lee, J. W. (2025). A review on multi-fidelity hyperparameter optimization in machine learning. *ICT Express*, 11(2), 245–257. <https://doi.org/10.1016/j.ict.2025.02.008>
- Benghazouani, S., Nouh, S., & Zakrani, A. (2025). A robust logistic regression approach enhanced by hyperparameter optimization techniques through swarm intelligence and genetic

- algorithms: Advancing cancer diagnosis. High-Confidence Computing, 100378. <https://doi.org/10.1016/j.hcc.2025.100378>
- Ghasemi, A., Hashtarkhani, S., Schwartz, D. L., & Shaban-Nejad, A. (2024). Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. Cancer Innovation, 3(5), e136. <https://doi.org/10.1002/cai2.136>
- Дёмин, К. С., & Гермашев, И. В. (2024). Использование методов машинного обучения при диагностике рака молочной железы // Demin, K. S., & Germashev, I. V. Ispol'zovanie metodov mashinnogo obucheniya pri diagnostike raka molochnoy zhelezy [The use of machine learning methods in breast cancer diagnosis]. Digital Diagnostics, 5(3), 578. <https://doi.org/10.17816/DD625866> (In Russ.)
- Cristea, D. M., Sima, I., & Iantovics, L. B. (2025). Comparative analysis of optimized logistic regression with state-of-the-art models for complex gastroenterological image analysis. Frontiers in Medicine, 12, 1655612. <https://doi.org/10.3389/fmed.2025.1655612>
- Khan, M. Z., Salahuddin, M., Khan, A. A., Sultan, S. A., Bhatti, S. D., Ullah, S., & Mohamed, M. A. (2026). Innovative hybrid machine learning methods for breast cancer predictive optimization. International Journal of Computational Intelligence Systems. <https://doi.org/10.1007/s44196-025-01135-5>
- Ahmad, M., Javed, N., Muzafar, A., Muzafar, M., Naseer, H., Huang, G., & He, D. (2026). Enhancing logistic regression performance through hyperparameter tuning: A comparative evaluation across datasets. <https://doi.org/10.21203/rs.3.rs-8304042/v1>
- Aryn, A., Мамырбаев, О., & Pavlov, S. (2025). Классификация рака молочной железы, используя методы машинного обучения // Aryn, A., Mamyrbaev, O., & Pavlov, S. Klassifikatsiya raka molochnoy zhelezy, ispol'zuya metody mashinnogo obucheniya [Breast cancer classification using machine learning methods]. Nauka intellektual'nykh sistem, 1(1), 25–37. <https://kazscience.kz/index.php/SIS/article/view/14/3> (In Russ.)

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Арын Арай Болатқызы – техника ғылымдарының магистрі, Каспий қоғамдық университеті, Алматы қ., Қазақстан
Арын Арай Болатқызы – магистр технических наук, Каспийский университет, г. Алматы, Казахстан
Aryn Aray Bolatkyzy – Master of Technical Sciences, Caspian University, Almaty, Kazakhstan
e-mail: arun_arai@mail.ru,
ORCID: <https://orcid.org/0000-0001-7023-0424>



Мамырбаев Өркен Жұмажанұлы – PhD, Ақпараттық және есептеу технологиялары институты, Алматы, Қазақстан
Мамырбаев Өркен Жұмажанович – PhD, Института информационных и вычислительных технологий, г. Алматы, Казахстан
Mamyrbayev Orken Zhumazhanovich – PhD, Institute of Information and Computing Technologies, Almaty, Kazakhstan
e-mail: morkenj@mail.ru,
ORCID: <https://orcid.org/0000-0001-8318-3794>

https://doi.org/10.51885/3134-8025_IICS_2026_2_4

XFTAP 14.35.09

БІЛІМ БЕРУДІҢ ЦИФРЛЫҚ ТРАНСФОРМАЦИЯСЫ ЖАҒДАЙЫНДА ОҚЫТУШЫЛАРДЫҢ БІЛІКТІЛІГІН АРТТЫРУ ЖҮЙЕСІНДЕ STEM-ӘДІСТЕРДІ ОНТОЛОГИЯЛЫҚ МОДЕЛЬДЕУ ЖӘНЕ АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР НЕГІЗІНДЕ ИНТЕГРАЦИЯЛАУ

ОНТОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ИНТЕГРАЦИЯ STEM-МЕТОДОВ НА ОСНОВЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В СИСТЕМЕ ПОВЫШЕНИЯ КВАЛИФИКАЦИИ ПРЕПОДАВАТЕЛЕЙ В УСЛОВИЯХ ЦИФРОВОЙ ТРАНСФОРМАЦИИ ОБРАЗОВАНИЯ

ONTOLOGICAL MODELING AND INTEGRATION OF STEM METHODS BASED ON INFORMATION TECHNOLOGIES IN TEACHER PROFESSIONAL DEVELOPMENT SYSTEMS UNDER THE CONDITIONS OF DIGITAL TRANSFORMATION OF EDUCATION

М.Ж. Базарова ^{1*}, Қ. Алибекқызы ¹, С. Адиканова ², Р. Komada ³

¹Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

²Сәрсен Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен қ., Қазақстан

³Люблин технологиялық университеті, Люблин қ., Польша

*Жауапты автор: Базарова Мадина Жомартовна, e-mail: madina_vkgtu@mail.ru

Түйінді сөздер:

STEM-білім беру,
оқытушылардың
біліктілігін арттыру,
кәсіби құзыреттілік, білім
берудің цифрлық
трансформациясы, STEM-
әдістер, әдіснамалық
модель, құзыреттілікті
бағалау.

ТҮЙІНДЕМЕ

Мақала білім берудің цифрлық трансформациясы жағдайында оқытушылардың біліктілігін арттыру жүйесіне STEM-әдістерді интеграциялау мәселелерін зерттеуге арналған. Жұмыстың өзектілігі мұғалімдердің заманауи білім беру ортасы мен цифрлық экономиканың талаптарына сәйкес келетін кәсіби құзыреттіліктерін қалыптастыру қажеттілігіне байланысты. Зерттеудің мақсаты оқытушылардың цифрлық, сыни, зерттеу және жобалау-инженерлік құзыреттерін тиімді дамытуды қамтамасыз ететін негізгі STEM-әдістерді талдау және іріктеу болып табылады. Зерттеу әдістері ретінде ғылыми дереккөздерді талдау, педагогикалық тәжірибелерді жүйелеу, STEM әдістерінің, құзыреттерінің, индикаторлары мен бақылау формаларының сәйкестік матрицасын модельдеу және әзірлеу қолданылды. Нәтижесінде STEM-әдістер кешені қалыптастырылды және оқыту нәтижелерінің өлшемділігі мен қайталануын қамтамасыз ететін оқытушылардың біліктілігін арттыру процесіне оларды интеграциялау моделі әзірленді. Зерттеудің ғылыми жаңалығы құзыреттілік және онтологиялық тәсілдер негізінде STEM әдістерін жүйелеуде жатыр. Жұмыстың практикалық маңыздылығы педагогикалық кадрлардың біліктілігін арттыру бағдарламаларын әзірлеу кезінде алынған нәтижелерді қолдану мүмкіндігінен тұрады.



Ключевые слова:

STEM-образование,
повышение
квалификации
преподавателей,
профессиональные
компетенции, цифровая
трансформация
образования, STEM-
методы,
методологическая
модель, оценка
компетенций.

АННОТАЦИЯ

Статья посвящена исследованию вопросов интеграции STEM-методов в систему повышения квалификации преподавателей в условиях цифровой трансформации образования. Актуальность работы обусловлена необходимостью формирования у педагогов профессиональных компетенций, соответствующих требованиям современной образовательной среды и цифровой экономики. Целью исследования является анализ и отбор ключевых STEM-методов, обеспечивающих эффективное развитие цифровых, критических, исследовательских и проектно-инженерных компетенций преподавателей. В качестве методов исследования использованы анализ научных источников, систематизация педагогических практик, моделирование и разработка матрицы соответствия STEM-методов, компетенций, индикаторов и форм контроля. В результате сформирован комплекс STEM-методов и разработана модель их интеграции в процесс повышения квалификации преподавателей, обеспечивающая измеримость и воспроизводимость результатов обучения. Научная новизна исследования заключается в систематизации STEM-методов на основе компетентностного и онтологического подходов. Практическая значимость работы состоит в возможности применения полученных результатов при разработке программ повышения квалификации педагогических кадров.

keywords:

STEM education, teacher
training, professional
competencies, digital
transformation of
education, STEM methods,
methodological model,
competence assessment.

ABSTRACT

The article is devoted to the study of the integration of STEM methods into the teacher training system in the context of the digital transformation of education. The relevance of the work is due to the need for teachers to develop professional competencies that meet the requirements of the modern educational environment and the digital economy. The purpose of the study is to analyze and select key STEM methods that ensure the effective development of digital, critical, research, and design engineering competencies of teachers. The research methods used are analysis of scientific sources, systematization of pedagogical practices, modeling and development of a matrix of compliance with STEM methods, competencies, indicators and forms of control. As a result, a set of STEM methods was formed and a model for their integration into the teacher training process was developed, ensuring the measurability and reproducibility of learning outcomes. The scientific novelty of the research lies in the systematization of STEM methods based on competence-based and ontological approaches. The practical significance of the work lies in the possibility of applying the results obtained in the development of professional development programs for teaching staff.

КІРІСПЕ

Жаһандық цифрлық трансформация және білім экономикасына көшу жағдайында білім беру жүйесін жаңғырту стратегиялық маңызға ие болады. Қазақстан Республикасында бұл процестер мемлекеттік саясаттың басым бағыты ретінде қарастырылады, бұл білім беру процесін белсенді цифрландыруда, инновациялық педагогикалық технологияларды енгізуде және білім беру ортасын ұйымдастыру нысандарын жаңартуда көрінеді. Бұл тұрғыда мұғалімдердің цифрлық және пәнаралық білім беру ортасында жұмыс істеуге дайындығын қамтамасыз ететін кәсіби құзыреттіліктерін қалыптастыру мәселесі ерекше өзекті болып отыр.

Заманауи зерттеулер білім берудегі цифрлық түрлендірулердің тиімділігі негізінен оқытушылардың кәсіби дайындық деңгейімен, олардың оқытудың инновациялық әдістерін қолдану және цифрлық құралдарды білім беру практикасына біріктіру қабілетімен анықталатынын атап көрсетеді. Ғылыми әдебиеттерде педагогтардың біліктілігін арттырудың құзыреттілік, жүйелік және цифрлық-бағдарланған тәсілдері, сондай - ақ STEM-және STEAM-білім беруді енгізуге арналған зерттеулер кеңінен ұсынылған. Сонымен қатар, STEM-әдістерді жүйелеу және оларды оқыту нәтижелерінің өлшемділігі мен қайталануын ескере отырып, оқытушылардың біліктілігін арттыру бағдарламаларына мақсатты интеграциялау мәселелерінің жеткіліксіз әзірленуі сақталуда (Bybee, R. W. (2013), Mishra, P., & Koehler, M. J. (2006), OECD. (2019)).

Оқытушылардың кәсіби қызметіне қойылатын талаптардың артуы жағдайында қолданылатын STEM әдістерін қалыптасқан құзыреттермен және оларды бағалау құралдарымен байланыстыруға мүмкіндік беретін әдіснамалық модельдерді әзірлеу ерекше маңызға ие. Осы процеске бірыңғай ресімделген тәсілдердің болмауы педагог кадрлардың біліктілігін арттырудың құрылымдық және бейімделу модельдерін әзірлеуге бағытталған әрі қарайғы зерттеулердің қажеттілігін негіздейді (OECD. (2021), European Commission. (2022), Li, Y., Schoenfeld, A. H., et al. (2020), Honey, M., Pearson, G., & Schweingruber, H. (2014)).

Осы зерттеудің мақсаты негізгі STEM-әдістерді талдау және іріктеу, сондай-ақ оларды ХХІ ғасырдың кәсіби құзыреттіліктерін қалыптастыруға бағытталған оқытушылардың біліктілігін арттыру процесіне интеграциялау моделін әзірлеу болып табылады. Жұмыста қойылған мақсатқа қол жеткізу үшін мынадай міндеттер шешіледі: құзыреттілік тәсіл тұрғысынан STEM-әдістерді жүйелеу; нысаналы кәсіби құзыреттер тізбесін және олардың қалыптасу индикаторларын айқындау; "STEM-әдіс – құзыреттілік – индикаторлар – құралдар – Бақылау нысандары"сәйкестік матрицасын әзірлеу; оқытушылардың біліктілігін арттырудың тиімділігін бағалаудың әдіснамалық негіздерін негіздеу (Сарсенбаева, Г. Ж. (2021)).

Зерттеу нысаны білім берудің цифрлық трансформациясы жағдайында оқытушылардың біліктілігін арттыру процесі болып табылады. Зерттеу пәні мұғалімдердің кәсіби құзыреттілігін қалыптастырудың STEM-әдістерін және оларды біліктілікті арттыру жүйесіне интеграциялау тетіктерін құрайды. Зерттеудің негізгі әдістері ретінде ғылыми дереккөздерді талдау, модельдеу, педагогикалық тәжірибелерді жүйелеу, сондай-ақ онтологиялық және құзыреттілік тәсілдердің элементтері қолданылды.

Зерттеудің ғылыми жаңалығы оқытушылардың кәсіби құзыреттіліктерін қалыптастыру тұрғысынан STEM әдістерін жүйелі жіктеуде және біліктілікті арттыру процестерін ресімдеу үшін онтологиялық модельдеуді қолдануда жатыр. Кіріспеде ұсынылған тәсіл зерттеудің ғылыми векторын белгілейді және мақаланың негізгі бөлігінде ұсынылған нәтижелерді одан әрі ұсынудың логикалық негізі болып табылады.

Мұғалімнің кәсіби дамуы әр түрлі іс-шара ретінде емес, дәйекті және өзара байланысты кезеңдерді қамтитын біртұтас және үздіксіз процесс ретінде қарастырылуы керек. Тұжырымдамалық модель шеңберінде мұғалімнің кәсіби құзыреттілігін дамыту циклдік жүйе түрінде ұсынылады, онда әр кезең алдыңғы кезеңнің нәтижелеріне сүйенеді және одан әрі кәсіби өсу үшін негіз қалыптастырады (Ибрагимова, Л. А. (2019)).

Тұжырымдамалық құрылым диагностикалық, жобалық, білім беру және аналитикалық процестерді біртұтас интеллектуалды жүйеге біріктіреді. Бұл біліктілікті арттыру тиімділігін арттырып қана қоймай, білім беруді цифрлық трансформациялау контекстінде оқытушылардың дербестендірілген кәсіби дамуы үшін тұрақты негіз құруға мүмкіндік береді.

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Зерттеу объектісі ретінде STEM-бағдарланған білім беру тәсілдерін енгізу жағдайында оқытушылардың біліктілігін арттыру процесі қарастырылады. Зерттеу пәні оқытушылардың кәсіби құзыреттілігін қалыптастырудың STEM әдістері, сондай-ақ оларды біліктілікті арттыру жүйесіне интеграциялау тетіктері болып табылады.

Зерттеу материалдары отандық және шетелдік авторлардың STEM-білім беру, білім беруді цифрлық трансформациялау және құзыреттілік тәсіл мәселелері бойынша ғылыми жарияланымдары, білім беру саласындағы нормативтік-құқықтық құжаттар, сондай-ақ оқытушылардың біліктілігін арттырудың білім беру бағдарламалары болды. Зерттеу барысында цифрлық білім беру платформалары мен бағдарламалық құралдар, соның ішінде оқытуды басқару жүйелері, бірлескен жобалық қызмет құралдары, цифрлық модельдеу және деректерді талдау құралдары пайдаланылды, олардың сипаттамалары заманауи білім беру тәжірибелерінің талаптарына сәйкес келеді.

Зерттеудің әдіснамалық негізін оқытушылардың кәсіби даму үдерісін тұтас қарауды қамтамасыз ететін жүйелі, құзыреттілік және онтологиялық тәсілдер құрайды (Ташкенбаева, А.Б., Ахметова, Д. К. (2023)). Зерттеу әдістері ретінде ғылыми дереккөздерді талдау және жалпылау, педагогикалық тәжірибелерді салыстырмалы талдау, кәсіби құзыреттерді модельдеу және құрылымдау, сондай-ақ "STEM – әдіс – құзыреттілік – индикаторлар – құралдар-Бақылау нысандары" сәйкестік матрицасын әзірлеу қолданылды.

Кәсіби құзыреттердің қалыптасуын бағалау үшін сараптамалық бағалау әдісі және құзыреттердің даму деңгейін (төмен, орташа, жоғары) сандық анықтауға мүмкіндік беретін индикаторлардың балдық шкаласы пайдаланылды. Нәтижелерді өңдеу математикалық статистика әдістерін және алынған деректердің объективтілігі мен қайталануын қамтамасыз ететін көрсеткіштерді нормалау әдістерін қолдану арқылы жүзеге асырылды.

Зерттеудің дәйектілігі аналитикалық кезеңді (теориялық және нормативтік дереккөздерді зерттеу), жобалау кезеңін (STEM әдістерін интеграциялаудың тұжырымдамалық және онтологиялық модельдерін әзірлеу) және ұсынылған модельдің тиімділігін талдауға бағытталған бағалау кезеңін қамтыды. Қолданылған әдістер мен процедуралар зерттеуді ұқсас жағдайларда қайталауға және алынған нәтижелерді оқытушылардың біліктілігін арттыру тәжірибесінде қолдануға мүмкіндік береді.

Оқытушылардың біліктілігін арттырудағы STEM-тәсіл қолданбалы есептерді шешуге, нақты процестерді модельдеуге және зерттеушілік ойлауды дамытуға бағытталған оқытудың белсенді және интерактивті әдістерін пайдалануды көздейді (Габдуллин, Р.Р., Сулейменова, А.Т. (2022), Кузьмин, Н.В., Сергеева, М.В. (2020)). STEM-дің негізгі әдістеріне жобалау-зерттеу қызметі, проблемалық-бағдарланған оқыту, кейстерге негізделген оқыту (case-based learning), инженерлік жобалау, сондай-ақ цифрлық симуляторларды, виртуалды зертханаларды және аналитикалық платформаларды пайдалану жатады.

Педагогтарға ғылым, технология, инженерия және математика элементтерін біріктіретін пәнаралық білім беру жобаларын әзірлеуге және сынақтан өткізуге мүмкіндік беретін жобалық оқыту әдісі ерекше маңызға ие. Біліктілікті арттыру процесінде мұндай жобаларды іске асыру топтық жұмыс, сыни талдау, жүйелік ойлау және деректер негізінде шешім қабылдау дағдыларын дамытуға ықпал етеді.

Проблемалық-бағдарланған және зерттеушілік оқыту әдістері оқытушылардың білім беру проблемаларын өз бетінше анықтау, гипотезаларды тұжырымдау, зерттеудің барабар құралдарын таңдау және алынған нәтижелерді түсіндіру қабілетін қалыптастыруды қамтамасыз етеді. Бұл өз кезегінде мұғалімдердің STEM және STEAM тәсілдерін нақты білім беру практикасына енгізуге дайындығын арттырады.

Виртуалды зертханаларды, деректерді талдау платформаларын, бағдарламалау орталарын және цифрлық модельдеуді қоса алғанда, цифрлық STEM құралдарын біріктіру оқытуды жекелеңдіру мүмкіндіктерін кеңейтеді және біліктілікті арттыру бағдарламаларын оқытушылардың жеке кәсіби қажеттіліктеріне бейімдеуге мүмкіндік береді. Осы құралдарды пайдалану цифрлық сауаттылықты және заманауи білім беру технологияларымен жұмыс істеудің тұрақты дағдыларын қалыптастыруға ықпал етеді.

Осылайша, оқытушылардың біліктілігін арттыру жүйесіне негізгі STEM-әдістерді анықтау және енгізу оқытудың практикалық бағыттылығын, оның тиімділігін арттыруды және білім берудің цифрлық трансформациясының қазіргі заманғы талаптарына сәйкестігін қамтамасыз ететін кәсіби дамудың тұжырымдамалық моделінің маңызды құрамдас бөлігі болып табылады (Schraw, G., Crippen, K. J., & Hartley, K. (2006), Абдрахманова, Г. И. (2021), Беспалова, О. В. (2020)).

Оқытушылардың біліктілігін арттыру жүйесіндегі STEM-әдістер, оқытушының қалыптастырылатын құзыреттері және оларды іске асыру құралдары 1-кестеде келтірілген.

1-кесте. Оқытушылардың біліктілігін арттыру жүйесіндегі
STEM-әдістер

STEM-әдіс	Қалыптастырылатын құзыреттер оқытушының	Аспаптар және қаражат іске асыру
Жобалық-зерттеушілік оқыту (Project-Based Learning)	Жобалық құзыреттілік, сыни тұрғыдан ойлау, пәнаралық интеграция, топтық жұмыс	Trello, Miro, Google Workspace, GitHub, LMS-платформалар
Проблемалық- бағдарлы оқыту (Problem-Based Learning)	Аналитикалық ойлау, кешенді есептерді шешу, педагогикалық дизайн	Кейс-платформалар, Padlet, цифрлық кейстер, интерактивті сценарийлер
Инженерлік жобалау (Engineering Design Process)	Инженерлік ойлау, жүйелік талдау, технологиялық сауаттылық	АЖЖ- жүйелері, Tinkercad, Arduino IDE, тренажерлар
Зерттеушілік оқыту (Inquiry-Based Learning)	Зерттеушілік құзыреттілік, мәліметтермен жұмыс, ғылыми ойлау	SPSS, Python, Jupyter Notebook, Excel, цифрлық зертханалар
Үлгілеу және модельдеу	Цифрлық құзыреттілік, абстрактілі ойлау, процестерді визуализациялау	PhET, AnyLogic, MATLAB, виртуалды зертханалар
Пәнаралық STEM/STEAM- модульдер	Креативтілік, интегративті ойлау, педагогикалық икемділік	LMS, VR/AR орталары, мультимедиялық платформалар
<i>Ескерту – автор Беспалов О.В. (2020) деректері негізінде жасаған</i>		

Құзыреттіліктің интегралды көрсеткішін есептеу әдістемесі.

Оқытушылардың кәсіби құзыреттіліктерінің қалыптасу деңгейін сандық тұрғыдан бағалау үшін интегралды көрсеткішті есептеу әдістемесі қолданылды. Бұл әдістеме STEM-әдістерді меңгеру нәтижелерін, цифрлық құралдарды пайдалану қабілетін, зерттеушілік және интегративтік дағдыларды кешенді түрде бағалауға мүмкіндік береді.

Бағалау бірнеше кезең бойынша жүргізілді. Бірінші кезеңде әрбір құзыреттілікке сәйкес келетін индикаторлар анықталды. Әр индикатор оқытушының белгілі бір әрекетті орындау деңгейін немесе нақты нәтижеге қол жеткізуін сипаттайды. Мысалы, цифрлық құзыреттілік үшін цифрлық платформаларды қолдану, білім беру ресурстарын әзірлеу, онлайн өзара әрекеттесуді ұйымдастыру сияқты индикаторлар алынды.

Екінші кезеңде әр индикатор балдық шкала бойынша бағаланды. Зерттеуде 0-ден 4 балға дейінгі шкала қолданылды: 0 балл – индикатор көрінбейді немесе қалыптаспаған; 1 балл – төмен деңгей; 2 балл – ортадан төмен деңгей; 3 балл – жеткілікті деңгей; 4 балл – жоғары деңгей. Мұндай шкала құзыреттіліктің көріну дәрежесін біріздендірілген түрде тіркеуге мүмкіндік береді.

Үшінші кезеңде әрбір жеке құзыреттілік бойынша жинақталған көрсеткіш есептелді. Ол (1) формуламен анықталды:

$$C_j = \frac{\sum_{i=1}^{n_j} w_i * x_i}{\sum_{i=1}^{n_j} w_i} \quad (1)$$

мұндағы

C_j – j -құзыреттілік бойынша жинақталған көрсеткіш;

x_i – i -индикатор бойынша алынған балл;

w_i – i -индикатордың салмақ коэффициенті;

n_j – j -құзыреттілікке кіретін индикаторлар саны.

Осы зерттеуде индикаторлардың маңыздылығы тең деп қабылданғандықтан, олардың салмақ коэффициенттері бірдей алынды, яғни $w_i = 1$. Бұл жағдайда формула қарапайым арифметикалық ортаға келтіріледі.

Төртінші кезеңде барлық құзыреттіліктер бойынша жалпы интегралды көрсеткіш есептелді. Ол оқытушының кәсіби дайындық деңгейін кешенді сипаттайды және (2) формула бойынша анықталады:

$$I = \frac{\sum_{j=1}^{m_j} v_j * x_i}{\sum_{j=1}^{m_j} v_j} \quad (2)$$

мұндағы

I – құзыреттіліктің интегралды көрсеткіші;

C_j – j -құзыреттілік бойынша жинақталған мән;

v_j – j -құзыреттіліктің салмақ коэффициенті;

m_j – бағаланатын құзыреттіліктер саны.

Егер барлық құзыреттіліктер тең мағыналы деп қарастырылса, онда интегралды көрсеткіш (3) формула түрде есептеледі:

$$I = \frac{1}{m} \sum_{j=1}^m C_j \quad (3)$$

Есептеу нәтижелерін түсіндіру үшін интегралды көрсеткіш мәндері деңгейлерге бөлінді: 0,00–0,39 – төмен деңгей; 0,40–0,69 – орта деңгей; 0,70–1,00 – жоғары деңгей. Мұндай интерпретация оқытушылардың кәсіби құзыреттіліктерінің қалыптасу деңгейін салыстырмалы түрде талдауға мүмкіндік береді.

Ұсынылған әдістеме бағалау үдерісін формалдандыруға, нәтижелерді сандық түрде салыстыруға және STEM-әдістерді оқытушылардың біліктілігін арттыру жүйесіне енгізу тиімділігін анықтауға жағдай жасайды. Сонымен қатар бұл тәсіл онтологиялық модель құрамындағы «құзыреттілік – индикатор – бағалау құралы – бақылау нысаны» арасындағы семантикалық байланыстарды практикалық бағалау рәсімімен ұштастыруға мүмкіндік береді.

НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

«Оқытушылардың біліктілігін арттыру процесінің онтологиясы» сыныптары мен қасиеттерінің тұжырымдамалық диаграммасы көрсеткі қасиеттерімен байланысқан тоғыз нөмірленген тіктөртбұрыштан (сыныптардан) тұрады. Әр блок онтологияның мәнін білдіреді (2-кесте).

2-кесте. Онтология сабақтарының сипаттамасы

№	Сыныптың атауы
1. Оқытушы	Орталық орталық. Біліктілікті арттыру үдерісіне қатысушының мүддесін білдіреді. Бастап оның шығады негізгі байланыстар басқаларға сыныптарға
2. Құзыреттілік	Оқытушының жеке кәсіби немесе цифрлық дағдыларын көрсетеді
3. Индикатор	Оқыту нәтижесіне қол жеткізудің нақты өлшенетін көрсеткіші
4. Бағалау	Құзыреттіліктерді тексеру немесе тестілеу нәтижелері
5. Білім беру модуль	Біліктілікті арттыру бағдарламасының элементі – белгілі бір құзыреттерді қамтитын курс немесе тақырып
6. Нәтиже оқытудың	Нұсқаушы меңгеруі тиіс нақты дағды немесе білім
7. Жеке дара траектория	Оқытушының кәсіби дамуының дербес бағыты
8. Тапшылық құзыреттер	Дамуды қажет ететін білімдегі немесе дағдылардағы анықталған кемшіліктер
9. Ұсыным	Тапшылықты жоюға арналған автоматтандырылған немесе сараптамалық курс/модуль ұсынысы
<i>Ескерту – автор Беспалов О.В. (2020) деректері негізінде жасаған</i>	

Әр көрсеткі екі сынып арасындағы семантикалық қатынасты (Объектілік қасиеттерді) көрсетеді. Көрсеткілер көзден нысанаға бағытталған. 3-кестеде онтологияның Объектілік қасиеттеріне сәйкес келетін барлық байланыстар келтірілген.

3-кесте. Онтология сыныптарының семантикалық қатынастары

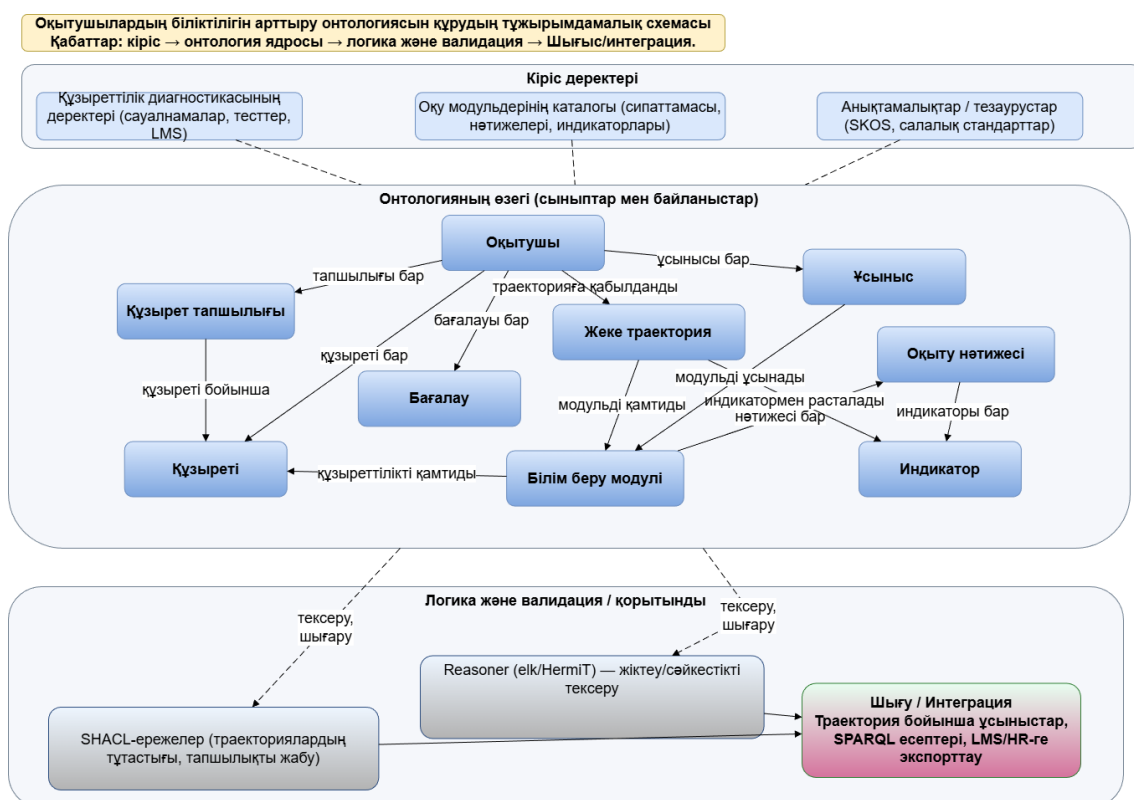
Байланыс бағыты	Объектілік қасиет атауы	Семантикалық түсіндірме
1	2	3
Оқытушы → Құзыреттілік	«ие болады құзыретімен»	оқытушының қазіргі мықты жақтарын көрсетеді
Оқытушы → Тапшылық құзыреттер	«бар тапшылық»	бекітеді облыстың, талап ететіндер дамытудың
Тапшылық құзыреттер → Құзыреттілік	«бойынша құзыреттер»	тапшылықтың қандай құзыретке жататынын нақтылайды
Жеке дара траектория → Білім беру модуль	«қамтиды модуль»	дербес жолды нақты курстармен байланыстырады
Білім беру модуль → Құзыреттілік	«жабады құзыретіне»	модульде қандай дағдылар дамитынын көрсетеді
Білім беру модуль → Нәтиже оқытудың	«бар нәтиже»	анықтайды күтілетін нәтижелер білім беру әсерлері

3-кестенің соңы

1	2	3
Нәтиже оқытудың → Индикатор	«бар индикатор»	нәтижені нақты өлшенетін көрсеткіштермен байланыстырады
Оқытушы → Индикатор	«расталады индикатормен»	бекітеді дерек жетістіктер құзыреттер
Оқытушы → Бағалау	«бар бағалау»	диагностика немесе тест нәтижелерін көрсетеді
Оқытушы → Жеке дара траектория	«есепке алынды арналған траекториясын»	мұғалімді өзінің жеке бағдарла- масымен байланыстырады
Оқытушы → Ұсыным	«бар ұсынымды»	жүйенің оқыту кеңесін құрғанын көрсетеді
Ұсыным → Білім беру модуль	«ұсынады модуль»	қандай курс тапшылықты жоюға көмектесетінін нақтылайды
Ескерту – авторлар Ташкенбаева, А.Б., Ахметова, Д.К. (2023) деректер негізінде құрастырған		

Оқытушылардың біліктілігін арттыру процестерінің бұл тұжырымдамалық моделі (1-сурет.) Protégé / WebProtégé құралдарында нақты онтологияны құруға негіз береді.

Әрбір тіктөртбұрыш (класс) owl: Class болады, әрбір көрсеткі owl:ObjectProperty болады және кіріс және шығыс блоктары интеграцияның семантикалық қабатының шекараларын анықтайды (Import/export RDF/CSV/LMS API).



1-сурет. Оқытушылардың біліктілігін арттыру процестерінің тұжырымдамалық моделі

Ескерту – авторлар Ташкенбаева, А.Б., Ахметова, Д.К. (2023) негізінде құрастырған

Мұғалімдердің кәсіби дамуының тұжырымдамалық моделі шеңберіндегі келесі кезең-оқытушылардың біліктілігін арттыру жүйесіне интеграциялау үшін ең үлкен әлеуетке ие негізгі STEM әдістерін анықтау. Бұл кезең педагогтерде қазіргі білім беру ортасының талаптарына сәйкес келетін тәжірибеге бағытталған, пәнаралық және цифрлық құзыреттіліктерді қалыптастыруды қамтамасыз ететін педагогикалық әдістер мен технологияларды іріктеуге және жүйелеуге бағытталған.

Тұжырымдамалық модельдегі әрбір STEM әдісі тиісті құзыреттілік сыныптарымен (DigitalCompetence, ResearchCompetence, CriticalThinking, CommunicationSkill және т.б.) және педагогикалық іс-әрекеттермен (LearningActivity, AssessmentMethod, EducationalTool) байланысты онтологиялық сынып ретінде әрекет етеді. Бұл құрылым мұғалімнің кәсіби даму процесін өзара байланысты білім беру субъектілері мен олардың арасындағы қатынастардың жиынтығы ретінде ресімдеуге мүмкіндік береді.

Онтологиялық тәсілді қолдану мыналарды қамтамасыз етеді:

- кәсіби құзыреттілік контекстінде STEM әдістерін жүйелеу;
- оқыту әдістері, қалыптасқан Дағдылар мен цифрлық құралдар арасындағы логикалық байланыстарды анықтау;
- оқытушылардың біліктілігін арттырудың жеке траекторияларын бейімдеп жобалау мүмкіндігі.

Әдістемелік аспектіде бұл STEM-бағытталған біліктілікті арттыруды білім берудің цифрлық трансформациясының заманауи талаптарына сәйкес келетін ресми білім үлгілеріне негізделген басқарылатын және қайталанатын процесс ретінде қарастыруға мүмкіндік береді.

Әдістемелік тұрғыдан алғанда, оқытушылардың біліктілігін арттыру жүйесіндегі STEM әдістерін анықтау және жіктеу тұжырымдамалық модельдің аспаптық деңгейінің функциясын орындайды. Олар құзыреттердің теориялық сипаттамасынан оларды қалыптастыру тетіктерін практикалық іске асыруға көшуді қамтамасыз етеді.

STEM әдістерін құзыреттіліктің онтологиялық моделіне біріктіру мүмкіндік береді:

- педагогикалық тәжірибені рәсімдеу;
- оқытушылардың кәсіби өсуін бағалаудың ашықтығын арттыру;
- сандық платформалар мен біліктілікті арттыруды қолдаудың интеллектуалды жүйелерін дамыту үшін негіз құру.

Осылайша, STEM әдістері педагогикалық технологиялар ретінде ғана емес, сонымен қатар педагогикалық кадрларды кәсіби дамытудың әдіснамалық моделінің құрылымдық элементтері ретінде де әрекет етеді.

Зерттеудің әдіснамалық моделі шеңберінде оқытушылардың кәсіби құзыреттіліктерін қалыптастыру STEM-әдістерді игеру деңгейін және педагогтердің оларды практикалық қолдануға дайындық дәрежесін көрсететін бағалау индикаторлары жүйесін әзірлеуді талап ететін өлшенетін процесс ретінде қарастырылады. Бұл жұмыста индикаторлар біліктілікті арттыру нәтижелерін сандық және сапалық бағалауға мүмкіндік беретін операциялық көрсеткіштер болып табылады.

Зерттеуде STEM/STEAM-білім беру логикасына және цифрлық трансформация талаптарына сәйкес келетін 6 негізгі оқытушы құзыреті көрсетілген. Әрбір құзырет операцияларды 3-суретте көрсетілген бақылау мен бағалауға болатын 4-6 индикатор арқылы орындайды.



2-сурет. STEM/STEAM-білім беру логикасына және цифрлық трансформация талаптарына сәйкес келетін оқытушының негізгі құзыреттері
Ескерту – авторлар Беспалов, О.В. (2020) негізінде құрастырған



4-кестеде келтірілген "STEM - әдіс - құзырет - индикаторлар - құрал-Бақылау нысаны" сәйкестік матрицасы оқытушылардың кәсіби құзыреттерімен қалыптастырылатын, олардың қалыптасу индикаторларымен өлшенетін, цифрлық және білім беру құралдарымен өлшенетін, сондай-ақ біліктілікті арттыру нәтижелерін бақылау және бағалау нысандары арасындағы пайдаланылатын педагогикалық әдістер арасындағы өзара байланыстардың логикалық құрылған жүйесін көрсетеді.

4-кесте. STEM әдістерінің сәйкестік матрицасы және құзыреттілікті бағалау

STEM-әдіс	Құзыреттілік	Түйінді индикаторлар	Аспаптар	Бақылау нысаны
Жобалау-зерттеушілік оқыту	Жобалау-инженерлік, коммуникативтік	Шешімдерді жобалау; топтық жұмыс; жобаны қорғау	Trello, Miro, LMS	Жобаны сараптамалық бағалау, рубрика
Проблемалық-бағдарланған оқыту	Сыни	Мәселелерді талдау; шешімдерді дәлелдеу	Case-платформалар, Padlet	Кейісті талдау, ашық жауаптары бар тест
Зерттеушілік оқыту	Зерттеушілік	Болжам; талдау деректердің; қорытындылар	SPSS, Excel, Python	Зерттеу есебі
Инженерлік жобалау	Жобалау-инженерлік	Прототиптеу; тестілеу	CAD, Tinkercad, Arduino	Прототипті демонстрациялау
Үлгілеу және модельдеу	Цифрлық	Модельдерді құру; модельдеудің дұрыстығы	PhET, MATLAB	Тәжірибелік тапсырма
Пәнаралық STEM/STEAM-модульдер	Интегративті	Интеграция пәндердің тізбесі; креативтілік	LMS, VR/AR	Портфолио, сараптамалық бағалау

Ескертү – автор Беспалов, О.В. (2020) деректер негізінде құрастырған

Әзірленген кәсіби құзыреттер тізбесі және индикаторлар жүйесі STEM-бағдарланған білім беру жағдайында оқытушылардың біліктілігін арттыру үдерісін жедел іске асыруды қамтамасыз етеді. Сәйкестік матрицасы "STEM-әдіс – құзыреттілік – индикаторлар – құралдар – Бақылау формалары" мұғалімдердің кәсіби дамуын дәлелді педагогика мен педагогикалық зерттеулер әдіснамасының талаптарына сәйкес келетін басқарылатын, өлшенетін және қайталанатын процесс ретінде қарастыруға мүмкіндік береді.

Онтологиялық модель контекстіндегі құзыреттерді бағалау әдістемесі. Құзыреттердің онтологиялық моделінде бағалау индикаторлары құзыреттілік сыныптарының және оқытушы қызметінің нәтижелерінің қасиеттері (data properties) ретінде ұсынылады. Мысалы, researchcompetence сыныбы isformedby және isevaluatedby типті қатынастар арқылы белгілі бір STEM әдістерімен байланысты "гипотезаны тұжырымдау", "деректерді талдау", "қорытындылардың дұрыстығы" индикаторларына сәйкес келеді.

Мұндай рәсімдеу мүмкіндік береді:

– оқытушылардың кәсіби дамуы туралы мәліметтерді жинау мен талдауды автоматтандыру;

– біліктілікті арттырудың әртүрлі кезеңдеріндегі бағалау нәтижелерінің салыстырмалылығын қамтамасыз ету;

– құзыреттіліктің қалыптасуының жеке көрсеткіштері негізінде оқытудың бейімделу траекторияларын жүзеге асыру.

Әдістемелік тұрғыдан алғанда, дамыған индикаторлар жүйесі тұжырымдамалық модельдің бағалау деңгейінің функциясын орындайды, олардың арасындағы байланысты қамтамасыз етеді:

- қолданылатын STEM әдістерімен;
- қалыптасқан Кәсіби құзыреттіліктер;
- біліктілікті арттыру нәтижелерімен өлшенеді.

Ұсынылған онтология оқытушылардың біліктілігін арттыру жүйесінде STEM-әдістерді қолдану үдерісін формализациялауға арналған. Модельдің негізгі мақсаты — оқытушының кәсіби дамуы, құзыреттері, оқу модульдері, бағалау құралдары, индикаторлар және дәлелдемелік артефактілер арасындағы семантикалық байланыстарды біртұтас құрылымға келтіру. Онтология білім беру үдерісінің негізгі компоненттерін өзара байланыстыра отырып, кәсіби даму нәтижелерін цифрлық ортада жүйелі түрде сипаттауға мүмкіндік береді.

Онтологияның орталық класы Teacher болып табылады, өйткені модельдегі көптеген ұғымдар оқытушының кәсіби қызметімен және оның даму траекториясымен байланысты. Осы класпен Teacher Profile, Competence, Competence level, Learning Trajectory, Module/course, Assessment, Recommendation және Monitoring Event сияқты кластар тікелей байланысады. Бұл байланыстар оқытушының бастапқы жағдайын сипаттаудан бастап, оқу модульдерін таңдауға, бағалау нәтижелерін тіркеуге және кейінгі ұсынымдарды қалыптастыруға дейінгі толық циклді көрсетуге мүмкіндік береді.

Онтологиядағы маңызды ұғымдардың бірі — Competence класы. Бұл класс оқытушының кәсіби, цифрлық, зерттеушілік және STEM-бағытталған құзыреттерін сипаттайды. Құзыреттердің қалыптасу деңгейі Competence level класы арқылы нақтыланады, ал нақты бағалау өлшемдері Indicator класы арқылы беріледі. Осылайша, онтологияда «құзыреттілік – деңгей – индикатор» логикалық тізбегі құрылады. Бұл тізбек оқытушының кәсіби даярлығын формалданған түрде бағалауға негіз болады.

Оқу мазмұнын ұйымдастыру бөлігі Training program, Module/course, Learning activity және STEM method кластары арқылы ұсынылған. Training program класы жалпы біліктілікті арттыру бағдарламасын сипаттайды, ал Module/course жекелеген оқу модульдерін көрсетеді. Әрбір модуль белгілі бір STEM method әдістерімен және нақты Learning activity түрлерімен байланысады. Мұндай құрылым STEM-тәсілдерді оқу мазмұнына жүйелі енгізуге және оларды оқытушының құзыреттерін дамыту мақсаттарымен сәйкестендіруге мүмкіндік береді.

Бағалау компоненті онтологияда ерекше орын алады. Бұл бөлік Assessment, Assessment tool, Control form, Indicator және Evidence/artifact кластары арқылы іске асырылады. Assessment класы бағалау рәсімін немесе бағалау үдерісін сипаттайды. Бағалау барысында қолданылатын нақты құралдар Assessment tool класында көрсетіледі, ал бақылаудың ұйымдастырушылық түрлері Control form класымен беріледі. Бағалаудың нәтижесі мен оны дәлелдейтін материалдар Evidence/artifact класы арқылы тіркеледі. Бұл шешім бағалау нәтижелерінің тек сандық емес, дәлелдемелік негізін де сақтауға жағдай жасайды.

Онтологияның басқарушылық және бейімделушілік бөлігі Recommendation, Adaptation Rule, Monitoring Event, Management Action, Evaluation model және Key Performance Indicator кластары арқылы сипатталады. Recommendation класы бағалау нәтижелеріне сүйене отырып, оқытушыға немесе білім беру бағдарламасына қатысты ұсынымдарды қалыптастыруға мүмкіндік береді. Adaptation Rule класы оқу траекториясын немесе модуль мазмұнын бейімдеу ережелерін сипаттайды. Monitoring Event класы аралық немесе қорытынды мониторинг рәсімдерін көрсетеді. Management

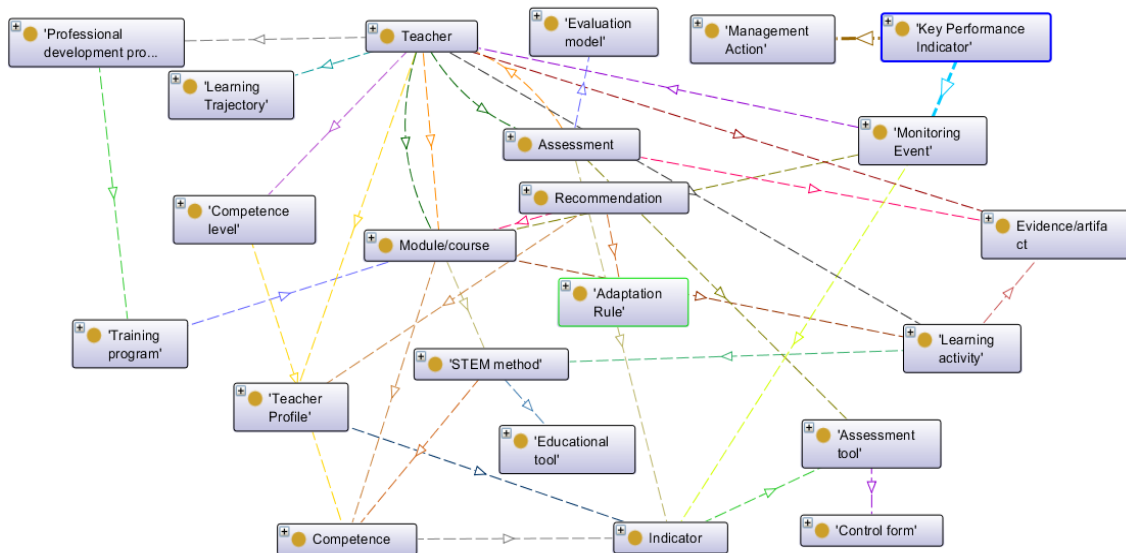
Action және Key Performance Indicator кластары онтологияны тек оқу-әдістемелік емес, сонымен қатар басқарушылық талдау деңгейінде де қолдануға мүмкіндік береді. Яғни модель арқылы кәсіби даму нәтижелерін бағалап қана қоймай, тиімділіктің негізгі көрсеткіштерін бақылауға және басқарушылық шешімдер қабылдауға болады.

Сонымен қатар, онтологияда Educational tool класы қарастырылған, ол цифрлық немесе дәстүрлі білім беру құралдарын сипаттайды. Бұл класс STEM method, Learning activity, Competence және Indicator ұғымдарымен байланыса отырып, оқу құралдарының нақты құзыреттерді дамытудағы рөлін көрсетуге мүмкіндік береді. Осы арқылы оқыту құралдары, әдістері және күтілетін нәтижелер арасындағы өзара тәуелділік семантикалық деңгейде беріледі.

Ұсынылған онтологиялық модельдің маңызды ерекшелігі — оның көпдеңгейлі және өзара байланысқан құрылымында. Модель бір жағынан оқытушының жеке кәсіби профилін сипаттаса, екінші жағынан оқу бағдарламасының мазмұнын, STEM-әдістерді, бағалау рәсімдерін, дәлелдемелік материалдарды және басқарушылық көрсеткіштерді бір ортақ білім кеңістігінде біріктіреді. Мұндай құрылым кәсіби дамуды жеке траектория негізінде ұйымдастыруға, құзыреттерді бағалауды нақты индикаторлармен байланыстыруға және оқыту нәтижелерін цифрлық түрде мониторингтеуге мүмкіндік береді.

Жалпы алғанда, ұсынылған онтология оқытушылардың біліктілігін арттыру жүйесінде STEM-әдістерді енгізудің тұжырымдамалық және формалдық негізін құрайды. Ол «оқытушы – құзыреттілік – индикатор – модуль – әдіс – бағалау – дәлелдеме – ұсыным» тізбегі бойынша семантикалық байланыстарды көрсетіп, білім беру үдерісін құрылымдау мен цифрлық платформаларда іске асыру үшін қолайлы негіз ұсынады.

Бұл оқытушылардың кәсіби даму процесін нақты білім беру нәтижелеріне қол жеткізуге бағытталған және дәлелді педагогика талаптарына сәйкес келетін басқарылатын және қайталанатын жүйе ретінде қарастыруға мүмкіндік береді. 2-суретте оқытушылардың STEM әдістерімен өзара әрекеттесуінің OntoGraf көрсетілген.



3-сурет. OntoGraf оқытушылардың STEM әдістерімен өзара әрекеттесуі

Ескерту – Honey, M., Pearson, G., & Schweingruber, H. (2014) негізінде авторлар құрастырған

Алынған нәтижелер оқытушылардың біліктілігін арттыру жүйесіне STEM-әдістерді интеграциялау мүмкіндіктері туралы қолданыстағы ғылыми түсініктерді кеңейтеді және құзыреттілікке бағдарланған және цифрлық-бағдарланған білім беру әдіснамасын

дамытуға үлес қосады. Кәсіби құзыреттіліктерді бағалаудың әзірленген моделі мен құралдары педагог кадрлардың кәсіби даму тетіктерін жетілдіруге бағытталған әрі қарайғы ғылыми зерттеулер үшін әдіснамалық негіз ретінде пайдаланылуы мүмкін.

Жұмыстың қолданбалы құндылығы алынған нәтижелерді біліктілікті арттыру бағдарламаларын жобалау және іске асыру практикасына енгізу, сондай-ақ оқытушылардың кәсіби өсуін бақылаудың цифрлық білім беру платформалары мен Зияткерлік жүйелерін әзірлеу мүмкіндігі болып табылады. Ұсынылған тәсілдерді қолдану білім беру процестерінің тиімділігін арттыруға, олардың өлшенуін және бейімделуін қамтамасыз етуге, сондай-ақ білім беру жүйесін цифрлық трансформациялау контекстінде STEM-және STEAM-білім беруді дамытуға мүмкіндік береді.

ҚОРЫТЫНДЫ

Зерттеу барысында білім берудің цифрлық трансформациясы жағдайында оқытушылардың біліктілігін арттыру жүйесіне STEM-әдістерді интеграциялау мәселелері қаралды. Оқытушылардың кәсіби құзыреттілігін қалыптастыруды қамтамасыз ететін негізгі STEM-әдістерді талдау және іріктеу мақсатына қол жеткізілді. Жұмыс барысында STEM әдістерін жүйелеу, мақсатты құзыреттер тізімін анықтау және оларды бағалау құралдарын әзірлеу бойынша міндеттер шешілді.

Зерттеудің негізгі нәтижесі құзыреттілік және онтологиялық тәсілдерге негізделген оқытушылардың біліктілігін арттыру процесіне STEM-әдістерді интеграциялаудың кешенді моделін қалыптастыру болды. Әзірленген сәйкестік матрицасы "STEM – әдіс – құзыреттілік – индикаторлар – құралдар-Бақылау нысандары" педагогтердің кәсіби даму процесін ресімдеуді және оны сандық бағалау мүмкіндігін қамтамасыз етуге мүмкіндік берді. Құзыреттіліктің қалыптасуының индикаторлары мен деңгейлерінің жүйесі, сондай-ақ құзыреттіліктің интегралды көрсеткішін есептеу әдістемесі ұсынылды, бұл STEM-білім беру саласындағы зерттеулердің әдіснамалық құралдарын кеңейтеді.

Зерттеудің шектеулеріне негізінен эксперименттік зерттеулердің қажеттілігін анықтайтын ауқымды эмпирикалық тестілеусіз талдаудың әдіснамалық және модельдік деңгейіне назар аудару керек. Әрі қарайғы зерттеулердің болашағы әртүрлі білім беру контексттерінде ұсынылған модельдің тиімділігін тексерумен, бағаланатын құзыреттер тізімін кеңейтумен, сондай-ақ онтологиялық және аналитикалық технологиялар негізінде оқытушылардың біліктілігін арттыруды қолдаудың интеллектуалды жүйелерін әзірлеумен байланысты.

Айта кету керек, зерттеудің қазіргі кезеңінде ұсынылған модельдің эмпирикалық апробациясы әлі жүргізілген жоқ, себебі онтологиялық модель әзірлеу және құрылымдық нақтылау сатысында тұр. Қазіргі уақытта оқытушылардың біліктілігін арттыру жүйесінде модельді кейіннен практикалық іске асыру үшін қажетті негізгі мәндер, байланыстар және бағалау параметрлері формалдандырылуда. Эмпирикалық апробацияны жүргізу онтологиялық модельді әзірлеу аяқталғаннан кейін, оның логикалық дұрыстығы тексеріліп, құзыреттілікті бағалау тетіктері бапталған соң жоспарланып отыр. Зерттеудің келесі кезеңінде модельді оқытушылар іріктемесінде апробациядан өткізу және оның сандық әрі сапалық нәтижелерін талдау көзделеді.

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Авторлар мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: Бұл зерттеуді Қазақстан Республикасы Білім және ғылым министрлігінің Ғылым комитеті қаржыландырады, № AP25793611 «STEM әдістері мен заманауи цифрлық технологияларды біріктіретін оқытушылардың біліктілігін арттыру процесінің бейімделген онтологиялық моделін әзірлеу» гранты.

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ:

Қолданылмайды.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: Қолданылмайды.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Жұмыс сандық модельдеу мен теориялық талдауға негізделгендіктен, жаңа эксперименттік деректер жиналмаған.

АЛҒЫС БІЛДІРУ: Жоқ.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Осы мақаланы дайындау кезінде авторлар ғылыми мәтінді тұжырымдауға және редакциялауға, материалды құрылымдауға, сондай-ақ академиялық стиль талаптарына сәйкес тұжырымдарды нақтылауға көмектесу үшін генеративті жасанды интеллект технологияларын пайдаланды. Жасанды интеллект технологияларын қолдану көмекші сипатта болды және авторлық ғылыми талдауды, нәтижелерді түсіндіруді және тұжырымдарды тұжырымдауды алмастырмады. Мақаланың мазмұны, ұсынылған деректердің дұрыстығы және Академиялық этика нормаларының сақталуы үшін авторлар толығымен жауап береді (OpenAI (2026)).

ӘДЕБИЕТТЕР ТІЗІМІ

- Bybee, R. W. (2013). The case for STEM education: Challenges and opportunities. NSTA Press. <https://doi.org/10.2505/9781936959821>
- European Commission. (2022). European framework for the digital competence of educators (DigCompEdu). <https://doi.org/10.2760/159770>
- Honey, M., Pearson, G., & Schweingruber, H. (2014). STEM integration in K–12 education. National Academies Press. <https://doi.org/10.17226/18612>
- Li, Y., Schoenfeld, A. H., DiSessa, A. A., Graesser, A. C., & Benson, L. C. (2020). On computational thinking and STEM education. *Journal for STEM Education Research*, 3(2), 147–166. <https://doi.org/10.1007/s41979-020-00044-0>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- OECD. (2019). OECD Future of Education and Skills 2030. OECD Publishing. <https://doi.org/10.1787/9789264313832-en>
- OECD. (2021). Teachers and school leaders as valued professionals. OECD Publishing. <https://doi.org/10.1787/9789264641089-en>
- Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education. *Metacognition and Learning*, 1(1), 1–15. <https://doi.org/10.1007/s11409-006-6897-4>
- Абдрахманова, Г. И. (2021). Цифровая трансформация образования: вызовы и перспективы. *Вопросы образования*, 3, 45–61. <https://doi.org/10.17323/1814-9545-2021-3-45-61> // Abdrakhmanova, G. I. (2021). Tsifrovaya transformatsiya obrazovaniya: vyzovy i perspektivy [Digital transformation of education: Challenges and prospects]. *Voprosy obrazovaniya*, 3, 45–61. <https://doi.org/10.17323/1814-9545-2021-3-45-61> (In Russ.)
- Беспалова, О. В. (2020). Компетентностный подход в системе повышения квалификации педагогов. *Педагогика*, 4, 58–65. <https://elibrary.ru/item.asp?id=42819764> // Bespalova, O. V. (2020). Kompetentnostnyy podkhod v sisteme povysheniya kvalifikatsii pedagogov [Competency-based approach in the system of teachers' professional development]. *Pedagogika*, 4, 58–65. <https://elibrary.ru/item.asp?id=42819764> (In Russ.)
- Габдуллин, Р. Р., & Сулейменова, А. Т. (2022). STEM-образование как фактор модернизации педагогической подготовки. *Образование и наука*, 24(6), 9–29. <https://doi.org/10.17853/1994-5639-2022-6-9-29> // Gabdullin, R. R., & Suleimenova, A. T.

- (2022). STEM-obrazovaniye kak faktor modernizatsii pedagogicheskoy podgotovki [STEM education as a factor in the modernization of teacher training]. *Obrazovaniye i nauka*, 24(6), 9–29. <https://doi.org/10.17853/1994-5639-2022-6-9-29> (In Russ.)
- Ибрагимова, Л. А. (2019). Инновационные модели повышения квалификации педагогов. Современные проблемы науки и образования, 5. <https://science-education.ru/ru/article/view?id=29216> // Ibragimova, L. A. (2019). Innovatsionnyye modeli povysheniya kvalifikatsii pedagogov [Innovative models of teachers' professional development]. *Sovremennyye problemy nauki i obrazovaniya*, 5. <https://science-education.ru/ru/article/view?id=29216> (In Russ.)
- Кузьмин, Н. В., & Сергеева, М. В. (2020). Методология оценки профессиональных компетенций преподавателей. Высшее образование в России, 29(7), 85–94. <https://doi.org/10.31992/0869-3617-2020-29-7-85-94> // Kuzmin, N. V., & Sergeeva, M. V. (2020). Metodologiya otsenki professional'nykh kompetentsiy prepodavateley [Methodology for assessing teachers' professional competencies]. *Vyssheye obrazovaniye v Rossii*, 29(7), 85–94. <https://doi.org/10.31992/0869-3617-2020-29-7-85-94> (In Russ.)
- Сарсенбаева, Г. Ж. (2021). STEM-технологии в системе непрерывного профессионального образования. Педагогическое образование и наука, 2, 112–118. <https://elibrary.ru/item.asp?id=46381724> // Sarsenbaeva, G. Zh. (2021). STEM-tehnologii v sisteme nepreryvnogo professional'nogo obrazovaniya [STEM technologies in the system of continuing professional education]. *Pedagogicheskoye obrazovaniye i nauka*, 2, 112–118. <https://elibrary.ru/item.asp?id=46381724> (In Russ.)
- Ташкенбаева, А.Б., & Ахметова, Д.К. (2023). Онтологический подход к моделированию компетенций педагога. Информационные технологии в образовании, 1, 33–41. <https://doi.org/10.25683/ITE.2023.01.05> // Tashkenbaeva, A.B., & Akhmetova, D.K. (2023). Ontologicheskoy podkhod k modelirovaniyu kompetentsiy pedagoga [An ontological approach to modeling teacher competencies]. *Informatsionnyye tekhnologii v obrazovanii*, 1, 33–41. <https://doi.org/10.25683/ITE.2023.01.05> (In Russ.)
- ChatGPT. (2026, January 4). Assistance with academic writing, methodological structuring, and editing of a scientific article [Large language model]. OpenAI. <https://chat.openai.com/>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Базарова Мадина Жомартовна – PhD, қауымдастырылған профессор, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан
Базарова Мадина Жомартовна – PhD, ассоциированный профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан
Bazarova Madina Zhomartovna – PhD, associate professor, East Kazakhstan Technical University named after D. Serikbaev, Ust-Kamenogorsk, Kazakhstan,
e-mail: madina_vkgtu@mail.ru,
ORCID: <https://orcid.org/0000-0003-2580-6580>,



Алибекқызы Карлыгаш – PhD, қауымдастырылған профессор, Д.Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Алибекқызы Карлыгаш – PhD, ассоциированный профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Alibekkyzy Karlygash – PhD, associate professor, East Kazakhstan Technical University named after D. Serikbaev, Ust-Kamenogorsk, Kazakhstan,

e-mail: karlygash.eleusizova@gmail.com,

ORCID: <https://orcid.org/0000-0002-6732-4363>,



Адиқанова Салтанат – PhD, қауымдастырылған профессор, Сәрсен Аманжолов Шығыс Қазақстан университеті, Өскемен қ., Қазақстан

Адиқанова Салтанат – PhD, ассоциированный профессор, Восточно-Казахстанский университет им. С. Аманжолова, г. Усть-Каменогорск, Казахстан

Adikanova Saltanat – PhD, Associate Professor, East Kazakhstan University named after Sarsena Amanzholova, Ust-Kamenogorsk, Kazakhstan

e-mail: ersal_7882@mail.ru,

ORCID: <https://orcid.org/0000-0003-0085-8384>,



Paweł Komada – PhD, қауымдастырылған профессор, Люблин технологиялық университеті, Люблин қ., Польша

Paweł Komada – PhD, ассоциированный профессор, Люблинский Технологический Университет, г. Люблин, Польша

Paweł Komada – PhD, Associate Professor, Lublin University of Technology, Lublin, Poland,

e-mail: p.komada@pollub.pl,

ORCID: <https://orcid.org/0000-0002-0843-8053>

https://doi.org/10.51885/3134-8025_IICS_2026_2_5

SRSTI 28.23.29, 68.85.85

INTERPRETABLE REMAINING USEFUL LIFE PREDICTION OF ROLLING BEARINGS FROM SINGLE-CHANNEL VIBRATION USING RANDOM FOREST, XGBOOST AND LSTM

RANDOM FOREST, XG BOOST ЖӘНЕ LSTM АЛГОРИТМДЕРІН
ПАЙДАЛАНА ОТЫРЫП, БІР АРНАЛЫ ДІРІЛГЕ НЕГІЗДЕЛГЕН
ЖЫЛЖЫМАЛЫ МОЙЫНТІРЕКТЕРДІҢ ҚАЛДЫҚ ҚЫЗМЕТ ЕТУ
МЕРЗІМІНІҢ ТҮСІНДІРЕЛЕТІН БОЛЖАМЫ

ИНТЕРПРЕТИРУЕМОЕ ПРОГНОЗИРОВАНИЕ
ОСТАТОЧНОГО СРОКА СЛУЖБЫ ПОДШИПНИКОВ КАЧЕНИЯ
НА ОСНОВЕ ОДНОКАНАЛЬНОЙ ВИБРАЦИИ С ИСПОЛЬЗОВАНИЕМ
АЛГОРИТМОВ RANDOM FOREST, XGBOOST И LSTM

A.N. Zharlykassova ^{1*}, O.S. Salykova ¹, S.B. Mausymbaeva ¹

¹NPJSC Akhmet Baitursynuly Kostanay Regional University, Kostanay, Kazakhstan

*Corresponding author: Zharlykassova Anara Nurlanovna, e-mail: anara2719@mail.ru

Keywords:

remaining useful life,
rolling bearings, predictive
maintenance, vibration
analysis, Random Forest,
XGBoost, LSTM, SHAP.

ABSTRACT

Interpretable data-driven remaining useful life (RUL) prediction is crucial for the practical prognostics of rolling bearings in industrial environments. This paper presents a two-stage framework in which Random Forest classifier first identifies anomalous operating regimes, followed by XGBoost and a Long Short-Term Memory (LSTM) regressor for RUL estimation on degradation segments. Features are extracted from fixed-length windows of single-channel vibration data and include a limited set of time-domain statistics and frequency-domain descriptors based on dominant spectral components. The proposed approach is validated using the NASA IMS bearing run-to-failure dataset. The Random Forest classifier achieves 98.6% accuracy on a held-out test set, with high precision and recall for the fault class. A hybrid RUL estimator combining XGBoost and LSTM predictions reduces MAE and RMSE by approximately 15–20% compared to individual models. Model interpretability is enhanced using SHAP, identifying RMS, crest factor, and dominant spectral features as key degradation indicators.

Түйінді сөздер:

қалған қызмет ету
мерзімі, жылжымалы
элементтердің
мойынтіректері,
болжамды техникалық
қызмет көрсету, дірілді

ТҮЙІНДЕМЕ

Деректерге негізделген қалдық қызмет ету мерзімін (Remaining Useful Life) интерпретацияланатын түрде болжау өнеркәсіптік жағдайларда домалау мойынтіректерінің прогностикасын практикалық қолданудың маңызды шарты болып табылады. Бұл жұмыста екі кезеңді әдістеме ұсынылады, онда бірінші кезеңде Random Forest классификаторы жұмыс істеудің аномальды режимдерін анықтай-



талдау, кездейсоқ орман,
XGBoost, LSTM, SHAP.

ды, ал екінші кезеңде XGBoost регрессиялық моделі мен Long Short-Term Memory рекурренттік нейрондық желісі деградация аймақтарындағы қалдық ресурсты бағалау үшін қолданылады. Белгілер бірөлшемді вибрациялық сигналдың тұрақты ұзындықтағы терезелерінен алынады және уақыттық облыстағы шектеулі статистикалық сипаттамаларды, сондай-ақ басым спектрлік компоненттерге негізделген жиіліктік дескрипторларды қамтиды. Эксперименттік тексеру подшипниктердің істен шығуына дейінгі сынақтарды қамтитын NASA IMS деректер жиынтығында жүргізілді. Random Forest классификаторы тесттік іріктемеде 98,6 пайыз дәлдікке қол жеткізіп, ақау класы үшін жоғары толықтық пен нақтылық көрсетеді. XGBoost пен LSTM болжамдарын біріктіретін гибриді RUL бағалауы жеке модельдермен салыстырғанда MAE және RMSE көрсеткіштерін шамамен 15-20 пайызға төмендетеді. Нәтижелердің интерпретациялануы модельді SHAP әдісі арқылы талдау нәтижесінде қамтамасыз етіледі. Бұл тәсіл өндірістік диагностика жүйелерінде қолдануға мүмкіндік береді.

Ключевые слова:

оставшийся срок службы,
подшипники качения,
прогнозирующее
техническое
обслуживание, анализ
вибрации, Random Forest,
XGBoost, LSTM, SHAP.

АННОТАЦИЯ

Интерпретируемое прогнозирование остаточного ресурса (Remaining Useful Life) на основе данных является важным условием практического применения методов прогнозтики подшипников качения в промышленных условиях. В работе предлагается двухэтапная методика, в которой на первом этапе классификатор Random Forest выявляет аномальные режимы работы, а на втором этапе регрессионные модели XGBoost и рекуррентная нейронная сеть Long Short-Term Memory используются для оценки остаточного ресурса на участках деградации. Признаки извлекаются из окон фиксированной длины одномерного вибрационного сигнала и включают ограниченный набор статистических характеристик во временной области, а также частотные дескрипторы, основанные на доминирующих спектральных компонентах. Экспериментальная проверка проведена на наборе данных NASA IMS с испытаниями подшипников до отказа. Классификатор Random Forest обеспечивает точность 98,6 процента на тестовой выборке и высокие значения полноты и точности для класса отказа. Гибридная оценка RUL, объединяющая прогнозы XGBoost и LSTM, снижает значения MAE и RMSE примерно на 15–20 процентов по сравнению с отдельными моделями. Интерпретируемость результатов обеспечивается методом SHAP анализа модели.

INTRODUCTION

Rolling bearings are critical components in rotating machinery and have a strong impact on overall equipment reliability, availability and life cycle cost. Numerous field studies report that 45–55% of failures in rotating equipment are related to bearing degradation, emphasizing the importance of reliable condition monitoring and remaining useful life (RUL) prediction methods for these components (Shen 2021 – Sanakkayala 2022). Unplanned bearing failures can lead to costly downtime, collateral damage to other components, and safety risks.

Conventional maintenance strategies based on fixed schedules or simple threshold alarms are often sub optimal: premature replacement wastes healthy life, while late maintenance increases the risk of catastrophic failure. In contrast, predictive maintenance (PdM) aims to schedule interventions based on the actual health state and estimated RUL of each asset, using data from condition monitoring systems (Sanakkayala 2022).

Among various sensing modalities, vibration analysis is widely recognized as one of the most informative approaches for bearing condition assessment (Li 2024). Increases in vibration amplitude, changes in spectral content, and the appearance of characteristic fault frequencies are well known indicators of defects in bearing elements (inner and outer races, rolling elements, cage).

In the last decade, machine learning (ML) and deep learning (DL) methods have attracted significant attention in vibration-based fault diagnosis and prognostics (Shen 2021 – Neto 2023 – Sun 2025). These methods can model complex nonlinear relationships in high dimensional data and reduce dependence on hand crafted rules. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), LSTM and hybrid architectures have been applied to bearing fault classification and RUL prediction with promising results (Sanakkayala 2022 - Neto 2023 – Sun 2025). However, many published solutions rely on multi-channel data, large deep models with high computational cost, and limited interpretability, which complicates deployment in industrial settings.

In this work we focus on three practical requirements that are often under addressed in the literature:

1. Low hardware complexity: use of a single accelerometer channel instead of multi sensor arrays.
2. Compact feature sets: reliance on simple time and frequency domain statistics computable in real time on edge devices.
3. Model interpretability: ability to explain predictions to domain experts using feature importance analysis.

We propose an interpretable ML framework that combines:

- a Random Forest classifier to distinguish between normal and faulty operating conditions;
- XGBoost and LSTM models for RUL regression on defective segments;
- SHAP analysis to interpret feature influence on both classification and regression outputs.

The models are trained and evaluated on the publicly available NASA Bearing dataset (Lee 2007), which contains run to failure vibration recordings for multiple bearings under accelerated degradation tests. The contributions of this paper are:

- Definition of a single channel, feature based pipeline for bearing RUL prediction, suitable for resource constrained monitoring systems.
- Design of a two-stage architecture combining Random Forest classification with XGBoost and LSTM regression for RUL estimation.
- Use of SHAP-based interpretability to quantify the impact of vibration features on model decisions.
- Experimental evaluation on the NASA Bearing dataset demonstrating 98.6% classification accuracy and 15–20% relative reduction in RUL prediction error for the hybrid model compared to individual regressors.

Paper organization. The remainder of this paper is organized as follows: Related Work reviews recent approaches to bearing fault diagnosis, RUL estimation and explainable AI. The Data Set and Feature Extraction section describes the NASA IMS bearing data and the feature construction procedure. Proposed Methods details the two-stage learning formulation, including the Random Forest fault classifier, the XGBoost and LSTM regressors, and SHAP-based interpretation. The Experimental Setup and Results section presents the evaluation protocol and quantitative results, followed by Discussion and concluding remarks.

MATERIALS AND METHODS

Data driven approaches to bearing fault diagnosis and RUL estimation can be broadly divided into three groups: classical ML based on hand crafted features, deep learning applied to raw or minimally processed signals, and hybrid methods that combine both.

Classical ML with hand crafted features. Early work relied on time domain statistics (mean, RMS, kurtosis, skewness), frequency domain indicators (spectral peaks, band energies) and envelope analysis, followed by classifiers such as support vector machines, k nearest neighbors or shallow neural networks (Shen 2021 – Li 2024). These approaches are relatively lightweight and interpretable but may require extensive feature engineering and may struggle with complex degradation patterns.

Deep learning for bearing RUL. Recent studies have applied deep neural architectures that learn feature representations directly from vibration signals. Physics informed deep learning architectures for bearing fault detection are proposed in (Shen 2021), while explainable deep learning for bearing fault prognosis is studied in (Sanakkayala 2022). CNN LSTM hybrids and attention mechanisms have been shown to improve RUL prediction quality for rolling bearings (Neto 2023 – Sun 2025). However, these models often require high computational resources and large labelled datasets, and their black box nature can hinder industrial acceptance.

Hybrid and ensemble approaches. Hybrid methods that combine ML models at different stages (e.g., feature based ML with deep sequence models) or aggregate multiple predictors have been explored for prognostics (Neto 2023). Gradient boosting methods such as XGBoost have been successfully applied to tabular health indicators and often provide strong performance with good interpretability. Combining tree-based models with LSTM can exploit both static statistical features and temporal patterns in degradation trajectories, but the design and analysis of such ensembles remain an active research area.

In parallel, there is a growing body of work on explainable AI (XAI) for prognostics, including SHAP based interpretation of bearing health models (Sanakkayala 2022). However, many studies focus either on fault classification or on RUL regression, and relatively few present a complete pipeline encompassing anomaly detection, RUL prediction and interpretability using a single vibration channel.

Our work builds on these developments by proposing an integrated, interpretable framework that:

- uses a single accelerometer channel with a compact feature set;
- combines Random Forest, XGBoost, and LSTM in a pragmatic two stage architecture;
- applies SHAP analysis to both classification and RUL regression outputs.

Table 1. Positioning of the proposed framework relative to representative related work

Study	Dataset	Input	Task	XAI
Shen et al. (Shen 2021)	Field + CWRU	Sub-band envelope spectra	Fault detection	Physics-informed
Sanakkayala et al. (Sanakkayala 2022)	Lab datasets	Spectrogram (STFT) + CNN	Fault + RUL	LIME
Neto et al. [(Neto 2023)]	PRONOSTIA	STFT + CNN-GRU	RUL	None
Sun et al. (Sun 2025)	PHM2012	FFT + CBAM-CNN-LSTM	RUL	None
This work	NASA IMS (Lee 2007)	Time/frequency statistics	Fault + RUL	SHAP

Note – Compiled by the author based on data from Zharlykassova, 2026

The comparison is intended to clarify methodological choices (feature-based vs. end-to-end deep learning) and the role of post-hoc explainability in bearing prognostics.

Data Set and Feature Extraction NASA Bearing Dataset

We use the NASA IMS bearing run-to-failure dataset from the NASA Ames Prognostics Data Repository (Lee 2007). The test rig continuously ran rolling-element bearings under constant speed and load until failure. Vibration was recorded as 1-second snapshots sampled at 20 kHz (20,480 points per record) at regular intervals (typically every 10 minutes) across three experiments. Each experiment contains multiple bearings monitored by accelerometers mounted on the bearing housings.

In this study, we focus on a single accelerometer channel to reflect realistic industrial deployments where installing multiple sensors is costly or infeasible. Unless stated otherwise, all results are reported using this single vibration channel.

Sliding Window Segmentation

Raw vibration signals are segmented into fixed-length windows. For the NASA IMS dataset, each record corresponds to a 1 s vibration snapshot sampled at 20 kHz (20,480 samples). In our experiments, we treat each snapshot as one window ($N = 20,480$) and do not apply overlap (overlap = 0), as summarized in Table 2. For each window j , we denote the acceleration samples by $\mathbf{a}_j = [a_{j,1}, a_{j,2}, \dots, a_{j,N}]$.

The remaining useful life for window j is defined as in Eq. (1):

$$R_j = T_{\text{fail}} - t_j, \quad (1)$$

where T_{fail} is the time (or cycle index) at which the bearing reaches failure, and t_j is the time corresponding to the centre of window j . This definition is consistent with common practice in bearing RUL prediction and ensures that RUL monotonically decreases as degradation progresses.

The segmented windows are transformed into a tabular dataset where each row corresponds to one window and columns correspond to engineered time- and frequency-domain features. In total, 7672 windows are used in the final analysis set. This number refers to the post-processed sample set rather than to the full raw archive of the NASA IMS repository: each 1 s vibration snapshot from the selected single-channel subset is treated as one non-overlapping window, and only windows for which a consistent RUL target can be assigned are retained. Each window is labelled with a binary health indicator (faulty = 0 for normal, 1 for faulty) and a continuous RUL value R_j . Because the resulting data are strongly time-correlated, the window-level results reported below should be interpreted as a baseline evaluation rather than as a leakage-resistant estimate of generalization.

Fault label definition. Because the public IMS dataset does not provide explicit per-window class labels, we derive a binary fault indicator y_j from the RUL target by thresholding: $y_j = 1$ (faulty) if $R_j \leq \tau$ and $y_j = 0$ otherwise. In the baseline experiments reported here, τ is defined on the training split as the 10th percentile of the RUL distribution, which operationally marks the late-degradation regime rather than the physical onset of damage. This choice yields an intentionally conservative fault label and can partly explain the combination of very high precision with only moderate recall. More stringent validation should therefore include sensitivity analysis with alternative τ values and/or health-state definitions derived from degradation trends or expert annotations.

Time and Frequency Domain Feature Engineering

For each window, we compute a compact set of statistical features from the time and frequency domains. Let $\mathbf{a}_i$ denote the vibration samples within a window (dropping the window index j for clarity). The time-domain features include (Eqs. (2)–(5)):

Mean value

$$\mu = \frac{1}{N} \sum_{i=1}^N a_i, \quad (2)$$

Root-mean-square (RMS) value

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N a_i^2}, \quad (3)$$

Peak amplitude

$$A_{\text{peak}} = \max_i |a_i|, \quad (4)$$

Crest factor

$$CF = \frac{A_{\text{peak}}}{\text{RMS}}, \quad (5)$$

Optionally, higher order statistics such as skewness and kurtosis.

To capture spectral characteristics, we compute the discrete Fourier transform (DFT) of the window and extract a small number of dominant spectral peaks in specified frequency bands. These peaks approximate the energy around characteristic fault frequencies and their harmonics without requiring full envelope analysis.

To enhance reproducibility, Table 2 summarizes the key properties of the NASA IMS bearing dataset and the data preparation settings used in this study.

Table 2. Dataset and preprocessing summary (to support reproducibility)

Item	Value / setting
Dataset	NASA IMS bearing run-to-failure dataset (Lee 2007)
Sensing modality	Single vibration channel (accelerometer)
Sampling rate	20 kHz (public IMS data)
Samples per record	20,480 (≈ 1 s snapshot)
Record interval	Typically every 10 minutes
Windowing	Each 1 s record is treated as one window; overlap = 0
Number of windows	7,672 in the final post-processed analysis set (one non-overlapping 1 s window per retained snapshot from the selected single-channel subset)
RUL definition	$R_j = T_{\text{fail}} - t_j$ (Eq. 1)
Fault label definition	$y_j = 1$ if $R_j \leq \tau$, else 0; in the baseline experiment τ is set as the 10th percentile of training-set RUL values, defining a conservative late-degradation class of approximately 10% of windows.
Train/test split	Baseline: stratified 80/20 split at window level; 3-fold CV on training set. Recommended additional validation: leave-one-bearing-out and chronological split to reduce leakage risk.
Feature scaling	z-score normalization (StandardScaler) fitted on the training split and applied to validation/test.
<i>Note – Compiled by the author based on data from Zharlykassova, 2026</i>	

Proposed Methods

An overview of the proposed two-stage workflow is illustrated in Figure 1.

Problem Formulation

We consider a two-stage learning problem:

1. Fault classification: Given a feature vector x_j extracted from window j , predict whether the bearing is in a healthy or faulty condition as in Eq. (6):

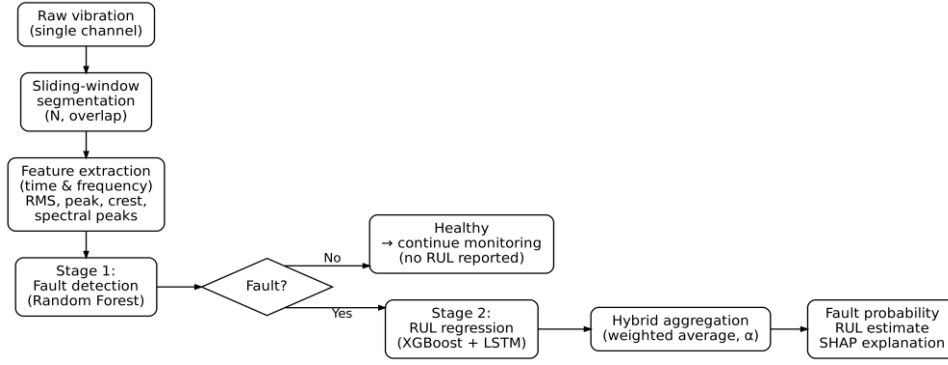


Figure 1. Overview of the proposed two-stage, interpretable RUL prediction pipeline.

Note – Compiled by the author based on data from Zharlykassova, 2026

$$y_j \in 0,1, \quad y_j = 1 \text{ indicates a faulty state.} \quad (6)$$

2. RUL regression: For windows corresponding to faulty or degrading conditions, estimate the remaining useful life R_j , defined in Eq. (1).

The overall framework first applies the fault classifier. For windows predicted as normal, RUL estimation can be skipped or treated as “large and uncertain”. For windows predicted as faulty, the regression models XGBoost and LSTM estimate R_i , and their outputs are aggregated into a hybrid RUL prediction.

Random Forest Fault Classifier

Random Forest (RF) is an ensemble of decision trees trained on bootstrap samples with random feature sub-sampling at each split; the final class prediction is obtained via majority voting across trees (Hu 2023).

Key hyperparameters include:

$n_estimators$: number of trees in the forest;

max_depth : maximum depth of each tree;

$min_samples_split$: minimum number of samples required to split an internal node;

$max_features$: number of features considered at each split.

Increasing $n_estimators$ typically reduces variance but increases computational cost. Constraining max_depth mitigates overfitting by limiting tree complexity. Larger $min_samples_split$ values prevent splits based on small sample counts, while $max_features$ controls the level of randomness in feature selection; a common heuristic is for d features.

We perform grid search with 3 fold cross validation on the training set to tune these hyperparameters. The search grid includes:

$n_estimators \in \{50, 100, 200\}$;

$max_depth \in \{5, 10, \text{None}\}$;

$min_samples_split \in \{2, 5, 10\}$;

$max_features \in \{2, 3, 4\}$ (heuristic: $\sqrt{d}$ for d features).

In total, 81 hyperparameter combinations are evaluated, using classification accuracy as the primary selection metric.

For a given model, the classification accuracy on a dataset of size M is (Eq. (7)):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN'} \quad (7)$$

where TP, TN, FP and FN denote the usual elements of the confusion matrix for the fault class (true positives, true negatives, false positives and false negatives, respectively).

XGBoost and LSTM for RUL Regression

To estimate RUL, we employ two complementary models:

1. XGBoost regressor: an optimized gradient boosting ensemble of regression trees with shrinkage and regularization, which often provides strong accuracy on structured/tabular data (Wiens 2025).

2. LSTM network: a recurrent neural network designed to model temporal dependencies in sequential data via gated memory cells (Krichen 2025). In our setting, the LSTM takes as input a sequence of window-level feature vectors ordered in time and outputs an RUL estimate for the latest window.

The LSTM model is trained to minimize the mean squared error (MSE) loss (Eq. (8)):

$$MSE = \frac{1}{N} \sum_{i=1}^N (R_i - \hat{R}_i)^2, \quad (8)$$

where N is the number of samples, R_i denotes the true RUL value, and $\hat{R}_i$ represents the RUL value predicted by the LSTM model for the i -th sample.

While XGBoost works directly on individual windows, the LSTM exploits the temporal ordering of windows to capture degradation trends. To combine their strengths, we form a hybrid RUL estimator by aggregating the individual predictions, e.g., via a weighted average (Eq. (9)):

$$\hat{R}_j^{(\text{hyb})} = \alpha \cdot \hat{R}_j^{(\text{XGB})} + (1 - \alpha) \cdot \hat{R}_j^{(\text{LSTM})}, \quad (9)$$

where $\alpha \in [0,1]$ is optimized on a validation set. More sophisticated stacking strategies with a meta-learner could be employed, but weighted averaging is preferred here due to its simplicity and transparency.

SHAP Based Model Interpretation

To enhance transparency, we apply SHAP (Shapley Additive exPlanations) to quantify the contribution of each feature to the model output (Ponce 2024). For tree-based models (Random Forest, XGBoost), we use the efficient TreeSHAP formulation; for sequence models, SHAP can be applied to the aggregated window-level representation or to the final regressor output (details in the implementation notes).

Given a feature vector, SHAP represents the model prediction relative to a baseline as (Eq. (10)):

$$f(x) = E[f(x)] + \sum_{k=1}^d \phi_k, \quad (10)$$

where ϕ_k is the SHAP value associated with feature k . Positive ϕ_k indicates that the feature increases the predicted probability of fault or reduces the predicted RUL, while negative values indicate the opposite.

Aggregating SHAP values over many windows yields global feature importance rankings, while temporal plots of SHAP values for individual bearings help visualize how feature influence evolves as degradation progresses.

RESULTS AND DISCUSSION

The full dataset is stratified into training (80%) and test (20%) sets, preserving the class ratio. Within the training set, 3-fold cross validation is used for hyperparameter tuning of the Random Forest and XGBoost models. The LSTM is trained using early stopping based on validation loss to prevent overfitting.

Performance of the fault classifier is evaluated using accuracy, precision, recall and F1 score, with a focus on the fault (positive) class. Performance of the RUL regressors (XGBoost, LSTM, hybrid) is evaluated using:

Mean absolute error (MAE) (Eq. (11)):

$$\text{MAE} = \frac{1}{M} \sum_{j=1}^M |\hat{R}_j - R_j|, \quad (11)$$

Root-mean-square error (RMSE) (Eq. (12)):

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{j=1}^M (\hat{R}_j - R_j)^2}. \quad (12)$$

All results reported below are obtained on the held-out test set that was not used for training or model selection.

Fault Classification Performance

Grid search identifies an effective Random Forest configuration with approximately 100 trees, unlimited depth, moderate minimum samples per split and a small number of features considered at each split. On the test set, the model achieves:

Accuracy: 0.986 (98.6%).

The confusion matrix is shown in Table 3.

Table 3. Confusion matrix of the Random Forest classifier (test set)

True class	Predicted Normal (0)	Predicted Fault (1)
Normal (0)	1380	2
Fault (1)	20	133
<i>Note – Compiled by the author based on data from Zharlykassova, 2026</i>		

From Table 3 we compute, for the fault class, precision, recall and F1-score as in Eqs. (13)–(15):

Precision

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{133}{133+2} \approx 0.99, \quad (13)$$

Recall

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{133}{133+20} \approx 0.87, \quad (14)$$

F1-score

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \approx 0.92. \quad (15)$$

These results indicate that the classifier detects the majority of late-degradation windows while maintaining a very low false alarm rate, which is crucial for avoiding unnecessary shutdowns in industrial practice. The moderate recall suggests that some earlier degrading windows are still missed; this is consistent with the conservative fault definition based on the lower-tail RUL threshold. In practice, this trade-off can be adjusted by tuning both the decision threshold of the classifier and the operational definition of the fault label.

RUL Prediction Performance

Relative RUL prediction performance is summarized in Table 4.

Table 4. Relative RUL prediction performance on the held-out test set (normalized)

Model	MAE (normalized)	RMSE (normalized)
Best single regressor (min of XGBoost, LSTM)	1.00	1.00
Hybrid (α -weighted)	0.80–0.85	0.80–0.85
<i>Note – Compiled by the author based on data from Zharlykassova, 2026</i>		

For RUL regression, Table 4 reports normalized MAE and RMSE, where the best single regressor is used as the reference. The hybrid estimator achieves an approximately 15–20% reduction in both metrics on the held-out test set.

XGBoost captures nonlinear relationships between engineered features and RUL on a per-window basis, whereas the LSTM can leverage temporal ordering to model degradation dynamics. The proposed hybrid estimator combines their strengths and yields consistently lower error than either single regressor in our experiments.

In practice, the most consequential errors occur close to end-of-life. Consistent with this requirement, the hybrid estimator tends to reduce large late-stage deviations relative to the individual regressors, which is particularly important for maintenance decision-making (Madin 2025)

SHAP Based Interpretation

SHAP analysis for the Random Forest and XGBoost models indicates that the dominant drivers of both fault classification and RUL regression are vibration-derived indicators of energy and impulsiveness.

RMS vibration and crest factor are consistently among the most influential features, aligning with the expected increase of vibration energy and impulsive events as damage progresses.

Frequency-domain descriptors (e.g., dominant spectral peaks and their harmonics) also contribute, suggesting growing periodic components and resonance effects near characteristic fault frequencies.

Temporal SHAP profiles for individual bearings typically show a gradual feature drift followed by sharp changes close to failure, supporting the use of these features as practical degradation indicators.

These observations are consistent with prior studies on vibration-based bearing diagnostics and prognostics (Sanakkayala 2021 – Li 2024 – Sun 2025).

The experimental results demonstrate that a single channel, feature based approach can achieve fault classification and RUL prediction performance comparable to more complex multi-channel deep learning architectures reported in the literature (Shen 2021 - Neto 2023 - Sun 2025).

The key advantages of the proposed framework are as follows:

Even simple statistics such as RMS and crest factor capture key aspects of bearing degradation when computed over appropriately chosen windows.

In practical deployments, contextual measurements (e.g., temperature, load, or speed) can improve robustness by distinguishing benign changes from true degradation. However, the benchmark evaluation in this paper focuses on single-channel vibration features to emphasize minimal hardware requirements.

Ensemble methods such as Random Forest and XGBoost are well suited for tabular health indicators and offer good bias-variance trade-offs.

Compared to end-to-end deep models, the proposed framework offers better interpretability and lower computational requirements, which is important for embedded implementations. The use of SHAP provides both global and local explanations that align well with domain knowledge, thereby increasing trust in model outputs among maintenance engineers.

Nevertheless, several limitations should be acknowledged:

The proposed framework was validated using on a single benchmark dataset under relatively controlled conditions. Real-world environments may introduce additional variability (e.g., varying loads, speeds, noise sources) that could degrade performance.

We treat the problem as a supervised learning task with labels derived from run-to-failure experiments. In practice, labeled RUL data may be scarce, which motivates the development of semi-supervised or physics-informed approaches.

Evaluation protocol. Prognostics datasets are strongly time-correlated; if windows from the same bearing trajectory appear in both training and test sets, performance can be overestimated due to temporal leakage. Therefore, the stratified 80/20 split at window level used in this study should be regarded as a baseline protocol for controlled comparison of models, not as the strongest estimate of deployment performance. For publication-grade benchmarking, leakage-resistant validation should additionally be reported using bearing-wise splits such as leave-one-bearing-out and chronological splits that train on earlier portions of each trajectory and test on later degradation segments. Such protocols better reflect the real prognostic setting and provide a more conservative estimate of generalization.

Computational considerations. The pipeline is designed for edge-friendly deployment: feature extraction involves simple time-domain statistics and a lightweight FFT for a small number of spectral peaks, while inference uses tree ensembles and a compact recurrent model. This makes the approach suitable for near-real-time condition monitoring where computational and memory budgets are limited.

The LSTM architecture and hybrid aggregation strategy are relatively simple; more sophisticated sequence models (e.g., attention mechanisms) or meta learning strategies could further improve performance at the cost of complexity.

CONCLUSION

This paper has presented an interpretable machine learning framework for predicting the remaining useful life of rolling bearings using a single accelerometer channel and a compact set of statistical features. The approach combines:

- a Random Forest classifier for robust detection of faulty operating states;
- XGBoost and LSTM regressors for RUL estimation;
- SHAP based analysis for model interpretability.

On the NASA Bearing dataset, the fault classifier achieves 98.6% accuracy and the hybrid RUL estimator improves MAE/RMSE by approximately 15–20% compared to the individual regressors. The SHAP analysis highlights RMS vibration, crest factor and dominant spectral components as key predictors of degradation, in agreement with engineering intuition.

Because the proposed pipeline relies on a single vibration channel, simple features and relatively lightweight models, it is well suited for deployment in resource constrained industrial monitoring systems.

Reproducibility. To facilitate independent verification, we report the full feature list, preprocessing settings, and model configurations. The code for data preparation and model training, together with the train/test split specification and random seeds, is available from the corresponding author upon reasonable request.

Future work will focus on:

- extending the approach to multi sensor configurations and variable operating conditions;
- integrating physics informed constraints and uncertainty quantification into the RUL models;
- validating the framework on real industrial datasets from different sectors (e.g., pumps, compressors, wind turbines), and exploring online learning to adapt to evolving conditions.

CONFLICT OF INTEREST: The authors declare no conflict of interest.

FUNDING: The article was prepared within the framework of the project IRN BR24992785 'Organization and conduct of comprehensive research to ensure a sustainable agro-industrial complex of the Kostanay region by creating a research and technological center'.

INSTITUTIONAL REVIEW BOARD STATEMENT: This study complied with applicable institutional and national ethical standards. As the research was purely technical and did not involve human participants or animals, the requirement for Institutional Review Board approval is Not applicable.

INFORMED CONSENT STATEMENT: Not applicable.

DATA AVAILABILITY STATEMENT: The data supporting the findings of this technical study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS: The authors express their gratitude to the anonymous reviewers for their valuable comments, which helped improve the quality of the article.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: During the preparation of this work, the authors did not use any artificial intelligence (AI) or AI-assisted technologies in the writing process.

REFERENCES

- Shen, S., et al. (2021). A physics-informed deep learning approach for bearing fault detection. *Engineering Applications of Artificial Intelligence*, 103, 104295. <https://doi.org/10.1016/j.engappai.2021.104295>
- Sanakkayala, D. C., et al. (2022). Explainable AI for bearing fault prognosis using deep learning techniques. *Micromachines*, 13(9), 1471. <https://doi.org/10.3390/mi13091471>
- Li, J., Liu, J., & Ren, B. (2024). The impact of random parameter distribution on RVFL model performance in bearing fault diagnosis. *International Journal of Machine Learning and Cybernetics*, 16(12), 9773–9785. <https://doi.org/10.1007/s13042-024-02319-9>
- Neto, D., Petrella, L., et al. (2023). A hybrid deep learning-based approach for rolling bearing fault prognostics. *IFAC-PapersOnLine*, 56(2), 6588–6593. <https://doi.org/10.1016/j.ifacol.2023.10.311>
- Nalbantov, G., et al. (2021). Predictive maintenance tool for non-intrusive inspection systems. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2103.01044>
- Sun, B., et al. (2025). Remaining useful life prediction of rolling bearings based on CBAM-CNN-LSTM. *Sensors*, 25(2), 554. <https://doi.org/10.3390/s25020554>
- Lee, J., Qiu, H., Yu, G., Lin, J., & Rexnord Technical Services. (2007). Bearing data set [Data set]. NASA Ames Prognostics Data Repository. <https://data.nasa.gov/dataset/ims-bearings>
- Койбагаров, М. К., Исабеков, Ж. Н., Курмангалиева, Л. А., Байтурганова, В. К., & Рахметова, П. М. (2025). Разработка системы предиктивного технического обслуживания на основе машинного обучения. *Вестник Университета Шакарима. Серия технические науки*, 2(18), 121–128. [https://doi.org/10.53360/2788-7995-2025-2\(18\)-14](https://doi.org/10.53360/2788-7995-2025-2(18)-14) // Koibagarov, M. K., Isabekov, Zh. N., Kurmangalieva, L. A., Baiturganova, V. K., & Rakhmetova, P. M. (2025). Razrabotka sistemy prediktivnogo tekhnicheskogo obsluzhivaniya na osnove mashinnogo obucheniya [Development of a predictive maintenance system based on

- machine learning]. Vestnik Universiteta Shakarima. Seriya tekhnicheskie nauki, 2(18), 121–128. [https://doi.org/10.53360/2788-7995-2025-2\(18\)-14](https://doi.org/10.53360/2788-7995-2025-2(18)-14) (In Russ.)
- Байгожанова, Д. С., Нартай, К. К., Ахитова, Р. С., Кемель, А. М., & Ибраева, Н. Н. (2025). Проектирование и реализация интеллектуальной системы управления на основе машинного обучения. Вестник КазАТК, 3(138), 414–426. <https://doi.org/10.52167/1609-1817-2025-138-3-414-426> // Baigozhanova, D. S., Nartai, K. K., Akhitova, R. S., Kemel, A. M., & Ibraeva, N. N. (2025). Proektirovanie i realizatsiya intellektualnoi sistemy upravleniya na osnove mashinnogo obucheniya [Design and implementation of an intelligent control system based on machine learning]. Vestnik KazATK, 3(138), 414–426. <https://doi.org/10.52167/1609-1817-2025-138-3-414-426> (In Russ.)
- Сагатова, Л. Б., Бортебаев, С. А., Колга, А. Д., Елемесов, К. К., & Басканбаева, Д. Д. (2025). Оценка технического состояния центробежных насосных агрегатов по вибрационным параметрам. Наука и техника Казахстана, 3, 135–153. <https://doi.org/10.48081/QGAN2347> // Sagatova, L. B., Bortebaev, S. A., Kolga, A. D., Elemesov, K. K., & Baskanbaeva, D. D. (2025). Otsenka tekhnicheskogo sostoyaniya tsentrobezhnykh nasosnykh agregatov po vibratsionnym parametram [Assessment of the technical condition of centrifugal pumping units by vibration parameters]. Nauka i tekhnika Kazakhstan, 3, 135–153. <https://doi.org/10.48081/QGAN2347> (In Russ.)
- Hu, J., & Szymczak, S. (2023). A review on longitudinal data analysis with random forest. Briefings in Bioinformatics, 24(2), bbad002. <https://doi.org/10.1093/bib/bbad002>
- Wiens, M., Verone-Boyle, A., Henscheid, N., Podichetty, J. T., & Burton, J. (2025). A tutorial and use case example of the eXtreme Gradient Boosting (XGBoost) artificial intelligence algorithm for drug development applications. Clinical and Translational Science, 18(3), e70172. <https://doi.org/10.1111/cts.70172>
- Krichen, M., & Mihoub, A. (2025). Long short-term memory networks: A comprehensive survey. AI, 6(9), 215. <https://doi.org/10.3390/ai6090215>
- Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. Clinical and Translational Science, 17(11), e70056. <https://doi.org/10.1111/cts.70056>
- Мадин, В. А., Комаров, Д. Н., Салыкова, О. С., Салыков, Б. Р., Зарубин, М. Ю., & Кобжасаров, Т. Ж. (2025). Предиктивное обслуживание шпиндельных подшипников на основе вибродиагностики на примере токарного станка СКЕ 6136 6150. Вестник КазУТБ, 3(28), 63–79. <https://vestnik.kaztbu.edu.kz/index.php/kazutb/article/view/873/544> // Madin, V. A., Komarov, D. N., Salykova, O. S., Salykov, B. R., Zarubin, M. Yu., & Kobzhasarov, T. Zh. (2025). Prediktivnoe obsluzhivanie shpindelnykh podshipnikov na osnove vibrodiagnostiki na primere tokarnogo stanka SKE 6136 6150 [Predictive maintenance of spindle bearings based on vibration diagnostics: The case of the CKE 6136 6150 lathe]. Vestnik KazUTB, 3(28), 63–79. <https://vestnik.kaztbu.edu.kz/index.php/kazutb/article/view/873/544> (In Russ.)

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Жарлыкасова Анара Нурлановна – докторант, Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті, Қостанай қ., Қазақстан

Жарлыкасова Анара Нурлановна – докторант, Костанайский Региональный университет имени Ахмет Байтұрсынұлы, г. Костанай, Казахстан

Zharlykassova Anara Nurlanovna – doctoral student, NPJSC Akhmet Baitursynuly Kostanay Regional University, Kostanay, Kazakhstan,
e-mail: anara2719@mail.ru,
ORCID: <https://orcid.org/0009-0002-7862-4182>,



Салыкова Ольга Сергеевна – техника ғылымдарының кандидаты, Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті, Қостанай қ., Қазақстан

Салыкова Ольга Сергеевна – кандидат технических наук, Костанайский Региональный университет имени Ахмет Байтұрсынұлы, г. Костанай, Казахстан

Salykova Olga Sergeevna – Candidate of Technical Sciences, NPJSC Akhmet Baitursynuly Kostanay Regional University, Kostanay, Kazakhstan,
e-mail: solga0603@mail.ru,
ORCID: <https://orcid.org/0000-0002-8681-4552>



Маусымбаева Самал Батырбековна – жаратылыстану ғылымдарының магистрі, Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті, Қостанай қ., Қазақстан

Маусымбаева Самал Батырбековна – магистр естественных наук, Костанайский Региональный университет имени Ахмет Байтұрсынұлы, г. Костанай, Казахстан

Maussymbaeva Samal Batyrbekovna – Master of Technical Sciences, NPJSC Akhmet Baitursynuly Kostanay Regional University, Kostanay, Kazakhstan,
e-mail: msamal_93@mail.ru,
ORCID: <https://orcid.org/0009-0006-3407-7702>

https://doi.org/10.51885/3134-8025_IJCS_2026_2_6

XFTAP 81.93.29

МЕТАУНИВЕРСИТЕТТЕГІ ҚҰПИЯЛЫЛЫҚ ЖӘНЕ ДЕРЕКТЕРДІ ҚОРҒАУ: VR СЫНЫПТАРЫ МЕН ӘЛЕМДІК ЧАТТЫ ҚОРҒАУ

КОНФИДЕНЦИАЛЬНОСТЬ И ЗАЩИТА ДАННЫХ В МЕТАУНИВЕРСИТЕТЕ: ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ VR-КЛАССОВ И ВНУТРЕННЕГО ЧАТА

PRIVACY AND DATA PROTECTION IN METAUNIVERSITY: SECURING VR CLASSROOMS AND IN-WORLD CHAT

М.Т. Ипалакова ¹, Д.О. Шарипов ¹, Д.Д. Цой ¹, Е.А. Дайнеко ², A.D. Samala ³

¹Халықаралық ақпараттық технологиялар университет, Алматы, Қазақстан

²Автоматика және ақпараттық технологиялар институты, Сәтбаев университеті, Алматы, Қазақстан

³Негери Паданг университеті, Батыс Суматра, Паданг, Индонезия

Жауапты автор: Цой Дана Дмитриевна, e-mail: d.tsoy@iitu.edu.kz

Түйінді сөздер:

Метауниверситет,
виртуалды шындық
аудиториясы, чат,
құпиялылық, деректерді
қорғау, RBAC, шифрлау,
модерация,
криминалистика.

ТҮЙІНДЕМЕ

Бұл мақалада VR сыныптарын, ішкі мәтіндік/дауыс чатын және оқу нәтижелерін университеттің ақпараттық жүйелерімен синхрондауды біріктіретін білім беру метауниверситетіндегі деректердің құпиялылығы мен қорғалуы қарастырылады. Онда мұндай платформалардың қауіпсіздігі келесі қауіптердің үйлесімімен анықталатыны көрсетілген: тіркелгі мен аватардың бұзылуы, XR телеметриясы мен биосигналдың ағып кетуі, интеграция және желілік синхрондау тәуекелдері және әлеуметтік-техникалық коммуникацияларды теріс пайдалану. Зерттеудің мақсаты - ақпараттың құпиялылығы мен тұтастығын нақты уақыт режимінде қамтамасыз ететін қайталанатын деректер мен коммуникация қауіпсіздігі архитектурасын әзірлеу. Әдістеме деректерді жіктеуді, қауіпті модельдеуді және құпиялылық бойынша дизайн парадигмасы шеңберінде кіру саясатын ресімдеуді қамтиды. Ұсынылған архитектураға SSO/MFA/RBAC, чат және VR сессиясын модерациялау, сыртқы ақпараттық жүйелермен қауіпсіз интеграция, криптографиялық қорғау (TLS, AES-GCM, KMS/HSM) және журнал жүргізу/криминалистикалық тізбек (SIEM, өзгермейтін аудит журналдары) кіреді.

Ключевые слова:

метауниверситет,
виртуальная аудитория,
чат, конфиденциальность,
защита данных, RBAC,
шифрование, модерация,
форензика.

АННОТАЦИЯ

Статья рассматривает обеспечение конфиденциальности и защиты данных в образовательном метауниверситете, объединяющем VR-классы, внутренний текстовый/голосовой чат и синхронизацию учебных результатов с университетскими информационными системами. Показано, что безопасность таких платформ определяется связкой угроз: компрометацией аккаунтов и аватаров, утечками XR-телеметрии и биосигналов, рисками интеграций и сетевой синхронизации, а также социально-техническими злоупот-



реблениями в коммуникациях. Цель работы – разработать воспроизводимую архитектуру защиты данных и коммуникаций, обеспечивающую конфиденциальность и целостность информации в режиме реального времени. Методология включает классификацию данных, моделирование угроз и формализацию политик доступа в парадигме privacy-by-design. Предложена архитектура, включающая SSO/MFA/RBAC, модерацию чата и VR-сессий, защищённую интеграцию с внешними ИС, криптографическую защиту (TLS, AES-GCM, KMS/HSM) и контур логирования/форензики (SIEM, неизменяемые аудит-логи).

keywords:

Metauniversity, virtual reality classroom, chat, privacy, data protection, RBAC, encryption, moderation, forensics.

ABSTRACT

This article examines privacy and data protection in metauniversity that integrates VR classrooms, internal text/voice chat, and the synchronization of learning outcomes with university information systems. It demonstrates that the security of such platforms is determined by a combination of threats: account and avatar compromise, XR telemetry and biosignal leaks, integration and network synchronization risks, and sociotechnical communications abuse. The goal of the study is to develop a reproducible data and communications security architecture that ensures the confidentiality and integrity of information in real time. The methodology includes data classification, threat modeling, and the formalization of access policies within the privacy-by-design paradigm. The proposed architecture includes SSO/MFA/RBAC, chat and VR session moderation, secure integration with external information systems, cryptographic protection (TLS, AES-GCM, KMS/HSM), and a logging/forensics circuit (SIEM, immutable audit logs).

КІРІСПЕ

VR сыныптары мен интеграцияланған байланыс арналары (мәтін, дауыс, медиа) бар заманауи білім беру метакеңістіктері XR сенсорларын, идентификациясын, оқу аналитикасын және байланыс қызметтерін біріктіреді. Сондықтан олардың қауіпсіздігі бір ішкі жүйемен емес, қауіптердің үйлесімімен анықталады. Жалпы метакеңістік архитектурасы деңгейінде негізгі тәуекелдерге тіркелгі мен аватардың бұзылуы, мазмұн мен көріністі ауыстыру, желілік синхрондау шабуылдары, үшінші тарап қызметінің осалдықтары және телеметрия деректерінің ағып кетуі жатады. Виртуалды ортада құдалау, қорқыту және манипуляцияны қоса алғанда, әлеуметтік қорлау да маңызды (Abedin et al., 2022). VR және AR-дың ерекше сипаты ағып кетулердің салдарын күшейтеді. Тек жеке деректер ғана емес, сонымен қатар мінез-құлық және биометриялық сигналдар да қауіп астында, және оларды оқиғадан кейін анонимдеу қиын (Al-kaeed, Qayyum & Qadir, 2024).

Метакеңістік тәуекелдерін жүйелі шолу VR құрылғыларының, жасанды интеллект модульдерінің және интеграциялардың қиылысуын көрсетеді. Білім беру үшін бұл оқушылардың әрекет деректеріне, мысалы, көзқарасқа, қимылдарға, қозғалыстарға және мінез-құлық үлгілеріне жоғары сезімталдықты білдіреді. Бағалау және мұғалім мен оқушының өзара әрекеттесуінің контексті де маңызды (Al-kfairy et al., 2025). Метаверстің шолу зерттеулері қауіптердің көбінесе жүйелік деңгейлер арасында «ағып» кететінін көрсетеді. Сондықтан, көп деңгейлі қауіпсіздік моделі пайдалы. Ол құрылғыларды, желіні, сәйкестендіруді, деректерді және қолданбаларды қамтиды (Laiz-Ibanez, Mendaña-Cuervo & Carus Candas, 2025). Мұны толықтыра отырып, метаверс үшін қауіпсіздік және құпиялылық талаптарын зерттеу оларға қол жеткізуді бақылау, деректерді қорғау, аудит және теріс пайдалануға қарсы тұру сияқты тексерілетін шарттар ретінде қарастыруды ұсынады. Бұл тәсіл талаптарды білім беру VR қызметтерін жобалауға аударуды жеңілдетеді (Ren et al., 2024).

VR сыныбы үшін «деректерді қорғау» ғана емес, сонымен қатар ақпарат ағындары үшін контекстік нормаларды сақтау да маңызды. Деректерге кімнің және қандай жағдайларда қол жеткізе алатынын анықтау қажет. Контекстік тұтастық тәсілі VR сабақтарының жаңа деректер арналарын — кеңістіктік оқиғалар мен жеке өзара әрекеттесулерді жасайтынын көрсетеді. Бұлар білім берудегі құпиялылыққа деген үміттерге қайшы келуі мүмкін (Brehm & Shvartzhnaider, 2024). Бұл практикалық талаптарға әкеледі. Рөлдер мен өкілеттік шекаралары, жиналған сигналдарды азайту, сондай-ақ қол жеткізуді тіркеу және анық хабарландырулар қажет (Brehm & Shvartzhnaider, 2024).

XR биометриясымен тәуекелдердің бөлек класы байланысты. Көз деректері әсіресе маңызды. Ол интерфейстерде және бейімделгіш оқытуда қолданылады. Зерттеулер деанонимизация қаупін көрсетеді. Сондықтан, ресми құпиялылық шаралары қажет. Дифференциалды құпиялылық пен деректерді синтездеуге негізделген тәсілдер перспективті (Dutta, Lanvin & Wunsch-Vincent, 2022). Сонымен қатар, интерактивті VR үшін жеке ағынды деректерді пайдалану әдістері ұсынылады. Олар тұрақтылық пен пайдаланушы тәжірибесіне әсер етуді ескере отырып бағаланады (Shen et al., 2019). Мета-университет үшін бұл қарапайым ережеге әкеледі. Оқу көрсеткіштері шикі биосигналдардан бөлінуі керек. Тек қажетті минимум сақталуы керек. Мүмкіндігінше агрегация және жеке түрлендірулер қолданылуы керек (Dutta, Lanvin & Wunsch-Vincent, 2022).

Кіріктірілген чат және медиа бөлісу құпиялылық пен қауіпсіздік арасында шиеленіс тудырады. Толық шифрлау ұстап алу қаупін азайтады. Дегенмен, бұл зиянды мазмұнды анықтауды қиындатады. Сондықтан «жеке тексеру» схемалары ұсынылады. Бұлар мазмұнды толық ашпай-ақ зиянды анықтауға мүмкіндік береді. Криптографиялық хаттамалар және жеке медиа өңдеу қолданылады. Бұл білім беру үшін маңызды сала. «Бақылау үшін құпиялылықты өшіруге» болмайды. Дизайндық компромисс пен ашық процедуралар қажет (Garg et al., 2023).

Байланыс қауіпсіздігі әлеуметтік-техникалық қабатты қамтиды. Оларға уыттылық, модерация қателері және кемсітушілік әсерлер жатады. Уытты сөйлеуді анықтаудағы бейтараптыққа шолулар пайдаланушы топтары мен контексттер бойынша жүйелі қателіктерді көрсетеді (Giaretta, 2024). Сондықтан, білім беру чаттары үшін жергілікті валидация, ашық критерийлер және адамның бақылауы маңызды. Әділдік көрсеткіштері де қажет (Giaretta, 2024). Уытты жіктеудегі бейтараптыққа қосымша шолулар бейтараптықтың себептерін және оны азайту жолдарын жүйелейді. Олар модельдер мен процедураларды үнемі қайта бағалау қажеттілігін атап көрсетеді (Wang et al., 2023). Метаверс кластарын қабылдаудың эмпирикалық зерттеулері сенім, құпиялылық және әділдік арасындағы байланысты көрсетеді. Бұл платформаны пайдалануға дайындыққа әсер етеді (Laiz-Ibanez, Mendaña-Cuervo & Carus Candas, 2025). Соңында, тәуекелдер, гарнитуралар мен сессиялар деңгейінде күшейеді. Қорғалған тіркелгі болса да, құрылғының бұзылуы мүмкін. Сондықтан, AR және VR-дағы аутентификация бойынша зерттеулер үздіксіз және байқалмайтын пайдаланушыны тексеру механизмдерін ұсынады. Бір мысал - гарнитура қауіпсіздігі үшін жүріске негізделген аутентификация (Wilson et al., 2024). Бұл дәстүрлі шараларды толықтырады және сессияны ұрлау қаупін азайтады.

VR қауіпсіздігі мен құпиялылығына жалпы шолулар қауіптер мен әсерлерді жүйелейді. Олар типтік шабуылдарды, осалдықтардың себептерін және ашық мәселелерді атап көрсетеді. Бұл зерттеулер білім беру платформасын жобалау алдында базалық тәуекел картасы ретінде пайдалы (Zheng, Zhou, Lu & Cai, 2023). Оқу мәтіндік және дауыстық байланысты қолдайтын виртуалды сыныптарда өтетін білім беру метакеңістіктерін дамыту аясында, Құпиялылық және деректерді қорғау мәселелері жүйелік сипатқа ие болуда. Дәстүрлі LMS платформаларынан айырмашылығы, VR сыныптары оқушылар үшін кеңейтілген сандық із қалдырады: оқу нәтижелері мен хабарламалармен қатар, өзара

әрекеттесу деректері, мінез-құлық телеметриясы және қатысу оқиғалары жазылады. Бұл рұқсатсыз кіру, ақпараттың ағып кетуі, оқу нәтижелерінің бұрмалануы және байланыс арналарын теріс пайдалану қауіпін арттырады. Сондықтан, мета-университеттегі қауіпсіздікті әртүрлі шаралар жиынтығы ретінде емес, кіруді басқаруды, коммуникация қауіпсіздігін, деректерді сақтау саясатын және оқиғаларға жауап беру процедураларын байланыстыратын тұтас архитектура ретінде қарастыру керек.

Бұл мақалада зерттеу нысаны метакеңістік, атап айтқанда, VR сыныптары мен ішкі чатты коммуникация құралы ретінде қамтитын сандық платформа ретіндегі мета-университет болып табылады.

Зерттеудің тақырыбы - VR сыныптарында және интеграцияланған чаттарда, соның ішінде университеттің ақпараттық жүйелерімен интеграциялау сценарийлерінде оқушылардың деректері мен байланыстарына құпиялылықты, тұтастықты және бақыланатын қол жеткізуді қамтамасыз ету әдістері мен архитектуралық шешімдері. Сондықтан, зерттеудің негізгі мақсаты - нақты уақыт режимінде деректердің құпиялылығы мен тұтастығын қамтамасыз ете алатын мета-университет үшін деректер мен коммуникация қауіпсіздігі архитектурасын әзірлеу және тексеру.

Осы мақсатқа жету үшін келесі зерттеу мақсаттары анықталды:

1. Мета-университет деректерін сезімталдық деңгейі бойынша жіктеу, сақтау және өңдеу талаптарын анықтау;
2. Білім беру сценарийлерін ескере отырып, қауіп моделін құру;
3. Аутентификация және кіруді бақылау механизмдерін (RBAC), арналардағы және тыныштықтағы деректерді шифрлауды, кілттерді басқаруды және маңызды әрекеттерді аудиттеуді қамтитын қауіпсіздік архитектурасын әзірлеу.
4. Құпиялылық қағидаттарын сақтай отырып, модерация және оқиғаларды тергеу механизмдері бар қауіпсіз ішкі чат жүйесін ұсыну;
5. Ұсынылған тәсілдің қолданылуын техникалық және ұйымдастырушылық критерийлерге сүйене отырып бағалау.

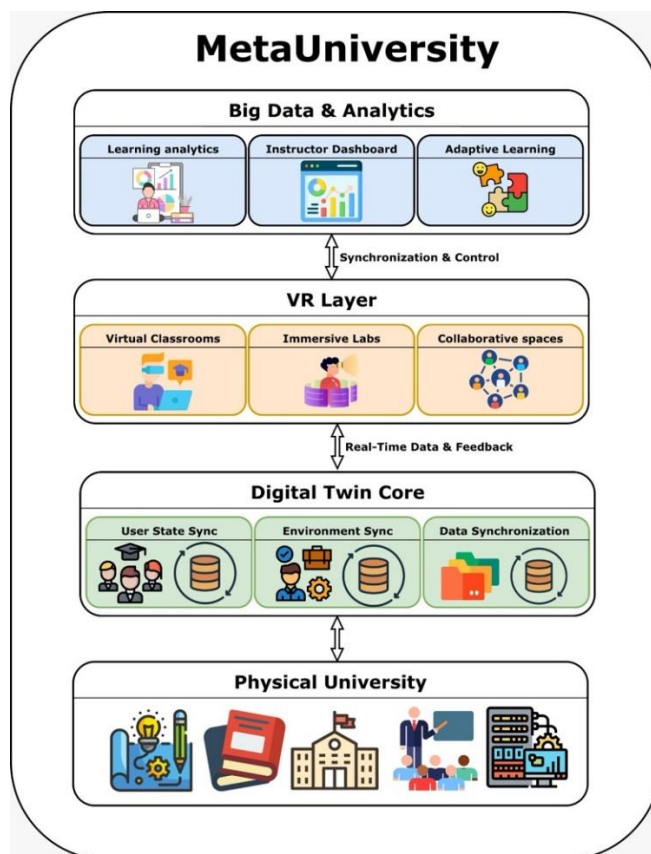
Бұл жобада қолданылатын әдістерге деректерді жіктеу және кіру саясатын ресімдеу, сондай-ақ қауіпсіздік механизмдерінің архитектуралық дизайны кіреді. Бұл тәсіл бізге қауіпсіздік шаралары жиынтығын сипаттауға ғана емес, сонымен қатар зерттеудің ғылыми жаңалығын анықтайтын қауіпсіз білім беру VR платформасын құрудың қайталанатын әдіснамасын ұсынуға мүмкіндік береді.

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Виртуалды шындық (VR) және кеңейтілген шындық (XR) технологиялары иммерсивті, интерактивті оқу тәжірибесін жеңілдету арқылы педагогикалық нәтижелерді жақсартудың айтарлықтай әлеуетін көрсетті. Білім беру мекемелері көбінесе географиялық жағынан шашыраңқы халыққа қызмет көрсететін және физикалық ресурстарға шектеулерге тап болатын Қазақстан жағдайында VR модельденген зертханаларды, тарихи реконструкцияларды және бірлескен топтық жұмысты қоса алғанда, озық оқыту әдістеріне тең қол жеткізу мүмкіндіктерін ұсынады.

Metauniversity платформасы білім берудегі қолданыстағы VR жүйелеріндегі анықталған олқылықтарды, әсіресе статикалық, алдын ала рұқсат етілген 3D активтерге және нақты уақыт режиміндегі көп пайдаланушылы өзара әрекеттесудегі шектеулерге тәуелділікті жоюға арналған. Жоба сұраныс бойынша 3D нысандарын генерациялауға арналған заманауи жасанды интеллект модельдерін интуитивті қимыл және дауыс негізіндегі интерфейстермен біріктіру арқылы тұтынушы деңгейіндегі автономды құрылғыларда динамикалық мазмұн жасауды және үздіксіз ынтымақтастықты қамтамасыз етуге бағытталған.

Бұл бастама информатика, инженерия, мәдениеттану және медициналық оқытуды қоса алғанда, әртүрлі пәндер бойынша қолданылатын масштабталатын құралды әзірлеу бойынша тәуелсіз зертханалық күш-жігерді білдіреді. Ол зертхананың VR өзара әрекеттесу жүйелеріндегі бұрынғы тәжірибесіне сүйене отырып, генеративтік жасанды интеллект пен желілік виртуалды ортадағы заманауи жетістіктерді біріктіреді.



1-сурет. Метауниверситет тұжырымдамасының модульдері

Ескерту – авторлар құрастырған

Зерттеу 1-суретте көрсетілген MetaUniversity тұжырымдамасын әзірлеуге негізделген, ол сандық егіз және аналитикалық тізбек арқылы физикалық университетке қосылған білім беру метакеңістігінің көп қабатты архитектурасы ретінде. Тұжырымдама төрт өзара байланысты қабаттар жүйесі ретінде қалыптасты: (1) Физикалық университет (нақты білім беру процестері мен инфрақұрылымы), (2) Сандық егіз ядро (пайдаланушы күйлерін, орталарын және деректерді синхрондау), (3) VR қабаты (виртуалды сыныптар, иммерсивті зертханалар және бірлескен кеңістіктер) және (4) Үлкен деректер және аналитика (оқу аналитикасы, нұсқаушының басқару тақталары және бейімделгіш оқыту). Тұжырымдаманың екі түрі байланыстардың кілті болып табылады: аналитикалық қабат пен VR қабаты арасындағы синхрондау және басқару (оқытуды басқару және бейімдеу) және VR қабаты мен сандық егіз ядросы арасындағы нақты уақыттағы деректер және кері байланыс (әрекеттерді, оқиғаларды және нәтижелерді нақты уақыт режимінде жазу). Бұл орнату бізге жүйелік құпиялылық талаптарын орнатуға мүмкіндік берді: платформа тек «VR көрінісін көрсетіп қана қоймай», сонымен қатар қабаттар арасында деректерді үздіксіз алмастырады, сондықтан қауіпсіздік тек бір қызметте ғана емес, бүкіл архитектура бойынша деректер ағындары мен рөлдер деңгейінде қамтамасыз етілуі керек.

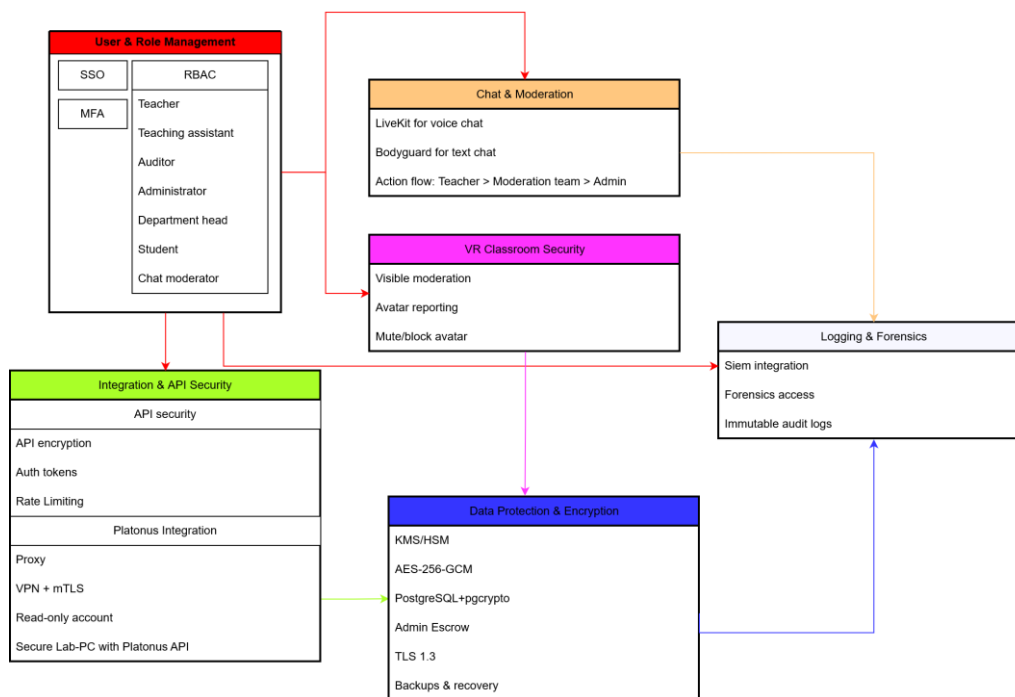
НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

Metaverse басқару тәсілдері екі архитектуралық санатқа бөлінеді: орталықтандырылған және орталықтандырылмаған. Орталықтандырылған платформалар (Web2) бір ғана нысанмен басқарылады, бұл жоғары өнімділікті, қатаң мазмұнды модерациялауды және оңай интеграцияны қамтамасыз етеді. Екінші жағынан, орталықтандырылмаған Web3 архитектуралары икемдірек және виртуалды активтерге иелік етуге мүмкіндік береді, бірақ кіруге үлкен кедергі келтіреді, тұрақсыз және әрқашан қауіпсіз бола бермейді (Chen, 2023). Платформаны таңдау мақсатты аудитория мен мақсаттарға байланысты. Орта және STEM білім беру үшін Minecraft Education және Roblox сияқты орталықтандырылған платформалар жүйелік талаптары төмен геймификацияланған ортаны ұсынған қолайлырақ (Minecraft, 2026; Roblox, 2026). Жоғары білім беру және корпоративтік оқыту үшін Engage VR маңыздырақ. 1-кестеде әртүрлі метаплатформалар олардың архитектурасына, бағытына, артықшылықтары мен кемшіліктеріне негізделе салыстырылады (Engage, 2026).

1-кесте. Метаплатформаларды салыстыру кестесі

Платформа	Сәулет өнері	Негізгі назар білім беруге аударылады	Артықшылықтары	Кемшіліктер
Minecraft білім беру / Roblox	Орталықтандырылған	Мектептегі білім беру, STEM, бағдарламалау, логика	Балаларға таныс орта, ойынға бейімделу, жүйелік талаптардың төмендігі	Жеңілдетілген «блок тәрізді» графика, ойын-сауық мазмұнының алаңдату қаупі
VR-ды қосыңыз	Орталықтандырылған	Жоғары білім, корпоративтік оқыту, күрделі модельдеу	Дәріскерлерге арналған кәсіби құралдар, реалистік физика, білім беру стандарттарымен интеграция	Толық тәжірибе алу үшін қымбат VR гарнитуралары қажет
Кеңістіктік	Орталықтандырылған	Жобалық жұмыс, виртуалды көрмелер, ынтымақтастық	Браузер арқылы (VR жоқ), кросс-платформалар, фотореалистік аватарлар арқылы оңай қол жеткізу	Ресми тестілеу мен емтихандарды өткізуге арналған шектеулі функционалдылық
Орталықсыздандырылған жер / Құмсалғыш	Орталықсыздандырылған (Web3)	Web3 технологияларын, экономикасын, цифрлық өнерді зерттеу	Жасалған кеңістікті толық бақылау, студенттік жобаларды монетизациялау	Кіру шегі жоғары (крипто әмияндар қажет), қажетсіз мазмұнды модерациялау қиын
Ескерту – авторлар құрастырған				

Meta-University қауіпсіз жұмысын қамтамасыз ету үшін 2-суретте көрсетілгендей, орталықтандырылған архитектуруны жүзеге асыратын пайдаланушы деректерін қорғау тәсілі әзірленді.



2-сурет. Метауниверситет тұжырымдамасының модульдері

Ескерту – авторлар құрастырған

Диаграммада метауниверситеттегі құпиялылық пен деректерді қорғауды қамтамасыз етудің ұсынылған архитектурасы көрсетілген, мұнда оқыту VR аудиториясында өткізіледі және ішкі коммуникациялармен (мәтін/дауыс) қолдау көрсетіледі, ал білім беру деректері сыртқы университет жүйесімен (Platonus) синхрондалады. Тұжырымдама өзара байланысты циклдар жиынтығы ретінде құрылымдалған, олардың әрқайсысы белгілі бір тәуекелдер жиынтығына жауап береді: сәйкестік пен құқықтарды басқару, VR-дағы коммуникациялар мен мінез-құлық қауіпсіздігі, интеграция және API қорғау, криптографиялық деректерді қорғау және, соңына келгенде, оқиғаларды тергеуге арналған тіркеу және криминалистика.

Жоғарғы сол жақ блокта сенім мен бақылаудың негізін қалайтын Пайдаланушы және рөлдерді басқару циклі көрсетілген. Аутентификация SSO арқылы жүзеге асырылады және MFA арқылы күшейтіледі, ал платформа функцияларына қол жеткізу RBAC көмегімен ұйымдастырылады. Рөлдік модель негізгі білім беру рөлдерін (Студент, Мұғалім) және операциялық және бақылау рөлдерін (Оқытушы көмекшісі, Чат модераторы, Аудитор, Әкімші, Кафедра меңгерушісі) қамтиды. Бұл жүйедегі деректер мен әрекеттерге кез келген қол жеткізуге (чат модерациясынан бастап сынып жазбаларын өңдеуге дейін) тек нақты анықталған құқықтар шеңберінде және ең аз артықшылық қағидаты шеңберінде рұқсат етілетінін, ал маңызды операцияларды жақсартылған аутентификация мен аудитпен байланыстыруға болатынын білдіреді. Орталық, жоғарғы блок, Чат және Модерация, ішкі коммуникациялық құрылымды және оны басқаруды сипаттайды. Дауыстық арна дауыстық чат үшін LiveKit көмегімен жүзеге асырылады, ал мәтіндік арна мәтіндік чат модерациясы үшін Bodyguard компонентін пайдаланады. Диаграммада анықталған

оқиғаға жауап беру процесі маңызды ерекшелік болып табылады: Мұғалім → Модерация тобы → Админи. Бұл бастапқы жауап беру және зиянның алдын алу мұғалім немесе сынып жетекшісі деңгейінде жүзеге асырылатынын, содан кейін модерация тобы қатысатынын және әкімші тек кеңейтілген өкілеттіктерді және ресми тергеуді қажет ететін істерге ғана қатысатынын білдіреді. Бұл схема әкімшілік құқықтарды теріс пайдалану қаупін азайтады, сонымен бірге жедел жауап беруді қамтамасыз етеді.

Төменде VR сынып қауіпсіздігі блогы берілген, ол VR сессиясындағы тікелей қауіпсіздік шараларын көрсетеді. Жүйеге көрінетін модерация (модератордың/мұғалімнің көрінетін қатысуы), мінез-құлық туралы хабарлау үшін аватар туралы хабарлау механизмдері және аватарларды өшіру/блоктау сияқты жылдам әрекет ету құралдары кіреді. Осылайша, VR сыныбының қауіпсіздігі тек техникалық шаралармен ғана емес, сонымен қатар нақты уақыт режимінде мінез-құлықты бақылаудың әлеуметтік және техникалық механизмдерімен де қамтамасыз етіледі, бұл қорқыту немесе қорлау қаупі бар жоғары интерактивті білім беру ортасы үшін өте маңызды.

Төменгі сол жақтағы Интеграция және API қауіпсіздігі блогы сыртқы қызметтермен және жүйелермен өзара әрекеттесу үшін қауіпсіз периметрді құрайды. API деңгейінде API шифрлау, аутентификация токендері және жылдамдықты шектеу қолданылады, бұл ұстап алу, сұранысты бұрмалау және автоматтандырылған шабуылдар қаупін азайтады. Бұл шеңберде прокси және қауіпсіз VPN + mTLS арнасы арқылы жүзеге асырылатын Platonus-пен интеграция да ерекшеленеді. Сонымен қатар, тәуекелді азайту принципі белгіленеді: тек оқуға арналған тіркелгіні пайдалану Platonus деректерін интеграция шеңберінен өзгерту мүмкіндігін шектейді, ал Platonus API бар Secure Lab-PC сценарийі сыртқы жүйеге тек белгіленген қауіпсіз жұмыс станциясынан кірудің практикалық шектеулерін қарастырады. Бұл дизайн интеграцияны басқарылатын, аудиттелетін және деректер алмасу интерфейсінің бұзылуына төзімді етеді.

Төменгі, орталық блок, Деректерді қорғау және шифрлау, деректердің құпиялылығы мен тұтастығының негізін білдіреді. Кілттерді басқару KMS/HSM жүйесінде жүзеге асырылады, ал қалған кезде деректерді шифрлау AES-256-GCM көмегімен жүзеге асырылады; сақтау PostgreSQL + pgcrypto жүйесінде қамтамасыз етіледі. TLS 1.3 қауіпсіз қосылымдар үшін қолданылады, ал ақауларға төзімділік сақтық көшірмелер мен қалпына келтіру механизмдерімен қолдау көрсетіледі. Әкімшінің деректерді кездейсоқ оқуының орнына, ресми және бақыланатын кіру процесі қажет болған ерекше жағдайларда (мысалы, оқиғаларды тергеу) қорғалған деректерге басқарылатын кіруді қамтамасыз ету үшін жеке механизм, әкімшілік эскроу, қарастырылған. Оң жақтағы журнал жүргізу және криминалистика блогы жүйенің бақылаулылығын және дәлелдемелерді қамтамасыз етеді. Оған орталықтандырылған оқиғаларды бақылау үшін SIEM интеграциясы, тергеу жүргізу үшін уәкілетті қызметкерлердің криминалистикаға кіруі және журналдарды өзгертуден немесе жоюдан қорғау механизмі ретінде өзгермейтін аудит журналдары кіреді. Диаграммадағы қосылымдар оқиғалар мен әрекет нәтижелерінің бұл циклге VR қауіпсіздігі мен деректерді қорғау ішкі жүйесінен кіріп, үздіксіз аудит тізбегін құрайтынын көрсетеді: бұзушылық немесе күдікті әрекет фактісінен бастап оны жазуға, талдауға және кейінгі әкімшілік шешімге дейін.

Осылайша, диаграмма тұтас модельді көрсетеді: рөлдерді басқару (SSO/MFA/RBAC) рұқсат етілген пайдаланушы әрекеттерін анықтайды; чат және VR сыныптары модерация және есеп беру механизмдерімен қорғалған; интеграциялар (соның ішінде Platonus) API және байланыс арнасы деңгейінде оқшауланған және қорғалған; деректер шифрланған және басқарылады.

Деректер KMS/HSM арқылы жиналады; барлық маңызды әрекеттер мен оқиғалар өзгермейтін журналдарда жазылып, бақылау және криминалистика үшін SIEM-ге жіберіледі.

Қауіпсіздік пен өнімділікті теңестіру үшін біз селективті шифрлауды енгізуді жоспарлап отырмыз: маңызды деректер (ID, токендер, аудио) AES-256-GCM арқылы қорғалады, ал XR телеметриясы жеңіл хаттамалармен қорғалады. Бұл VR иммерсиясын сақтай отырып, кідіріс уақытын қолайлы шектерде сақтайды.

Edge-негізделген модерация тұжырымдамасы толық журналдардың орнына тек шифрланған оқиға «жалаушаларын» алмасу арқылы серверге берілетін деректер көлемін азайтады, бұл қоршаған ортаның қауіпсіздігіне нұқсан келтірмей құпиялылықты қамтамасыз етеді.

Журналдың өзгермейтіндігін қамтамасыз ету үшін хэш ағаштары қолданылады. Блоктарды тізбекке хэштеу өнімділік немесе бұзушылық деректерін ретроспективті түрде өңдеудің алдын алады, бұл жүйенің аудиттер үшін ашықтығы мен сенімділігіне кепілдік береді.

Төмендегі 2-кестеде мета-университет ішінде жиналған және өңделген деректер көрсетілген. Әрбір деректер класы үшін мысалдар, сезімталдық деңгейлері, сақтау мерзімдері және рөл бойынша қолжетімділік берілген. Сезімталдық деңгейі ақпараттың ағып кетуі немесе дұрыс пайдаланылмауы жағдайындағы ықтимал залалды көрсетеді. Қолжетімділік үлгі бойынша анықталады: студент, оқытушы, әкімші және аудитор немесе ақпараттық қауіпсіздік қызметі. Деректер ішкі, құпия және өте сезімтал болып жіктеледі. Өте сезімтал деректер үшін қатаң ережелер қолданылады. Мұндай деректерді жинау шектеулі, сақтау мерзімдері қысқартылған және қол жеткізу тек қажет болған жағдайда ғана беріледі. Ұсынылған сақтау мерзімдерін университет саясаты мен ақпараттық қауіпсіздік талаптарына сәйкес реттеуге болады. Бұл матрица қауіп моделін құру және қауіпсіздік шараларын таңдау үшін одан әрі қолданылады.

2-кесте. Метауниверситетте жиналған және өңделетін деректер түрлері

Класс деректер	Жиналған деректер түрі	Сезімтал-дық	Жарамдылық мерзімі	Рөлдерге кіру мүмкіндігі
1	2	3	4	5
Идентификациялық деректер	SSO - идентификатор, рөл, топ, сессия токендері	Құпия (PDn)	Университет/IS ережелері: тіркелгі белсенді болған кезде + мұрағат	IAM әкімшісі, аудитор; топтық тізімдер бойынша мұғалім
Профиль және аватар	Лақап ат, құпиялылық параметрлері, аватар опциялары	Ішкі/ құпия	Тіркелгі белсенді болған кезде; сұраныс бойынша жою	Пайдаланушы; модератор/әкімші - қажет болған жағдайда
VR сеансының деректері	идентификаторы, қатысушылар, уақыт белгілері, көрініс/модуль	Ішкі	90–365 күн (сапа/оқиғаны шешу)	Мұғалім (жеке сыныптар), әкімші, аудитор
XR мінез-құлық телеметриясы	Позициялар/траекториялар, қимылдар, өзара әрекеттесулер	Өте сезімтал (мінез-құлықтық)	Әдепкі: 30–90 күн агрегаттары; "шикі" - сақтамаңыз немесе жөндеу үшін ≤7 күн	Аналитика - тек агрегаттар; қолжетімділік шектеулі

2-кестенің соңы

1	2	3	4	5
Көзді бақылау / көзқарас	Фиксациялар, назар аудару үлгілері	Әсіресе сезімтал (биометриялық/ мінез-құлық)	Әдепкі бойынша сақтамаңыз; қажет болса, синтез/ DP немесе агрегаттар ≤30 күн	Тек рұқсат + аудитор арқылы
Хабарламалар (мәтіндік)	Хабарламалар, реакциялар, арна метадеректері	Құпия	90–180 күн; тәртіптік істер үшін – ережелерге сәйкес	Қатысушылар; шағым модераторы; тергеу кезіндегі аудитор
Байланыс (дауыс)	Аудио ағыны, транскрипция (егер қосылған болса)	Өте сезімтал	Аудио әдепкі бойынша сақталмайды; транскрипция 30-90 күн бойы міндетті емес.	Қатысушылар; модератор/аудитор - ережелерге сәйкес
Оқу нәтижелері	Бағалар, тапсырмалар, қатысу	Құпия (PDn)	Университеттің/IS ережелеріне сәйкес	Студент (меншік), мұғалім (курс), ақпараттық технологиялар әкімшісі, аудитор
Оқиғалар және модерация	Есептер, шешімдер, дәлелдемелер базасы	Құпия	1-3 жыл (немесе ережелерге сәйкес)	Модератор, қауіпсіздік әкімшісі, аудитор
Техникалық журналдар және аудит	Кіру, рөлді өзгерту, API оқиғалары	Ішкі/ құпия	180 күн – 2 жыл	Қауіпсіздік/ аудитор
Ескерту – авторлар құрастырған				

Тәуекелдер VR сессияларының, сәйкестендірудің, оқу аналитикасының, коммуникациялардың және сыртқы интеграциялардың қиылысында туындайды. Сондықтан, бұл зерттеуде архитектуралық тәсіл таңдалды. Ол жүйені деректер ағындары мен рөлдерімен байланыстырылған қауіпсіздік ілмектерінің жиынтығы ретінде сипаттайды.

Басқа зерттеулермен салыстыру тұрақты негізгі тұжырымды көрсетеді. Бұл сигналдар бұзушылықтардың әсерін арттырады. Сондай-ақ, дербестендіру мен аналитика деректерді жинауды басқаратыны, құпиялылықты дизайн бойынша және қатаң кіруді бақылауды талап ететіні расталды. Бұл тәсілдің айырмашылығы - бұл ұғымдар бірыңғай операциялық құрылымға біріктірілген. Ол рөлдерді, деректер арналарын, шифрлауды, модерацияны және тергеулерді байланыстырады.

Нәтижесінде пайда болған әсер архитектуралық құрылыммен түсіндіріледі. Циклдерді бөлу бұзушылықтардың ықтималдығын азайтады. Әрбір қауіп класына нақты шаралар тағайындалады. VR сыныбына көрінетін модерация және жылдам әрекет ету механизмдері қосылды. Чат үшін жауап тізбегі анықталды. Интеграциялар үшін қауіпсіз периметр мен шлюз қарастырылған. Деректер үшін криптографиялық қорғау және

кілттерді басқару қолданылады. Оқиғалар үшін өзгермейтін журналдар мен криминалистика енгізілді. Бірге алғанда, бұл жеке шаралар жиынтығы емес, басқарылатын жүйені жасайды.

Құпиялылық пен модерация арасындағы тепе-теңдік ерекше маңызды. Бұл мақсаттар арасындағы қайшылық әдебиеттерде жиі талқыланады. Толық тексеру бақылауды арттырады, бірақ сенімді азайтады. Бұл зерттеуде ымыраға келу қабылданған. Сақталған деректерді шифрлау сақталады. Шифрланған деректерге қол жеткізу тек ережелерге сәйкес рұқсат етіледі. Модерация автоматты таңдау және адами тексеру арқылы гибриді процесс ретінде құрылады. Дауыс беру үшін транскриптке негізделген тәсіл ұсынылады. Бұл процедураларды стандарттауды және шешімдерді аудиттеуді жеңілдетеді.

Ол әдетте бөлек қарастырылатын бес міндетті біріктіреді:

- 1) Рөлдерді басқару;
- 2) Әлеуметтік-техникалық орта ретіндегі VR сессияларының қауіпсіздігі;
- 3) Бақыланатын модерациясы бар жеке чат;
- 4) Сыртқы ақпараттық жүйелермен қауіпсіз интеграциялар;
- 5) Оқу және тәртіп процесінің бөлігі ретінде журнал жүргізу және криминалистика.

Операциялық жауап беру ағыны да жаңа. Ол әкімшілік артықшылықтарды теріс пайдалану қаупін азайтады және сыныптағы жауап беру жылдамдығын арттырады.

Гипотеза архитектуралық негіздемемен расталады. Деректерді жіктеу, қауіп модельдеу және RBAC, шифрлау, аудит және модерацияның үйлесімі құпиялылық пен тұтастықты жақсартады. Оқиғаларды басқару басқарылатын болып қала береді. Дегенмен, эмпирикалық тексеру әлі де шектеулі. Қазіргі түрінде бұл архитектуралық және әдіснамалық валидация. Ірі көлемді пайдаланушы эксперименттері мен жүктеме өлшемдері ұсынылған жоқ. Сондықтан сандық тиімділікті тек прототиптеуден кейін ғана бағалауға болады. Нәтижелердің дәлдігі мен сенімділігі стандартты криптографиялық примитивтер, кіруді бақылау және аудит және API үшін желілік өлшемдер сияқты қайталанатын механизмдерге негізделген. Бұл шешімдері тексерілуі мүмкін. Олардың дұрыстығы енгізуге және қолдауға байланысты. Негізгі шарттарға әкімшілік тәртіп, сапалы кілттерді басқару және кешенді журнал жүргізу жатады.

Қолданылу шектеулері жүйенің жұмыс істеу ерекшеліктеріне байланысты. Оқиғаларды тергеу процедуралары және өкілеттіктерді нақты бөлу қажет. Рөлдік модельді орнату және қызметкерлерді оқыту өте маңызды. Модерация модельдерін жергілікті түрде тексеру де қажет, себебі қателіктер мен әділетсіз тәжірибелер орын алуы мүмкін. Сыртқы ақпараттық жүйелермен интеграция интерфейстердің қолжетімділігіне және университет саясатына байланысты. Дауыстық деректерді жазуға және материалдарды сақтауға қатысты заңды талаптар өте маңызды.

Бір ықтимал шектеу - өнімділік пен қауіпсіздікті теңестіру қажеттілігі. Тым күрделі нақты уақыттағы шифрлау алгоритмдері пайдаланушы тәжірибесіне кері әсер етуі мүмкін.

Жүйені пайдаланушылар санының көп болуына масштабтау әкімшілік және модерация тұрғысынан жүйені басқаруға қысым жасайды. Бұл мәселені шешу үшін модераторлар құрамын кеңейту немесе автоматтандырылған бақылау жүйелерін енгізу қажет болуы мүмкін.

Тағы бір осалдық - сандық университет ішіндегі өзара әрекеттесу кезінде жиналған деректердің ерекше сипаты: биометриялық параметрлер, қозғалыс үлгілері және көзді бақылау жеке деректер туралы заңдарды сақтау тұрғысынан құқықтық қиындықтар тудырады. Сондықтан, метакөңілсіздіктегі бұзушылықтар үшін құқықтық жауапкершілікті бөлу қажеттілігі бар.

Зерттеудің шектеулері бағалаудың толық еместігіне қатысты. Қазіргі уақытта кідіріс және қосымша көрсеткіштер, сондай-ақ қауіпсіздік шараларының VR сессиясының тұрақтылығына әсерін бағалау жетіспейді. Пайдаланушылардың хабарландырулар мен белгіленген ережелерге сенімі зерттелмеген. Бұл шектеулерді прототип әзірлеу, жүктемені тестілеу және пайдаланушыларды зерттеу арқылы шешу ұсынылады. Модерация процедуралары мен апелляциялық механизмдердің әділдігі бөлек қарастыруды қажет етеді.

Жұмысты одан әрі дамыту үш негізгі бағытта жоспарланған. Бірінші бағыт XR телеметриясының құпиялылықты сақтайтын аналитикасына қатысты және агрегация және деректерді азайту режимдерін пайдалануды қамтиды. Екінші бағыт әділ модерацияны қамтамасыз етуге арналған және тиісті көрсеткіштерді әзірлеуді, жергілікті деректер жиынтығын қалыптастыруды және адам-циклдегі тәсілдерді қолдануды қамтиды. Үшінші бағыт цифрлық әлемнің қауіпсіздігін қамтамасыз етуге бағытталған.

ҚОРЫТЫНДЫ

VR/AR-дың ерекше сипаты құпиялылықтың бұзылуының салдарын күшейтеді, себебі тек жеке деректер ғана емес, сонымен қатар фактіден кейін сенімді түрде жасыру қиын мінез-құлық және биометриялық сигналдар (соның ішінде көзді бақылау) да қауіп астында. VR құрылғыларының, аналитикалық және бейімделгіш модульдердің және сыртқы интеграциялардың қиылысуын ескере отырып, құпиялылықты қамтамасыз ету құпиялылық бойынша дизайн тәсілдерін, бақыланатын қолжетімділікті, криптографиялық және статистикалық қорғау механизмдерін пайдалануды талап етеді.

Зерттеудің негізгі нәтижесі MetaUniversity тұжырымдамасын (1-сурет) сандық егіз және аналитикалық құрылым арқылы физикалық университетпен байланысты көп қабатты архитектура ретінде әзірлеу болды. Бұл тұжырымдама бізге жүйелік талапты ресімдеуге мүмкіндік берді: қауіпсіздік бүкіл архитектура бойынша деректер ағындары мен рөлдер деңгейінде қамтамасыз етілуі керек, себебі платформа қабаттар арасында үздіксіз деректер алмасады (VR сессиялары, күйді синхрондау, аналитика, кері байланыс) және оқушы үшін кеңейтілген сандық із қалдырады. Осы тұжырымдамаға сүйене отырып, пайдаланушы деректерін қорғаудың кешенді архитектурасы ұсынылады (2-сурет), оған сәйкестендіру және кіру құқықтарын басқару, VR класындағы қауіпсіздік, қауіпсіз байланыс және модерация, интеграция және API қорғау, криптографиялық деректерді қорғау, сондай-ақ бақылау және зерттеу инфрақұрылымы кіреді.

Ұсынылған модель нақты университет ортасында қауіпсіздіктің практикалық мүмкіндігін оны тізбектерге бөлу және айқын операциялық логика арқылы қамтамасыз етеді.

Алынған нәтижелердің практикалық маңыздылығы ұсынылған архитектура білім беру үшін қауіпсіз VR платформасын құрудың қайталанатын әдіснамасын анықтайды, ол бір уақытта мыналарды ескереді: (i) білім беру деректері мен байланыстарын қорғау, (ii) VR және чаттағы әлеуметтік-техникалық тәуекелдерді басқару, (iii) сәйкестік және процедуралық ашықтық талаптары. Болашақта зерттеуді одан әрі дамыту бейімделгіш оқыту сапасын сақтай отырып және жиналған сигналдарды азайта отырып, XR телеметриясын деанонизациялаудың қалдық тәуекелдерін ресми бағалауға және құпиялылықты сақтайтын аналитиканы (агрегация, дифференциалды құпиялылық, федеративті тәсілдер) масштабтауға бағытталуы керек. Құпиялылық пен байланыс қауіпсіздігі арасындағы тепе-теңдікке ерекше назар аудару қажет, әсіресе жақсартылған шифрлау және әділ модерация қажеттілігі аясында зиянды мазмұнды бақылау механизмдерін енгізу кезінде, бұл жергілікті деректер мен адамның циклдегі процедураларындағы модельдерді тексерудің маңыздылығын анықтайды.

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Автор мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: Бұл мақала AP23484442 бойынша Ғылым және жоғары білім министрлігінің ғылыми зерттеулерге арналған гранттық қаржыландыру жобасы аясында дайындалды.

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ:
Қолданылмайды.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: Қолданылмайды.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Осы зерттеудің нәтижелері негізінен сандық модельдеу мен теориялық талдауға сүйенеді. Жаңа эксперименттік деректер жиналмаған. Қажет болған жағдайда зерттеу нәтижелерін растайтын материалдар тиісті сұраныс бойынша жауапты автордан алынуы мүмкін.

АЛҒЫС БІЛДІРУ: Авторлар зерттеуді орындау барысында әдістемелік қолдау көрсеткен әріптестеріне және жұмыстың сапасын арттыруға ықпал еткен анонимді рецензенттерге алғыс білдіреді.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Авторлар бұл қолжазбаны дайындау кезінде генеративті жасанды интеллект құралдарын пайдаланбаған.

ӘДЕБИЕТТЕР ТІЗІМІ

- Abedin, B., Meske, C., Junglas, I., & Osmundsen, K. (2022). Designing and managing human–AI interactions. *Information Systems Frontiers*, 24(3), 691–697. <https://doi.org/10.1007/s10796-022-10313-1>
- Al-kaeed, M., Qayyum, A., & Qadir, J. (2024). Privacy preservation in Artificial Intelligence and Extended Reality (AI-XR) metaverses: A survey. *Journal of Network and Computer Applications*, 231, Article 103989. <https://doi.org/10.1016/j.jnca.2024.103989>
- Alhloul, A., & Kiss, E. (2022). Industry 4.0 as a challenge for the skills and competencies of the labor force: A bibliometric review and a survey. *Sci*, 4(3), Article 34. <https://doi.org/10.3390/sci4030034>
- Brehm, K., & Shvartzhaider, Y. (2024). Understanding privacy in virtual reality classrooms: A contextual integrity perspective. *IEEE Security & Privacy*, 22(2), 62–69. <https://doi.org/10.1109/MSEC.2023.3336802>
- Dutta, S., Lanvin, B., & Wunsch-Vincent, S. (Eds.). (2022). Global innovation index 2022: What is the future of innovation-driven growth? World Intellectual Property Organization. https://www.wipo.int/global_innovation_index/en/2022/
- Garg, T., Masud, S., Suresh, T., & Chakraborty, T. (2023). Handling bias in toxic speech detection: A survey. *ACM Computing Surveys*, 56(3), Article 52. <https://doi.org/10.1145/3580494>
- Giarretta, A. (2024). Security and privacy in virtual reality: A literature survey. *Virtual Reality*, 28(2), Article 79. <https://doi.org/10.1007/s10055-024-01079-9>
- Laiz-Ibanez, H., Mendaña-Cuervo, C., & Carus Candás, J. L. (2025). The metaverse: Privacy and information security risks. *Forensic Science International: Digital Investigation*, 52, Article 100373. <https://doi.org/10.1016/j.jjime.2025.100373>
- Ren, X., Fan, J., Xu, N., Wang, S., Dong, C., & Wen, Z. (2024). DPGazeSynth: Enhancing eye-tracking virtual reality privacy with differentially private data synthesis. *Information Sciences*, 675, Article 120720. <https://doi.org/10.1016/j.ins.2024.120720>
- Shen, Y., Wen, H., Luo, C., Xu, Y., Zhang, X., & Hu, J. (2019). GaitLock: Protect virtual and augmented reality headsets using gait. *IEEE Transactions on Dependable and Secure Computing*, 19(1), 418–432. <https://doi.org/10.1109/TDSC.2019.2925774>
- Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H., & Shen, X. (2023). A survey on metaverse: Fundamentals, security, and privacy. *IEEE Communications Surveys & Tutorials*, 25(1), 319–352. <https://doi.org/10.1109/COMST.2022.3202047>

- Wilson, E., Ibragimov, A., Proulx, M. J., Tetali, S. D., Butler, K., & Jain, E. (2024). Privacy-preserving gaze data streaming in immersive interactive virtual reality: Robustness and user experience. *IEEE Transactions on Visualization and Computer Graphics*, 30(5), 2257-2268.
- Zheng, T., Zhou, T., Lu, K., & Cai, Z. (2023). Inspecting end-to-end encrypted communication differentially for the efficient identification of harmful media. *IEEE Transactions on Information Forensics and Security*, 18, 4601–4615. <https://doi.org/10.1109/TIFS.2023.3315067>
- Chen, H., Duan, H., Abdallah, M., Zhu, Y., Wen, Y., & Cai, W. (2023). Web3 metaverse: State-of-the-art and vision. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 20(4), 1–42.
- Microsoft. (n.d.). Minecraft Education [Computer software]. <https://education.minecraft.net/>
- Roblox Corporation. (n.d.). Roblox Education [Virtual reality platform]. <https://education.roblox.com/>
- ENGAGE XR Holdings. (n.d.). Engage VR [Virtual reality platform]. <https://engagevr.io/>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Ипалакова Мадина Тулегеновна – қауымдастырылған профессор, инженерия ғылымдарының кандидаты, Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан.

Ипалакова Мадина Тулегеновна – ассоциированный профессор, кандидат технических наук, Международный университет информационных технологий, Алматы, Казахстан.

Ipalakova Madina – Associate Professor, cand. of technical sciences, International Information Technology University, Almaty, Kazakhstan,

e-mail: m.ipalakova@edu.iitu.kz,

ORCID: <https://orcid.org/0000-0002-8700-1852>,



Шарипов Данил Олегович – Халықаралық ақпараттық технологиялар университетінің компьютерлік технологиялар және киберқауіпсіздік факультетінің ақпараттық технологиялар мамандығы бойынша бірінші курс магистранты, Алматы қ., Қазақстан

Шарипов Данил Олегович – магистрант 1 курса специальности «Информационные технологии» факультета Компьютерные технологии и кибербезопасность, Международный университет информационных технологий, г. Алматы, Казахстан

Sharipov Danil – first-year master's student in the Information Technology specialty, Faculty of Computer Technology and Cybersecurity, International University of Information Technology, Almaty, Kazakhstan

e-mail: 43760@iitu.edu.kz

ORCID: <https://orcid.org/0009-0004-7780-6034>



Цой Дана Дмитриевна – Инженерия ғылымдарының магистрі, Халықаралық ақпараттық технологиялар университетінің «Аралас шындық зертханасы» ғылыми-зерттеу зертханасының меңгерушісі, Алматы қ., Қазақстан

Цой Дана Дмитриевна – магистр технических наук, руководитель научно-исследовательской лаборатории «Лаборатория смешанной реальности» Международный университет информационных технологий, г. Алматы, Казахстан

Tsoy Dana – Master of Engineering Sciences, Head of the Research Laboratory "Mixed Reality Laboratory" International University of Information Technologies, Almaty, Kazakhstan

e-mail: d.tsoy@iitu.edu.kz

ORCID: <https://orcid.org/0000-0003-0172-9760>



Дайнеко Евгения Александровна – қауымдастырылған профессор, PhD, Автоматика және ақпараттық технологиялар институты, Сәтбаев университеті, Алматы қ., Қазақстан

Дайнеко Евгения Александровна – ассоциированный профессор, PhD, Институт автоматизации и информационных технологий, Сатпаев Университет, г. Алматы, Казахстан

Daineko Yevgeniya – Associate Professor, PhD, Institute of Automation and Information Technologies, Satpayev University, Almaty, Kazakhstan

e-mail: y.daineko@satbayev.university

ORCID: <https://orcid.org/0000-0001-6581-2622>



Agariadne Dwinggo Samala – Инженерлік факультетінің ассистенті, Негери Паданг университеті, Батыс Суматра, Паданг қ., Индонезия

Agariadne Dwinggo Samala – доцент инженерного факультета Университета Негери Паданг, Западная Суматра, г. Паданг, Индонезия

Agariadne Dwinggo Samala – Assistant Professor at the Faculty of Engineering, Universitas Negeri Padang, West Sumatra, Padang, Indonesia

Email: agariadne@ft.unp.ac.id

ORCID: <https://orcid.org/0000-0002-4425-0605>

https://doi.org/10.51885/3134-8025_IICS_2026_2_7

MPHTI 28.23.15

СРАВНИТЕЛЬНЫЙ АНАЛИЗ И ОЦЕНКА ЭФФЕКТИВНОСТИ СВЕРТОЧНЫХ И ТРАНСФОРМЕРНЫХ АРХИТЕКТУР ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ МНОГОКЛАССОВОЙ КЛАССИФИКАЦИИ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

МЕДИЦИНАЛЫҚ КЕСКІНДЕРДІ КӨП КЛАСТЫ ЖІКТЕУ ҮШІН ТЕРЕҢ ОҚЫТУДЫҢ КОНВОЛЮЦИЯЛЫҚ ЖӘНЕ ТРАНСФОРМЕРЛІК АРХИТЕКТУРАЛАРЫНЫҢ ТИІМДІЛІГІН САЛЫСТЫРМАЛЫ ТАЛДАУ ЖӘНЕ БАҒАЛАУ

COMPARATIVE ANALYSIS AND EVALUATION OF THE EFFECTIVENESS OF CONVOLUTIONAL AND TRANSFORMER DEEP LEARNING ARCHITECTURES FOR MULTI-CLASS CLASSIFICATION OF MEDICAL IMAGES

З.Г. Кабдрахманова ^{1*}, А.С. Тлебадинова ¹, С.К. Кумаргажанова ¹,
М.А. Карменова ², Zbigniew Omiotek ³

¹Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

²Восточно-Казахстанский университет им. С. Аманжолова, г. Усть-Каменогорск, Казахстан

³Люблинский технический университет, г. Люблин, Польша

*Автор-корреспондент: Кабдрахманова Зауре Гадильжановна, e-mail: zkabdrahmanova@vku.edu.kz

Ключевые слова:

иерархический
трансформер зрения,
Swin Transformer,
классификация степени
ожога, анализ
медицинских
изображений, глубокое
обучение, перенос
обучения, компьютерная
диагностика, механизмы
внимания.

АННОТАЦИЯ

Точное определение глубины и площади ожогового поражения – это ключ к правильному выбору метода лечения. Обычно врачи полагаются на визуальную оценку, что зачастую бывает субъективным и может привести к разногласиям между специалистами, а также снизить точность диагностики. Хотя за последние годы активно развиваются методы глубокого обучения для анализа медицинских изображений, потенциал трансформерных моделей в определении степени тяжести ожогов остается недостаточно исследованным.

В нашей работе мы предложили новый подход к автоматической классификации ожогов, основанный на современных архитектурах глубокого обучения. Мы использовали технологии трансферного обучения с предварительно обученными весами, а также применяли аугментацию данных для повышения надежности модели. Для обучения выбрали оптимизатор AdamW и внедрили динамическое управление скоростью обучения, чтобы сделать модель более устойчивой и способной к обобщению. В рамках исследований протестировали разные архитектуры – и сверточные нейронные сети, и трансформеры. Лучший результат показала



© 2026 З.Г. Кабдрахманова, А.С. Тлебадинова, С.К. Кумаргажанова,
М.А. Карменова, Zbigniew Omiotek

Данная работа распространяется на условиях лицензии

Creative Commons «С указанием авторства» 4.0 Международная (CC BY 4.0).

<https://creativecommons.org/licenses/by/4.0/>

трансформерная модель, достигшая точности 94,2 % благодаря способности эффективно учитывать как локальные текстурные детали, так и глобальные контекстные признаки.

Түйінді сөздер:

иерархиялық көру трансформері, Swin Transformer, күйік дәрежесін жіктеу, медициналық кескінді талдау, терең оқыту, оқытуды тасымалдау, компьютерлік диагностика, назар аудару механизмдері.

ТҮЙІНДЕМЕ

Күйіктің тереңдігі мен ауданын дәл анықтау – дұрыс емдеу әдісін таңдаудың кілті. Әдетте дәрігерлер визуалды бағалауға сүйенеді, бұл көбінесе субъективті және мамандар арасындағы келіспеушіліктерге әкелуі мүмкін, сонымен қатар диагностиканың дәлдігін төмендетеді. Медициналық кескіндерді талдау үшін терең оқыту әдістері соңғы жылдары белсенді дамып келе жатқанымен, күйіктердің ауырлығын анықтаудағы трансформаторлық модельдердің әлеуеті әлі зерттелмеген.

Біздің жұмысымызда біз терең оқытудың заманауи архитектураларына негізделген күйіктерді автоматты түрде жіктеудің жаңа әдісін ұсындық. Біз алдын ала дайындалған салмақтармен (веса) трансферлік оқыту технологияларын қолдандық, сонымен қатар модельдің сенімділігін арттыру үшін деректерді күшейтуді қолдандық. Оқыту үшін AdamW оңтайландырушысы таңдалды және модельді тұрақты және жалпылауға қабілетті ету үшін динамикалық оқу жылдамдығын басқаруды енгізді. Зерттеулер шеңберінде әртүрлі архитектуралар – конволюциялық нейрондық желілер де, трансформаторлар да сыналды. Жергілікті текстуралық бөлшектерді де, жаһандық контекстік белгілерді де тиімді есепке алу қабілетінің арқасында 94,2 % дәлдікке жеткен трансформаторлық модель ең жақсы нәтиже көрсетті.

keywords:

hierarchical vision transformer, Swin Transformer, burn severity classification, medical image analysis, deep learning, transfer learning, computer diagnostics, attention mechanisms.

ABSTRACT

Accurate determination of the depth and area of burn damage is key to choosing the right treatment method. Doctors usually rely on visual assessment, which is often subjective and can lead to disagreements between specialists, as well as reduce the accuracy of diagnosis. Although deep learning methods for medical image analysis have been actively developed in recent years, the potential of transformer models in determining the severity of burns remains understudied.

In our work, we proposed a new approach to automatic burn classification based on modern deep learning architectures. We used transfer learning technologies with pre-trained weights and applied data augmentation to improve model reliability. We chose the AdamW optimizer for training and implemented dynamic learning rate control to make the model more stable and capable of generalization. As part of our research, we tested different architectures, including convolutional neural networks and transformers. The transformer model showed the best results, achieving 94.2% accuracy, thanks to its ability to effectively take into account both local texture details and global contextual features.

ВВЕДЕНИЕ

Ожоги являются огромной проблемой для здравоохранения многих стран и всего мирового сообщества из-за их распространенности, сложности терапии и значительного влияния на клинические исходы пациентов. Оценка степени и глубины поражения очень важна для выбора правильных лечебных стратегий, прогноза выздоровления и рационального использования медицинских ресурсов. Однако в настоящее время методы оценки во многом, а иногда и в основном, зависят от зрительного осмотра и опыта врача,

вследствие чего наблюдаются значительная индивидуальная субъективность, а также разногласия между разными специалистами по поводу одних и тех же клинических случаев. Различия в опыте врачей, условиях проведения осмотра, а также исключительные для каждого пациента факторы, могут приводить к несогласованному оцениванию и, как следствие, к непредсказуемым ошибкам в ходе принятия лечебных решений. Все вышеизложенное обуславливает потребность в автоматических и объективных системах поддержки принятия решений, которые позволят врачам проводить более последовательную и надежную оценку тяжести ожогов.

Настоящая работа посвящена вопросам применения глубокого обучения (deep learning) и трансферного обучения (transfer learning) для задачи оценки тяжести кожных ожогов. Обсуждаются особенности и результаты каждого из представленных нами алгоритмов, а также их потенциальная роль в практике диагностики ожоговых травм. По нашим результатам методы трансферного обучения могут быть эффективно использованы в диагностике и, соответственно, внедрены в клиническую практику. В последующем это будет важным шагом на пути к развитию новых инструментов, которые позволят своевременно выявлять пациентов за счет совершенствования методов, настройки гиперпараметров и повышения качества данных.

Обзор литературы. Развитие искусственного интеллекта и глубокого обучения (deep learning) произвело сдвиг в области анализа медицинских изображений, позволив автоматизировать извлечение признаков и распознавание образов, что превосходит традиционные методы ручного анализа. Особенно успешными стали сверточные нейронные сети (Convolutional Neural Networks, CNN), которые применяются в различных задачах медицинской визуализации. Так, в исследованиях D. P. Yadav и соавт. (2022), а также Priya Mittal и соавт. (2024) были продемонстрированы высокий потенциал и надежность сверточных архитектур в извлечении локальных признаков и текстурных особенностей медицинских изображений, а также их способности обеспечивать высокоточную классификацию и сегментацию за счет иерархического извлечения признаков и увеличения рецептивного поля по мере углубления сети.

В последнее время архитектуры на основе трансформаторов (transformer, vit и vit в частном случае) приобрели большую популярность как эффективный механизм самовнимания (self-attention), способный захватывать как локальные, так и глобальные взаимосвязи в визуальных данных. Одним из таких архитектурных решений является Swin Transformer, представляющий собой иерархическую структуру трансформера с окном внимания со сдвигом. Это позволяет достигать хорошей эффективности вычислений за счет размеров окна, а также поддерживать масштабируемость решения для изображений с высоким разрешением.

В трудах Hu Cao и соавт. (2021) и Kumar Anuj и соавт. (2025) рассматривается применение архитектуры Swin Transformer к анализу медицинских изображений, где одной из главных сложностей является необходимость извлечения высокоразмерных признаков при ограниченном объеме размеченных данных. Модель реализует многозадачное обучение, что позволяет успешно решать несколько разных задач медицинской визуализации, одновременно формируя обобщенные представления, применимые к новым данным и к новым задачам. Благодаря адаптивному механизму внимания (attention), архитектура способна фокусировать общее внимание только на значимых частях изображений, что позволяет снизить эффект шума, артефактов и низкоконтрастных изображений.

Кроме того, в работе предложен повторный encoder с иерархической структурой и многоуровневым анализом, что повышает захват локальных и глобальных сценариев. По сути, эксперименты показали, что эффективность модели достигает точности 80,76%.

Таблица 1 содержит обзор некоторых связанных работ (в таблице указаны используемые методы и достигнутая точность классификации изображений ожогов).

Таблица 1. Обзор литературы и оценка методов и моделей глубокого обучения для распознавания ожоговых травм

Авторы, название, год	Методы и модели распознавания ожоговых поражений	Результаты исследования
Hassanipour S. et al., Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: A systematic review and metaanalysis, 2019	ANN и LR	Уровень точности ANN и LR в моделях случайных эффектов составил 90,5 (95 % CI 87,6–94,2) и 83,2 (95 % CI 75,1–91,2) соответственно
Cirillo M. D., Mirdell R., Sjöberg F. et al., Time-independent prediction of burn depth using deep convolutional neural networks, 2019	CNN: VGG-16, GoogleNet, ResNet-50, и ResNet-101	ResNet-101: средняя точность, чувствительность и специфичность для четырех видов ожогов составили 90,54 %, 74,35 % и 94,25 % соответственно
Uğur Şevik, Erdiñç Karakullukçu, Tolga Berber, Yeşim Akbaş, Serdar Türkyılmaz. Automatic classification of skin burn color images using texture-based feature extraction, 2019	SegNet и CNN	Имея F-оценку 74,28 % при классификации изображений, полученных без протокола, предложенная схема смогла достичь аналогичных результатов с глубоким обучением, которое имело F-оценку 80,50 %.
Боб Чжан, Цзяньхан Чжоу Multi-feature representation for burn depth classification using burn images, 2020	Модель Stacked Sparse Autoencoder (SSAE) применяется для извлечения скрытых признаков (latent features) Random Forest (основная модель для классификации)	точность 85,86 % и 76,87 % путем классификации изображений ожогов на два класса и три класса соответственно, что превосходит обычные методы в определении глубины ожога. Результаты показывают, что данный подход эффективен и может помочь медицинским экспертам в определении различной глубины ожогов.
D.P. Yadav, A. S. Jalal, V. Prakash – Human burn depth and grafting prognosis using ResNeXt topology based deep learning network, 2022	Deep learning модель на основе ResNeXt topology; CNN с residual connections для классификации глубины ожога и прогноза grafting	Deep learning модель на основе ResNeXt topology; CNN с residual connections для классификации глубины ожога и прогноза grafting. (Название модели BNeXtmodel, точность 97,17 %)
Ze Liu et al. – Swin Transformer: Hierarchical Vision Transformer using Shifted Windows, 2021	Hierarchical Swin Transformer с механизмом shifted window attention	Top-1 Accuracy = 0,873 (ImageNet benchmark)
Hu Cao et al. – Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation, 2021	Transformer-based encoder-decoder архитектура (Swin-Unet) для медицинских изображений	Transformer-based encoder-decoder architecture (Swin-Unet)
Dosovitskiy et al. – An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale (Vision Transformer), 2021	Vision Transformer (ViT); чистая transformer архитектура для классификации изображений	Top-1 Accuracy = 0,884 (ImageNet pretraining with large-scale data)
Примечание – составлено авторами		

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Предлагаемая в данной работе структура глубокого обучения базируется на предварительно обученной иерархической архитектуре трансформатора зрения и позволяет автоматически классифицировать степень тяжести ожоговых травм. Методология также объединяет перенос обучения, структурированное увеличение количества данных, адаптивные стратегии оптимизации и сложное оценивание с целью обеспечения надежности и обобщающей способности при работе с медицинскими данными. Всю структуру данной работы можно представить в виде следующей последовательности: предварительная обработка данных; извлечение характеристик; обучение с учителем; выбор модели, основываясь на валидации; оценка, основываясь на интерпретируемости. Входной датасет состоит из RGB-изображений, сгруппированных на три класса ожогов. Данные разделены на тренировочные изображения – 70%, валидации – 15% и тестирования – 15% в обучении с учителем. Все изображения были форматированы для приведения к фиксированному двумерному разрешению, чтобы их можно было встраивать в механизм патча основного трансформатора. Так как использовались предобученные модели, обученные на ImageNet, входные изображения были нормализованы с использованием средних и стандартных отклонений набора данных для выравнивания распределения функций между предобученной и медицинской областями. Чтобы повысить обобщение и содействовать противопереобучению, вызванному ограниченностью медицинских данных, во время обучения использовалась стохастическая аугментация данных. Эта стратегия приводит к геометрическим аугментациям, таким как отображение по горизонтали и вращение, а также фотометрическим аугментациям, таким как яркость, контраст и насыщенность. Формально, если x – это наше исходное изображение, T – некоторое случайное преобразование, выбранное из множества преобразований $T \sim T$, а цель обучения задается на основе преобразованных вариантов изображения $\tilde{x} = T(x)$, то возможности разнообразия используемого для обучения набора расширяются.

Ядром модели для извлечения признаков является Swin Transformer иерархическая архитектура трансформатора зрения (vision transformer), которая внедряет локализованное self-attention в сдвинутых оконных областях. В отличие от классических трансформаторов Vision Transformer (ViT), вычисляющих глобальное self-attention между всеми токенами, Swin Transformer ограничивает внимание неперекрывающимися окнами и периодически сдвигает их, чтобы обеспечить взаимодействие между соседними областями изображения. Такой подход снижает вычислительную сложность, сохраняя при этом моделирование долгосрочной зависимости. Входное изображение разбивается на неперекрывающиеся патчи размером $P \times P$, формируя последовательность вложений токенов:

$$Z_0 = \text{Embed}(X) \in \mathbb{R}^{N \times D}, \quad (1)$$

где $N = \frac{HW}{P^2}$ обозначает количество патчей, D – размерность embedding space.

В каждом блоке Swin вычисляется оконное multi-head self-attention, при котором формируются запросы (Q), ключи (K) и значения (V):

$$Q = ZW_Q, K = ZW_K, V = ZW_V, \quad (2)$$

операция self-attention внутри локального окна определяется следующим образом:

$$\text{Attn}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}} + B\right)V, \quad (3)$$

где d – размерность головы внимания, B – относительное позиционное смещение (relative position bias). Каждый блок включает residual connections и feed-forward network (MLP):

$$\begin{aligned} Z' &= Z + W - \text{MSA}(\text{LN}(Z)), \\ Z_{out} &= Z' + \text{MLP}(\text{LN}(Z')). \end{aligned} \quad (4)$$

Иерархическое представление строится с помощью операций объединения патчей, которые уменьшают разрешение по пространству и увеличивают размерность канала, таким образом формируя пирамиду признаков, подобную сверточным нейронным сетям. Головка классификации использует глобальное усреднение и линейный слой:

$$o = W_{cls}g + b_{cls}, \quad (5)$$

где g является агрегированным представлением признаков. Вероятности классов вычисляются с помощью softmax:

$$\hat{y}_c = \frac{e^{o_c}}{\sum_{k=1}^C e^{o_k}}. \quad (6)$$

Целью оптимизации служит категориальная кросс-энтропийная ошибка:

$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log \hat{y}_{i,c}, \quad (7)$$

где θ обозначает параметры модели. Обучение производится с помощью оптимизатора AdamW, который разделяет затухание весов и обновления на основе градиента:

$$\theta_{t+1} = \theta_t - \eta_t \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} - \eta_t \lambda \theta_t. \quad (8)$$

Изменение скорости обучения реализовано с помощью ReduceLROnPlateau, при котором темп снижается, если валидационные потери перестают уменьшаться, что позволяет «переключаться» от грубой к более тонкой настройке.

Вычислительная сложность классического глобального самоуправления непомерна для изображений с высоким разрешением. Swin Transformer снижает сложность до:

$$O(N \cdot M^2), \quad (9)$$

где M обозначает размер окна. Эта стратегия локального внимания значительно улучшает масштабируемость, сохраняя при этом выразительную силу.

Были внесены определённые изменения, направленные на оптимизацию результатов при работе с медицинскими изображениями. Во-первых, перенос обучения с использованием весов, предварительно обученных на ImageNet, даёт возможность эффективно захватывать визуальные признаки низкого и среднего уровней, что способствует быстрому обучению и стабильности при недостаточности данных. Во-вторых, аугментация медицинскими данными обеспечивает инвариантность к нейтральным к диагнозу преобразованиям, сохраняя при этом изменения, специфичные для патологии. В-третьих, адаптивное управление оптимизацией с помощью AdamW совместно с регулировкой скорости обучения на основе результатов валидации стабилизирует процесс обучения. В-четвёртых, выбор модели на основе её точности на валидационном наборе обеспечивает способность к обобщению, а не к переобучению на обучающем наборе. И наконец, использование различных метрик, в том числе матриц погрешностей и мультиклассовых метрик, даёт возможность оценивать модель с учётом различных аспектов, что необходимо для её внедрения в клиническую практику. Оптимизированный процесс обучения отличается от классического Swin Transformer в основном стратегиями оптимизации и предварительной обработки данных, но не архитектурными изменениями. Эти улучшения настраивают модель на специфические особенности медицинских изображений, что и даёт высокий результат при задачах классификации. Возможность Swin Transformer строить иерархические представления, а также использовать локальное внимание, позволяет эффективно улавливать как детали текстур, так и широкие контексты, влияющие на содержание снимков при оценке степени ожогов. Такой подход особенно эффективен при работе с медицинскими текстурами, что и объясняет высокий результат

данного метода на задаче, на которую он был применён. На рисунке 1 показана архитектура Swin трансформера для классификации ожогов.

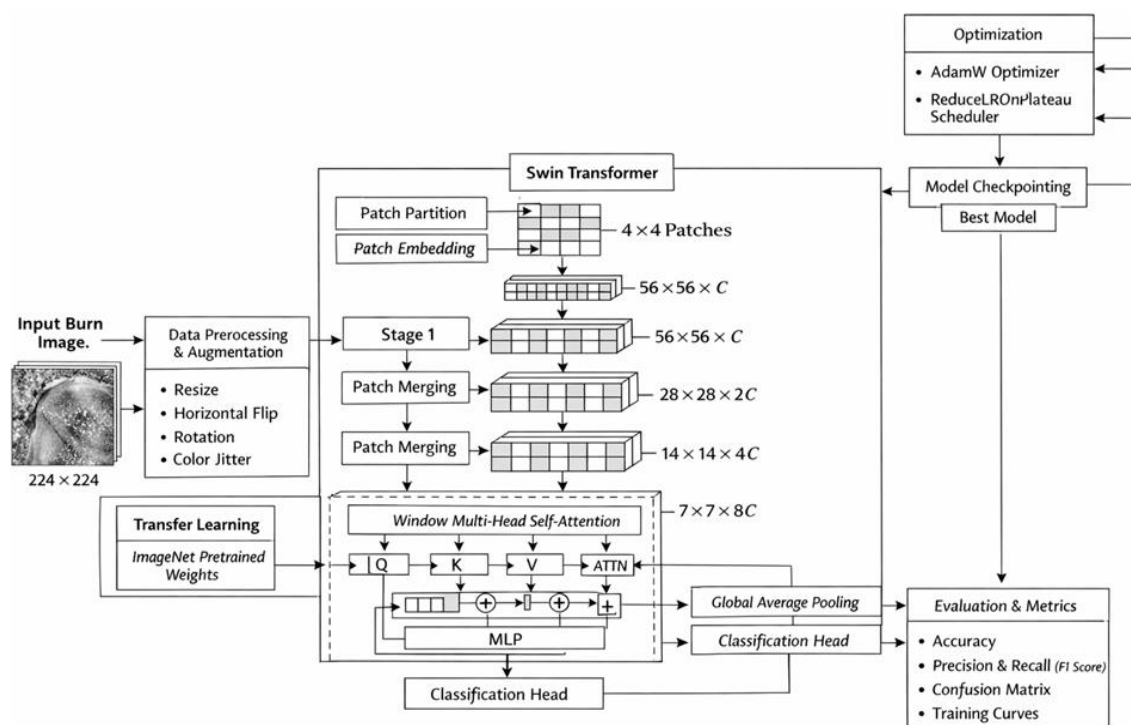


Рисунок 1. Архитектура Swin трансформерной модели,
предлагаемая для классификации ожогов кожи

Примечание – составлено авторами

Реализация алгоритма. Модель классификации ожогов по степени тяжести, предложенная на базе Swin Transformer, была реализована с помощью PyTorch фреймворка глубокого обучения и библиотеки моделей Timm для трансформерных моделей зрения. Разработка и тесты проводились в научной вычислительной среде Spyder, что позволило сделать процессы обучения, оценки и визуализации воспроизводимыми. Данные были разделены на три поднабора для обучения, валидации и тестирования, соответственно, при этом изображения были классифицированы на три категории степени тяжести ожогов. Все входные изображения были приведены к фиксированному разрешению 224×224 пикселей в соответствии с требованиями архитектуры базовой структуры Swin Transformer. Перед обучением изображения были нормализованы на основе статистики ImageNet, чтобы достичь совместимости с предварительно обученными весами модели. Для увеличения обобщающей способности модели и снижения переобучения, что является распространенной проблемой в задачах медицинской визуализации, в процессе обучения использовались техники расширения данных. Среди них были горизонтальное отражение, случайное вращение и преобразования цветового дрожания, которые добавляют небольшую вариативность, сохраняя при этом важные для лечения признаки визуализаций. В этом исследовании была использована базовая модель Swin Transformer Tiny (Swin_Tiny_Patch4_Window7_224), которая была инициализирована весами, предобученными на ImageNet, для применения переноса обучения. Классификационная головка была автоматически перенастроена в соответствии с количеством выходных классов. Оптимизация модели была выполнена с помощью оптимизатора AdamW, который различает затухание весов и обновления градиента для улучшения стабильности сходимости при работе с трансформерными

архитектурами. Обучение модели производилось в течение 100 эпох при размере батча 8 и начальной скорости обучения $2 \cdot 10^{-4}$ и коэффициентом затухания веса $1 \cdot 10^{-4}$. В качестве функции ошибки при обучении использовалась категориальная кросс-энтропия. Для динамического корректирования значения скорости обучения применялась стратегия планирования ReduceLROnPlateau, которая уменьшает скорость обучения при отсутствии улучшений значения функции потерь на валидационном наборе. Обновление значения скорости обучения может быть выражено следующим уравнением:

$$\eta_{t+1} = \begin{cases} \gamma \cdot \eta_t, & \text{if } \mathcal{L}_{val} \text{ не улучшается} \\ \eta_t, & \text{иначе,} \end{cases} \quad (10)$$

где η_t обозначает скорость обучения при итерации t , γ – коэффициент затухания (установлен на 0,3), а $\mathcal{L}_{val}$ представляет собой потерю валидации.

Во время обучения по мере улучшения метрик обучения и валидации, таких как точность и потери, лучшая модель сохранялась на диск. На этапе тестирования были собраны и рассчитаны значения всех метрик с accuracy, precision, recall, F1-score и построена матрица неточностей для более глубокого анализа поведения модели в отношении разных классов. Все эксперименты и тренировки были проведены на системе с процессорами Intel(R) Core (TM) i5-12500H 12-го поколения с тактовой частотой 2, 50 ГГц, 8 ГБ оперативной памяти и 64-разрядной операционной системой x64.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Результаты экспериментов показывают, что при учёте и локальной текстуры, и глобальных контекстных связей в медицинских изображениях архитектуры на базе трансформеров оказываются эффективнее, чем традиционные свёрточные архитектуры. В эксперименте использовались определенные конфигурации моделей и стратегия обучения; основные гиперпараметры, использованные в данной работе, приведены в табл. 2.

Таблица 2. Конфигурация модели и гиперпараметры обучения для архитектуры Swin Transformer

Категория	Параметр	Значение	Описание
Конфигурация архитектуры	Базовая модель (Backbone)	swin_tiny	Hierarchical vision transformer with shifted windows attention
	Размер входного изображения	224×224	Стандартизованный размер входного изображения
	Размер патча	4×4	Размер патча для разбиения изображения
	Размер окна	7×7	Размер окна локального self-attention
	Размер embedding	96	Начальный размер embedding
	Количество иерархических стадий	4	Глубина иерархической обработки
	Предобучение	weights ImageNet	Использование предобученной модели на ImageNet
Параметры оптимизации	Оптимизатор	AdamW	Используемый оптимизатор
	Начальная скорость обучения	2×10^{-4}	Начальная скорость обучения
	Weight decay	1×10^{-4}	Коэффициент регуляризации
	Декремент скорости обучения	ReduceLROnPlateau	Снижение скорости обучения при отсутствии улучшения метрики

Окончание таблицы 2

Категория	Параметр	Значение	Описание
	Коэффициент снижения LR	0.3	Коэффициент уменьшения learning rate
	Минимальный learning rate	1×10^{-6}	Нижняя граница learning rate
Настройки обучения	Размер batch	8	Размер batch при обучении
	Количество эпох	100	Максимальное количество эпох при обучении
	Функция потерь	CrossEntropyLoss	Тип функции потерь
	Random seed	42	Случайное зерно для воспроизводимости результатов
Процедура предобработки данных	Resize	224×224	Изменение размера изображений
	Normalize	mean/std of ImageNet	Используются те же средние значения и стандарты отклонения, что и при предобучении на ImageNet
Аугментация данных	Горизонтальное отражение	Включено	Для увеличения объема тренировочных данных
	Случайное вращение	$\pm 10^\circ$	Для увеличения объема тренировочных данных
	Color jitter	Brightness (0,15), Contrast (0,15), Saturation (0,1), Hue (0,02)	Для увеличения объема тренировочных данных
Оценка	Метрики	Accuracy, Precision, Recall, F1-score	Множество метрик для оценки качества классификации
	Матрица ошибок	Используется	Для изучения характера ошибок классификатора
Примечание – составлено авторами			

Для оценки надежности и воспроизводимости результатов модели их модель была обучена 3 раза с различными случайными инициализациями и разными random seed. Результаты представлены в виде средних значений с соответствующими стандартными отклонениями. Низкая изменчивость при повторных запусках указывает на стабильность и устойчивость модели к случайным факторам обучения.

Модели, обученные в этом эксперименте, были оценены с использованием количественных метрик производительности для классификационной эффективности по всем категориям. Результаты экспериментов предоставляют сравнительный обзор для выделения моделей, улучшая при этом набор критериев точности классификации, precision, recall и F1-score. Таким образом, можно более объективно оценить, насколько различные архитектурные решения влияют на предсказательную эффективность в рамках решения конкретной задачи. Accuracy - это процент объектов, по которым правильно предсказана их категория из общего числа примеров. В задачах многоклассовой классификации показатель accuracy (ACC) вычисляется так же, как и в бинарном случае и отражает долю правильно классифицированных образцов от общего числа наблюдений. При этом каждый образец вносит равный вклад в итоговое значение метрики. Такой показатель позволяет оценить, как справляется модель в целом, а также получить представление о её поведении по отдельным классам. Особенно актуально это, если классы несбалансированы или ошибки классификации имеют разную важность.

$$\text{Accuracy: ACC} = \frac{TP+TN}{TP+TN+FP+FN}, \quad (11)$$

где: TP – количество правильно предсказанных положительных результатов; TN – количество правильно предсказанных отрицательных результатов; FP – количество неправильно предсказанных положительных результатов; FN – количество неправильно предсказанных отрицательных результатов.

Точность определяется как доля объектов, которые классифицированы как положительные и действительно являются положительными. В задаче многоклассовой классификации объекты классифицируются более чем по двум классам (степени ожогов: легкая, средняя, тяжелая). Для каждого класса k метрики рассчитываются отдельно в соответствии с принципом «один против всех». Точность для класса k определяется как отношение числа объектов, корректно отнесённых к классу k , к общему количеству объектов, предсказанных моделью как принадлежащих классу k :

$$\text{Precision: } PRE_k = \frac{TP_k}{TP_k + FP_k}, \quad (12)$$

Возврат для класса k – это доля объектов в классе k , которые были правильно распознаны.

$$\text{Recall: } REC_k = \frac{TP_k}{TP_k + FN_k}, \quad (13)$$

F1 – показатель для класса k , является средним гармоническим показателей точности и полноты для класса k :

$$\text{F1-score: } F1_k = 2 \frac{PRE_k \cdot REC_k}{PRE_k + REC_k}. \quad (14)$$

Числовые результаты экспериментов приведены в следующей таблице.

Таблица 3. Сравнение результатов использования различных моделей глубокого обучения для задачи классификации ожогов кожи

№	Название модели	Accuracy	Precision	Recall	F1-score
1	Xception	0.7565	0.7607	0.7562	0.7468
2	VGG16	0.9188	0.9193	0.9187	0.9189
3	VGG19	0.9041	0.9053	0.9040	0.9044
4	ResNet50	0.8967	0.8992	0.8966	0.8973
5	DenseNet201	0.8708	0.8769	0.8708	0.8723
6	MobileNetV2	0.8469	0.8467	0.8468	0.8459
7	Inception_resnet_v2	0.8704	0.8804	0.8704	0.8701
8	EffNetB0	0.7915	0.7929	0.7911	0.7903
9	AttnMS_VGG16	0.9041	0.9052	0.9040	0.9044
10	Vgg16_TTA	0.9004	0.9053	0.9004	0.9015
11	Vit_TTA	0.9059	0.9071	0.9058	0.9063
12	Convnext_tiny	0.7897	0.8018	0.7897	0.7902
13	Swin_tiny	0.9428	0.9426	0.9427	0.9425

Примечание – составлено авторами

Сравнительный анализ показывает, что архитектуры, основанные на трансформерах, лучше традиционных сверточных нейронных сетей справляются с классификацией степени тяжести ожогов. Из всех моделей наилучшие результаты были у иерархического Swin Transformer, что подтверждает его способность эффективно улавливать как локальные текстуры, так и общие контекстные связи в медицинских изображениях. Классические архитектуры CNN, такие как VGG16 и VGG19, показали хорошие базовые результаты и подтвердили свою эффективность в задачах классификации на основе текстуры в медицине, однако они были слегка уступлены моделям на трансформерах.

Легкие и ориентированные на эффективность модели, такие как MobileNetV2 и ConvNext_tiny, показали сравнительно низкую производительность, что, вероятно, связано с тем, что снижение сложности модели ограничивает ее возможности представить данные. Варианты с улучшенным вниманием и стратегия увеличения времени тестирования повысили стабильность и надежность, что еще раз подчеркивает важность применения передовых методов для характеристики моделей.

Качество моделей классификации оценивалось с помощью теста МакНемара, применяемого для сравнения парных прогнозов на одном и том же тестовом наборе данных. В частности, были сравнены модели Swin Transformer и VGG16, показавшие максимальную точность. Тест был проведен на основе анализа несовпадающих пар прогнозов. Статистическая значимость признавалась при $p < 0,05$.

Ниже представлена сравнительная диаграмма оценок метрик для каждой модели.

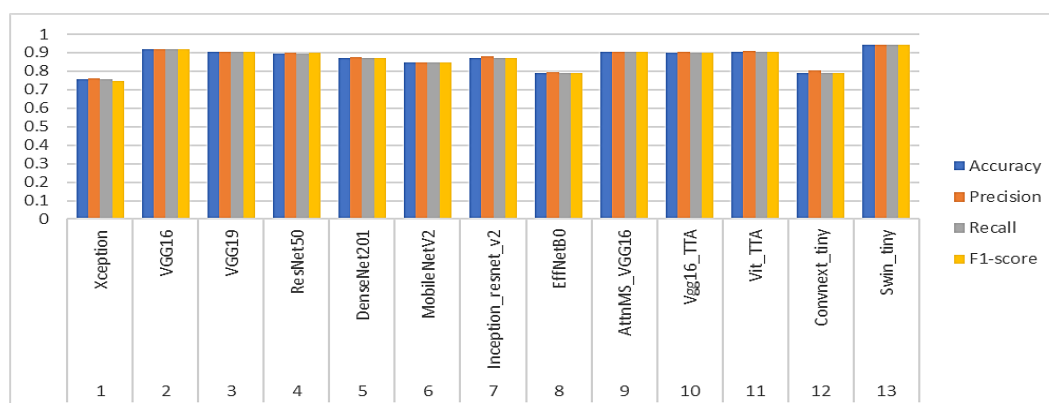


Рисунок 2. Сравнительная диаграмма показателей

Примечание – составлено авторами

Результаты, показанные на рисунке 3 (а), подтверждаются и диаграммой ошибок, на которой видно, что классификация проведена качественно, о чем свидетельствуют многочисленные правильные прогнозы на главной диагонали для всех трех классов. Особенно хорошо выявляются классы 1 и 3: они с высокой вероятностью определяются верно, незначительно перепутываются, между ними не происходит прямого замешательства, что говорит о хорошей отделимости их признаков. Большинство ошибок связаны со вторым классом: он перепутан с первым, порой с третьим, что указывает на то, что в признаках этого класса или в промежуточных признаках соседних классов есть что-то общее. Несмотря на эти проблемы, общее количество ошибок невелико, что говорит о том, что модель способна обобщать знания по сбалансированным классам, правильно выделяя как характерные особенности, так и контекст. Однако до сих пор модель не может хорошо отделить второй класс от двух других. В целом же по матрице ошибок можно сказать, что модель справляется с нечеткими ситуациями и делает сбалансированный выбор, что говорит о возможности использовать ее для достаточно надежной классификации медицинских изображений в нескольких классах.

Рисунок 3 (б) показывает кривые обучения, по которым мы можем следить за изменением значений функций потерь и достигаемых точностей в процессе обучения с возрастанием числа эпох. Потери при обучении уменьшаются очень быстро, приближаясь к нулю в течение нескольких эпох и тем самым доказывая, что модель успешно адаптируется под обучающую выборку.

Одновременно с этим точность при обучении тоже быстро растёт и достигает очень высокого значения, демонстрируя высокую способность модели к обучению.

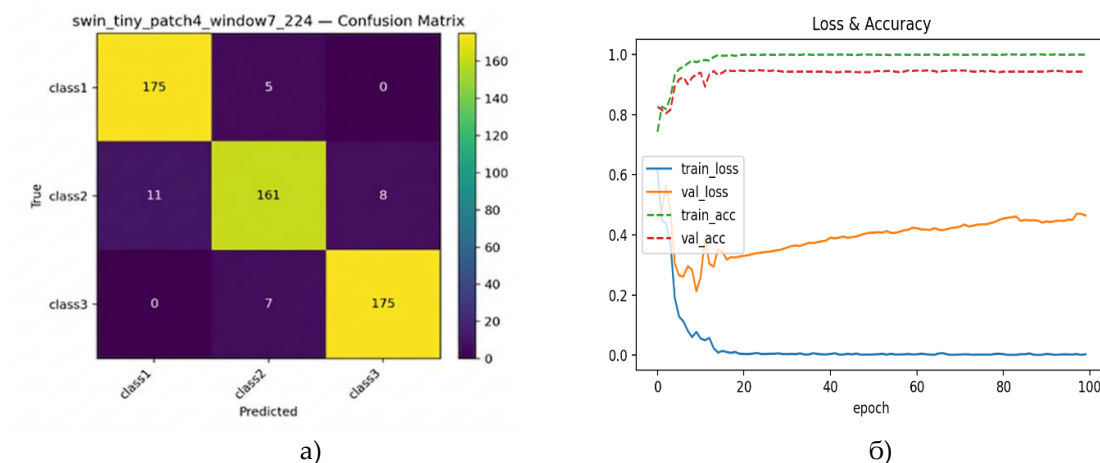


Рисунок 3. Матрица ошибок и loss модели Swin tiny

Примечание – составлено на основе расчетов авторов

ЗАКЛЮЧЕНИЕ

В данной статье представлены результаты экспериментов с использованием глубокого обучения при автоматическом определении степени тяжести ожогов на основе медицинских изображений, собранных из общедоступных источников. Сбалансированный набор из трех классов, представляющих разные степени тяжести ожогов, был собран из открытых репозиторий, таких как Kaggle и Roboflow. Для каждого класса было отобрано ровно 1200 образцов, из них около 840-848 изображений на класс были использованы для обучения, остальные изображения были отложены для валидации и тестирования модели. Благодаря сбалансированной структуре набора данных результаты модели можно оценивать объективно. Данный баланс минимизирует проблему смещения модели из-за дисбаланса классов во время обучения.

Были протестированы тринадцать архитектур глубокого обучения. Рассматривались классические сверточные нейронные сети (VGG16, VGG19, ResNet50, DenseNet201, MobileNetV2, Xception и Inception-based модели) и новые сверточные архитектуры (ConvNeXt), а также модель с улучшенным вниманием и модели на основе трансформеров (Vision Transformer (ViT) и Swin Transformer). Такая работа позволила детально проанализировать, как различные архитектурные схемы влияют на качество решения задачи классификации степени ожога по медицинским изображениям. Для обеспечения эффективности и стабильности модели был внедрен ряд стратегий и методов оптимизации. В том числе: перенос обучения, при котором используются веса предварительно обученной модели; увеличение объема данных с учетом домена; планирование скорости обучения с теплым перезапуском; улучшение характеристик с помощью внимания; увеличение объема данных во время тестирования для улучшения способности к обобщению.

Эффективность моделей также оценивалась с помощью комплексных показателей и метрик, таких как общая точность, точность, воспроизводимость, показатель F1 и анализ матрицы путаницы. Результаты экспериментов показывают, что учет как локальной текстуры, так и глобальных контекстуальных отношений в медицинских изображениях с помощью архитектур на основе трансформаторов превосходит по эффективности традиционные архитектуры на основе свертки. Вклад данной статьи заключается в разработке интеллектуальных систем поддержки принятия клинических решений путем предоставления структурированного подхода к оценке современных архитектур глубокого

обучения для классификации глубины ожогов. Новизна работы заключается в том, что в ней исследуется сравнительная эффективность различных глубоких архитектур, включая CNN и модели на основе внимания, для классификации степени повреждения с использованием цветных медицинских изображений. Были предложены соответствующие стратегии обучения для оптимизации производительности для конкретных задач, включая настройку гиперпараметров, аугментацию данных, интеграцию моделей на основе трансформаторов, реализованных в сочетании с оптимизированными стратегиями обучения, иерархические механизмы внимания, адаптированные к области задачи анализа ожогов. Результаты, полученные в этой работе, являются алгоритмической основой и научно-методическим руководством для последующей разработки информационной системы для поддержки диагностики и лечения ожогов.

В итоге результаты показывают, что архитектуры на основе трансформеров являются многообещающим направлением для повышения точности, объективности и надежности автоматической классификации степени ожогов и, несомненно, будут стимулировать будущие исследования в области создания клинически применимых систем искусственного интеллекта.

КОНФЛИКТ ИНТЕРЕСОВ: Авторы заявляют об отсутствии конфликта интересов.

ФИНАНСИРОВАНИЕ: Работа выполнена при поддержке финансирования проекта №AP23486396 «Модели и методы распознавания анатомических структур на изображениях МРТ в задачах компьютерной диагностики» на 2024–2026 годы Министерства науки и высшего образования Республики Казахстан.

ЗАЯВЛЕНИЕ ОБ ОДОБРЕНИИ ИНСТИТУЦИОНАЛЬНЫМ ЭТИЧЕСКИМ КОМИТЕТОМ (IRB): Данное исследование имеет исключительно техническую направленность и не предусматривает участие людей, животных или проведение биологических экспериментов, поскольку все эксперименты выполняются исключительно на изображениях ожоговых ран, в связи с этим данный раздел отмечен как «не применимо»

ЗАЯВЛЕНИЕ ОБ ИНФОРМИРОВАННОМ СОГЛАСИИ: *Так как все эксперименты выполняются исключительно на изображениях ожоговых поражений, данный раздел считается «не применимым»*

ЗАЯВЛЕНИЕ О ДОСТУПНОСТИ ДАННЫХ: Данные, использованные в этом исследовании, были получены из открытых источников, таких как Kaggle и Roboflow, и находятся в свободном доступе для использования. Все наборы данных могут быть загружены и использованы без ограничений. В рамках данного исследования новые экспериментальные данные не создавались.

БЛАГОДАРНОСТИ: Авторы выражают признательность коллегам за методологическую поддержку и конструктивные обсуждения, а также анонимным рецензентам за ценные замечания, способствовавшие повышению качества представленной работы.

УВЕДОМЛЕНИЕ ОБ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: В данной публикации не использовались инструменты искусственного интеллекта (ИИ)

СПИСОК ЛИТЕРАТУРЫ

- Rozo, A., Miskovic, V., Rose, T., Keersebilck, E., Iorio, C., & Varon, C. (2023). A deep learning image-to-image translation approach for a more accessible estimator of the healing time of burns. *Artificial Intelligence in Medicine*, 140. <https://doi.org/10.1016/j.artmed.2023.102570>
- Yadav, D. P., Jalal, A. S., & Prakash, V. (2022). Human burn depth and grafting prognosis using

- ResNeXt topology based deep learning network. Multimedia Tools and Applications, 81(13), 18897-18914. <https://doi.org/10.1007/s11042-022-12446-0>
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (ICLR). <https://doi.org/10.48550/arXiv.1409.1556>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2016.90>
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2017.195>
- Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2017.243>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR.2018.00474>
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4 Inception-ResNet and the impact of residual connections on learning. In AAAI Conference on Artificial Intelligence. <https://doi.org/10.1609/aaai.v31i1.11231>
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning (ICML). <https://doi.org/10.48550/arXiv.1905.11946>
- Dosovitskiy, A., et al. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. In International Conference on Learning Representations (ICLR). <https://doi.org/10.48550/arXiv.2010.11929>
- Liu, Z., et al. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In IEEE International Conference on Computer Vision (ICCV). <https://doi.org/10.1109/ICCV48922.2021.00986>
- Liu, Z., et al. (2022). ConvNeXt: A ConvNet for the 2020s. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/CVPR52688.2022.01167>
- Zhang, B., & Zhou, J. (2021). Multi-feature representation for burn depth classification using burn images. Artificial Intelligence Review. <https://doi.org/10.1007/s10462-021-10055-6>
- Kumar, A. (2025). An improved feature extraction algorithm for robust Swin transformer model in high-dimensional medical image analysis. Computers in Biology and Medicine, 188. <https://doi.org/10.1016/j.combiomed.2025.109822>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Кабдрахманова Зауре Гадильжановна – «Ақпараттық жүйелер» білім беру бағдарламасының докторанты, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан.
Кабдрахманова Зауре Гадильжановна – докторант образовательной программы «Информационные системы», Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан.

Kabdrakhmanova Zaure Gadylzhanovna – doctoral student of the educational program "Information Systems", D. Serikbaev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan,
e-mail: zkabdrakhmanova@vku.edu.kz,
ORCID: <https://orcid.org/0000-0003-4763-9437>



Тлебалдинова Айжан Солтангалиевна – PhD, қауымдастырылған профессор, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен, Қазақстан.

Тлебалдинова Айжан Солтангалиевна – PhD, ассоциированный профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева, Усть-Каменогорск, Казахстан.

Tlebaldinova Aizhan – PhD, associate professor, D. Serikbaev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan
e-mail: atlebaldinova@edu.ektu.kz,
ORCID: <https://orcid.org/0000-0003-1271-0352>



Кумаргажанова Сауле Кумаргажановна – техникалық ғылымдар кандидаты, профессор, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан.

Кумаргажанова Сауле Кумаргажановна – кандидат технических наук, профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан.

Kumargazhanova Saule Kumargazhanovna – candidate of Technical Sciences, professor, D. Serikbayev East Kazakhstan Technical University, Oskemen, Kazakhstan,
E-mail: skumargazhanova@gmail.com,
ORCID: <https://orcid.org/0000-0002-6744-4023>



Карменова Мархаба Ахметоллиновна – PhD, қауымдастырылған профессор, С.Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен қ., Қазақстан.

Карменова Мархаба Ахметоллиновна – PhD, ассоциированный профессор, Восточно-Казахстанский университет им. С. Аманжолова, г. Усть-Каменогорск, Казахстан.

Karmenova Markhaba – PhD, Associate Professor, S. Amanzholov East Kazakhstan University, Ust-Kamenogorsk, Kazakhstan.
e-mail: mkarmenova@vku.edu.kz,
ORCID: <https://orcid.org/0000-0003-1271-0352>



Zbigniew Omiotek – қауымдастырылған профессор, Люблин технологиялық университеті, Люблин, Польша.

Zbigniew Omiotek – ассоциированный профессор, Люблинский Технологический Университет, Люблин, Польша.

Zbigniew Omiotek – associate Professor, Lublin University of Technology, Lublin, Poland,
e-mail: z.omiotek@pollub.pl,
ORCID: <https://orcid.org/0000-0002-6614-7799>

https://doi.org/10.51885/3134-8025_IJCS_2026_2_8

XFTAP 28.23.00

МУЛЬТИМОДАЛЬДЫ СИГНАЛДАРДАҒЫ ЭМОЦИЯЛАРДЫҢ УАҚЫТША ТҰРАҚТЫЛЫҒЫН БАҒАЛАУ

ОЦЕНКА ВРЕМЕННОЙ СТАБИЛЬНОСТИ ЭМОЦИЙ В МУЛЬТИМОДАЛЬНЫХ СИГНАЛАХ

EVALUATION OF TEMPORAL STABILITY OF EMOTIONS IN MULTIMODAL SIGNALS

С.Ә. Кудубаева ^{1,2}, Е. Сексенбаев ², Ә.Н. Дайырбаева ^{1,3*}, А.С. Еримбетова ¹,
Ж.А. Бурибаев ^{1,4}

¹ҚР ҒЖБМҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан

²Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана қ., Қазақстан

³Қ.И.Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы қ., Қазақстан

⁴Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

*Жауапты автор: Дайырбаева Эльмира Нурбекқызы, e-mail: nurbekkyzelmira@gmail.com

Түйінді сөздер:

Эмоцияны
мультимодальды тану,
уақытша тұрақтылық,
кросс-модальды зейін,
трансформаторлар,
уақытша консистенцияны
жүйелеу, Херст
көрсеткіші,
автокорреляция,
бұрыштық жылдамдық
(траектория), шартты
кездейсоқ өрістер(CRF),
модальділік-беріктіктің
болмауы.

ТҮЙІНДЕМЕ

Бұл мақалада мультимодальды мәліметтердегі эмоциялық күйлердің уақытша тұрақтылығын бағалауға арналған әдістемелік тәсіл ұсынылады. Басқа жұмыстардан айырмашылығы, біз көп деңгейлі формализацияны (сигнал → функция → жасырын кеңістік → болжау) және көрсеткіштер жиынтығын (класс ауыстырғыштарының саны, эпизодтардың ұзақтығы және Херст көрсеткіші) енгіземіз. Тұрақтылық «Плутчик» моделі негізінде түсіндірілетін полярлық кеңістіктегі траекториялар деңгейінде бағаланады. Жұмыста уақытша тұрақтылықты сипаттайтын көрсеткіштер жүйесі ұсынылып, олар бірнеше мультимодальды архитектура-ларда салыстырмалы түрде талданады. Бағалау хаттамасына динамиктерден тәуелсіз бөлімдер, стандартты көрсеткіштер (F1/UAR, CCC/MAE) және уақытша тұрақтылық көрсеткіштері енгізілді. Қосымша ретінде MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығында text-only pilot-эксперимент жүргізілді. Ұсынылған бағалау тәсілі модельдердің уақыт бойынша тұрақтылық қасиеттерін дәлірек сипаттауға мүмкіндік береді. Практикалық құндылық адам мен компьютердің өзара әрекеттесу жүйелеріндегі және көмекші технологиялардағы мінез-құлықтың болжамдылығы мен түсіндірілуін арттыруда жатыр.

Ключевые слова:

Распознавание эмоций в
мультимодальных
данных, временная
стабильность, кросс-
модальное внимание,

АННОТАЦИЯ

В данной статье предлагается методический подход к оценке временной стабильности эмоциональных состояний в мультимодальных данных. В отличие от других работ, мы вводим многоуровневую формализацию (сигнал → функция → скрытое пространство → предсказание) и набор показателей, включающий



трансформеры,
систематизация
временной
консистентности,
показатель Херста,
автокорреляция, угловая
скорость (траектория),
условные случайные поля
(CRF), отсутствие
модальной устойчивости

число переключений классов, длительность эпизодов и показатель Херста. Стабильность оценивается на уровне траекторий в полярном пространстве, интерпретируемом на основе модели Плутчика. В работе предлагается система показателей, описывающих временную стабильность, и проводится их сравнительный анализ на нескольких мультимодальных архитектурах. Протокол оценки включает разбиения, независимые от дикторов, стандартные метрики (F1/UAR, CCC/MAE) и показатели временной стабильности. В качестве дополнительной проверки был проведён text-only pilot-эксперимент на сбалансированной небольшой подвыборке датасета MELD. Предложенный подход к оценке позволяет точнее характеризовать свойства временной устойчивости моделей. Практическая ценность работы заключается в повышении предсказуемости и интерпретируемости поведения систем человеко-компьютерного взаимодействия и вспомогательных технологий.

keywords:

Multimodal emotion
recognition, temporal
stability, cross-modal
attention, transformers,
systematization of temporal
consistency, Hurst
exponent, autocorrelation,
angular velocity
(trajectory), conditional
random fields (CRF), lack of
modality robustness.

ABSTRACT

This paper proposes a methodological approach for evaluating the temporal stability of emotional states in multimodal data. Unlike previous studies, we introduce a multi-level formalization (signal → function → latent space → prediction) and a set of indicators including the number of class switches, episode duration, and the Hurst exponent. Stability is assessed at the level of trajectories in a polar space interpreted on the basis of Plutchik's model. The study proposes a system of metrics describing temporal stability and provides a comparative analysis of these metrics across several multimodal architectures. The evaluation protocol includes speaker-independent splits, standard metrics (F1/UAR, CCC/MAE), and temporal stability indicators. As an additional validation step, a text-only pilot experiment was conducted on a balanced small subset of the MELD dataset. The proposed evaluation approach makes it possible to characterize the temporal stability properties of models more accurately. The practical value of the work lies in improving the predictability and interpretability of behavior in human-computer interaction systems and assistive technologies.

КІРІСПЕ

Эмоцияларды тану – адам мен машина арасындағы өзара әрекеттесуді табиғи әрі тиімді ету үшін маңызды зерттеу бағыты. Соңғы жылдары эмоцияны тану әдістері бейне, аудио және мәтін сияқты әртүрлі деректер көздерін қамтитын мультимодальды тәсілдерге негізделіп дамып келеді. Дегенмен, эмоциялар уақыт бойынша тұрақты құбылыс емес: олар бір сәтте өзгеріп, әртүрлі қарқындылықта көрінуі мүмкін екендігі көптеген зерттеулерде көрсетілген. Сондықтан эмоциялардың уақытша тұрақтылығын бағалау эмоцияны дұрыс модельдеу мен сенімді тануда шешуші рөл атқарады. Бұл мәселе эмоционалдық күйлердің динамикасын сипаттайтын көрсеткіштерді (мысалы, ауысулар саны, эпизод ұзақтығы, автокорреляция, Херст көрсеткіші) қолдануды талап етеді. Жоғарыда келтірілген тәсілдер тек тану дәлдігін арттырып қана қоймай, эмоциялардың табиғи ағымын ескеруге мүмкіндік береді, соның нәтижесінде мультимодальды эмоцияларды талдау жүйелері неғұрлым сенімді әрі түсінікті бола түседі. Мақала мазмұнын ашу барысында мультимодальды эмоцияларды талдау жүйелеріне негізделген зерттеулерге талдау жасадық.

Jayasinghe, Wong және Nugaliyadde (2026) өз кезегінде Speech Emotion Recognition (SER) – адамның сөйлеуінен эмоционалды күйлерді анықтау әдісін қарастырды. Негізінен,

Automatic SER (ASER) саласында эмоцияны болжауда қолданылатын машиналық оқыту (ML) тәсілдерінің интерпретабельділігі мен түсіндірілуіне назар аударылды. Авторлар ASER модельдерінің сенімділігі мен ашықтығы психикалық денсаулықты бақылау сияқты нақты қолданбалар үшін маңызды екенін көрсетті. Сол сияқты бағалау метрикаларын стандарттау, мультимодальды деректерді біріктіру және модельдердің тұрақтылығын арттыру мәселелері әлі толық шешілмегені де мақалада көрсетілген. Осылайша, түйіндейтін болсақ, интерпретабельділік пен түсіндіруді жетілдірудің ASER жүйелерін сенімді енгізу үшін шешуші маңызға ие екенін көрсетілген.

Қарастырып отырған келесі мақалада (Wang және басқ., 2026) мультимедиялық технологиялардың қарқынды дамуы адамдардың интернетте өз пікірін бейнелер арқылы жиі жеткізуіне ықпал еткені туралы айтылған. Осыған байланысты қазіргі таңда эмоцияларды визуалды талдау (VEA) барлық әлем зерттеушілерінің назарын аударған. Эмоциялар субъективті құбылыс болғандықтан, олар бейненің түсі, жарықтығы, бет әлпеті, объектілері мен көріністері сияқты түрлі атрибуттарға тәуелді болатындығын ескеріміз керек. Осы мәселені шешу үшін авторлар эмоцияға қатысты атрибуттық ақпаратты кешенді түрде алуға арналған атрибуттарға негізделген терең оқыту желісін (AttDNet) ұсынды. Ұсынылған желі бірнеше модульдерден тұратын атрибуттық тармақ пен бет әлпеті туралы ақпарат алатын тармақтан құралады. Атрибуттық тармақтың модульдері әртүрлі ақпаратты өңдейді, атап айтсақ CSCAM модулі бейненің түсі мен негізгі объектілерін, LEconv модулі объектілер контекстін, ал MLPooling модулі көрініс туралы мәліметтерді шығарады. Сонымен қатар, бір эмоциялық категорияға жататын үлгілерді жақындату үшін CenterLoss әдісін қолданған.

Zhang және басқ. (2026) мақаласында мультимодальды сарказмды анықтау мәселесін қарастырды. Аталған бағыттағы негізгі қиындық – әртүрлі модальдар арасындағы өзара әрекеттесуді толық ескеру қажеттілігі болып табылады. Авторлар сарказмды тануда тақырыптардың маңызды рөл атқаратынын негізге ала отырып, тақырыпқа негізделген енгізу мүмкіндіктерін жетілдіру әдісін ұсынды. Ұсынылған әдіс екі деректер жиынтығында (бейне негізіндегі және твиттер негізіндегі) және он екі негізгі мультимодальды сарказмды анықтау моделінде сынақтан өткізілген. Эксперименттік нәтижелер барлық базалық модельдерде көрсеткіштердің едәуір жақсарғанын көрсетті. Осылайша, авторлар ұсынған тәсілдің мультимодальды сарказмды анықтауда тиімділігін дәлелдеген.

Pradhan және Srivastava (2026) мультимодальды физиологиялық деректер негізінде эмоцияларды тануға арналған жаңа модельді ұсынған. Алдымен деректер арнайы сүзгілеу әдістері арқылы тазартылып, кейін SCA-GAN генеративті желісі арқылы аугментация жүргізілген. Модель кеңістікке бағытталған назар (SA) мен арнаға бағытталған назарды (CA) қолданып, деректерді теңгерімдеуді және ерекшеліктерді толық қамтуды қамтамасыз етті. Жүргізілген зерттеудің соңғы этапында эмоциялар HPC-CNN гибридті нейрондық желісі арқылы жіктелген. Зерттеу барысында ашық WESAD деректер жиынтығы пайдаланылды. Нәтижелер ұсынылған әдістің жоғары тиімділігін көрсетті: дәлдік – 99,4 %, нақтылық – 98,04 %, ерекшелік – 99,56 %, F1 көрсеткіші – 98,44 %, ал сезімталдық – 98,84 % болды. Қорытындылай келе, (Pradhan және Srivastava, 2026) ұсынған тәсіл мультимодальды физиологиялық сигналдар арқылы эмоцияларды сенімді әрі дәл анықтауға мүмкіндік беретінін дәлелдеді.

Қарастырылған (Qi, Huang және Zhang, 2022) жұмыста сөйлеу эмоцияларын тануға арналған әдіс ретінде CapsNet капсульді нейрондық желісі қолданылған. Модельдің тиімділігін арттыру үшін сөйлеу деректеріне «ақ» шу қосу және дауыс өзгерту арқылы аугментация жүргізілген. Жүйеге енгізілген параметрлерге MEL-спектр бейнелерімен қатар дыбыс көзі, дауыс жолы сияқты ерекшеліктер енгізілген. Эксперимент нәтижелері көрсеткендей, неміс тіліндегі EMODB корпусында төрт эмоцияны (бейтарап, зеріккен, ашу,

қуаныш) тану дәлдігі 99,69%, ал вьетнам тіліндегі ВКЕмо корпусында (бейтарап, мұң, ашу, қуаныш) 94,23% болды. Ең жоғары көрсеткіштер MEL-спектр бейнелерін негізгі жиілікке қатысты сипаттамалармен біріктіргенде алынды.

Van, Nguyen және Le (2022) жұмысында эмоцияларды тануға арналған капсулалық нейрондық желі негізіндегі тәсіл қарастырылған. Авторлар модельдің эмоциялық белгілерді ықшам әрі тиімді ұсыну мүмкіндігін көрсетіп, оның дәстүрлі архитектуралармен салыстырғанда перспективалы екендігін негіздейді.

Авторлар (Liu және басқ. 2023) сөйлеу эмоцияларын танудағы деректердің әртүрлі адамдар арасында тең бөлінбеу мәселесін шешу үшін теңгерімді аугментацияланған үлгілерді қолданды. Бұл тәсіл үш арналы LOG-MEL спектрограммаларын негізге алып, эмоциялық категориялар арасындағы үлгі теңгерімсіздігін азайтып, нейрондық желіге жеткілікті деректер ұсынуға мүмкіндік берді. Ерекшеліктерді әртараптандыру үшін уақыттық және жиіліктік сүзгілер қолданылды. Модельде белгілерді алу үшін CNN мен назарға негізделген екі бағытты ұзақ-қысқа мерзімді жады желісі (ABLSTM) біріктірілді, ал өнімділікті арттыру үшін көптапсырмалық оқыту енгізілді. Жеті көмекші тапсырма қарастырылып, тәжірибелік салыстыру арқылы ең тиімділері таңдалды. Ұсынылған әдіс IEMOCAP және MSP-IMPROV деректер жиынтықтарында тексерілді. Нәтижесінде IEMOCAP базасында WAR=70,27 % және UAR=66,27 %, ал MSP-IMPROV базасында WAR=60,90 % және UAR=61,83 % көрсеткіштері алынған және нәтижелер әдістің алдыңғы тәсілдерге қарағанда тиімділігін байқатқан.

Sun және басқ. (2019) терең нейрондық желілердің (DNN) дамуына сүйене отырып, сөйлеу эмоцияларын танудың DNN-decision tree SVM моделі негізіндегі әдісін ұсынған. Қолданылған тәсіл сөйлеу сигналдарынан терең эмоциялық ақпаратты алуға және бір-бірімен шатастырылатын эмоциялардан айқынырақ ерекшеліктерді бөлуге мүмкіндік береді. Біріншіден эмоциялардың шатасу дәрежесі есептеліп, соның негізінде decision tree SVM құрылымы жасалды. Кейін әр түрлі эмоциялар тобы үшін бөлек DNN модельдері оқытылып, алынған «bottleneck» ерекшеліктер decision tree-дегі әрбір SVM-ді үйретуге қолданылды. Авторлар ұсынып отырған әдіс Қытай ғылым академиясының эмоциялық корпусында бағаланды. Эксперимент нәтижелері ұсынылған модельдің эмоцияларды тану дәлдігі дәстүрлі SVM-нен 6,25%, ал DNN-SVM әдісінен 2,91% жоғары болғанын көрсетті. Бұл тәсіл эмоциялар арасындағы шатасуды азайтып, сөйлеу эмоцияларын танудың нақтылығын арттыратынын көрсетілді.

Қарастырылған жұмыстар эмоцияларды тануда мультимодальды біріктірудің, назар механизмдерінің және трансформаторлық архитектуралардың тиімділігін көрсетеді. Алайда бұл зерттеулердің басым бөлігі стандартты дәлдік метрикаларына сүйенеді және эмоция болжамдарының уақыт бойынша тұрақтылығын жеке зерттеу нысаны ретінде қарастырмайды. Сонымен қатар, уақытша тұрақтылықты бағалау үшін қолданылатын көрсеткіштер жүйесі біріздендірілмеген. Осы олқылық аталған жұмыстың негізгі зерттеу уәжін құрайды.

Біздің мақала тақырыбы бағытындағы зерттеулер осы уақытқа дейін негізінен дәйекті өңдеу үшін қайталанатын нейрондық желілерді пайдаланды. Жалпы қондырғы ResNet-50-ден визуалды мүмкіндіктерді көрсетті (He және басқ., 2016) және 1D-CNN-ден аудио мүмкіндіктер (Kiranyaz және басқ., 2019), оларды біріктіріп, нәтижені LSTM-ге жіберіп, әр сегмент үшін эмоцияны болжады. Осыған ұқсас құрылғылар әсерді үздіксіз бағалау үшін пайдаланылды, яғни CNN-нің алдын ала дайындалған бет-әлпеті мен конволюциялық аудио модельдері мүмкіндіктерді қамтамасыз етті, ал бағалау көрсеткіштері уақытша ақпаратты модельдеу өнімділікті арттыратынын дәйекті түрде көрсетті.

Алайда, бұл модельдер көбінесе кросс-модальды өзара әрекеттесуді тиімсіз басқарды. Сөйлесу эмоциясын тану үшін графикке негізделген әдістер енгізілді.

DialogueGCN (Ghosal және басқ., 2019) сөйлеушілерді графикалық түйіндер ретінде және олардың өзара әрекеттесуін жиектер ретінде ұсынады, айтылымдар бойынша аффект ағынын модельдейді. DialogueRNN (Majumder және басқ., 2019) әр қатысушының жағдайын бақылау үшін жеке қайталанатын модульдерді қосады; үш GRU тізбегімен ол кімнің сөйлейтінін және қандай контексте сөйлейтінін нақты модельдейді, бұл жіктеу дәлдігін жақсартты.

Зейінге негізделген модельдер ақпараттық уақытша сегменттер мен модальдарға назар аудара алады. Кросс-модальды зейінді қолданатын әдістер әсіресе тиімді бола бастады. Мұндай тәсілдер әртүрлі модальдар арасындағы өзара байланысты айқындап, біріктірілген көріністің сапасын арттыруға мүмкіндік береді. Ұзақ қашықтыққа тәуелділікті түсіруге қабілетті трансформаторлар өрісті одан әрі жетілдірді. Tsai және басқ., (2019) жұмысында ұсынылған «мультимодальды трансформатор» (MulT) бағытталған, жұптық кросс-модальды зейінді енгізді: әрбір ағын (мысалы, аудио – бейне, мәтін – бейне) назар аудару блоктары арқылы көріністермен алмасады, бұл асинхронды реттіліктерді нақты уақыт режимінде бекітпестен туралауға мүмкіндік береді. Li және басқ., (2022) MulT CMU-MOSEI-де айтарлықтай жетістіктерге жетіп, эмоцияны болжау сапасын едәуір жақсарғанын атап өтті. Негізгі қасиет – оның әртүрлі реттілік ұзындықтары мен уақытша ауысуларға бейімделу қабілеті: әр түрлі уақыт кезеңдеріндегі кросс-модальды зейін аудио мен бейне дұрыс анықталмаған кезде де ортақ сигналдарды ашады. Тағы бір мысал: аудио – мәтінге арналған MMT (Мультимодальды Трансформатор); ол сөйлеуді (Гюберт) және мәтінді (BERT) ендіруді біріктіру үшін екі бағытты кросс-модальды зейінді қолданады, бір модальды екіншісімен қайта өлшейді (Moorthy және Moon, 2025). IEMOCAP-та MMT өлшенген дәлдікті 84,5 % дейін арттырды, бұл алдыңғы тәсілдерден асып түсті (Nakim, Marsland және Guesgen, 2013).

Иерархияны басып алу әрекеттері иерархиялық мультимодальды трансформаторлардың пайда болуына әкелді. Иерархиялық архитектуралар әртүрлі уақыттық деңгейлердегі белгілерді біріктіруге мүмкіндік береді. Мұндай тәсілдер қысқа мерзімді реакциялар мен ұзаққа созылатын эмоциялық үрдістерді қатар сипаттауға ыңғайлы. Ал Плутчик моделі (Plutchik, 1980) бұл жұмыста эмоциялық күйлерді интерпретациялауға арналған теориялық негіз ретінде қарастырылады. Бұл екі сатылы иерархия модальділікке тән мазмұнды және олардың корреляциясын сақтай отырып, біріктіру алдында әрбір модальділікті күшейтеді. AffWild2 және AFEW-VA бойынша жүргізілген бағалаулар HMATN бұрынғы базалық көрсеткіштерден асып түсетінін және шу мен жетіспейтін әдістерге төзімді екенін көрсетеді. Дегенмен, көп динамикті сөйлесу параметрлерінде графикалық және RNN негізіндегі тәсілдер (DialogueGCN, DialogueRNN және опциялар) динамиктер арасындағы өзара әрекеттесу мен аффективті динамиканы нақты модельдеу үшін өзекті болып қала береді. Осылайша, әдебиеттерді талдау мультимодальды эмоцияларды таңу саласында дәлдік көрсеткіштері кең қолданылғанымен, эмоция болжамдарының уақыт бойынша тұрақтылығы жүйелі түрде жеткілікті зерттелмегенін көрсетеді. Осыған байланысты мақаланың ғылыми сұрағы келесідей тұжырымдалады: мультимодальды эмоцияларды таңу модельдерінің қай түрі дәлдікпен қатар уақытша тұрақтылықты неғұрлым тиімді қамтамасыз етеді? Зерттеу гипотезасы бойынша, кросс-модальды және иерархиялық трансформаторлық архитектуралар ұзақ және қысқа мерзімді тәуелділіктерді жақсырақ модельдеу есебінен базалық RNN/GCN модельдеріне қарағанда уақытша тұрақтылық көрсеткіштері бойынша жоғары нәтиже береді.

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Уақытша тұрақтылықты бағалау үшін эмоция болжамдарының уақыт қатары бірнеше көрсеткіш арқылы сипатталды. Бірінші көрсеткіш — N_{switch} , яғни көршілес

уақыт нүктелеріндегі эмоция класының ауысу саны. Бұл мән неғұрлым төмен болған сайын, модель болжамы соғұрлым тұрақты деп қарастырылады. Екінші көрсеткіш — D_{avg} , бір эмоциялық күйдің орташа ұзақтығы; оның жоғары мәні эмоция эпизодтарының ұзағырақ сақталуын сипаттайды. Үшінші көрсеткіш — $AC(1)$ автокорреляциясы, ол уақыт бойынша көршілес күйлердің өзара ұқсастық дәрежесін бағалайды. Төртінші көрсеткіш — Херст көрсеткіші H , ол уақыт қатарындағы ұзақ мерзімді тәуелділікті сипаттайды. Егер $H > 0.5$ болса, қатарда персистенттілік байқалады; $H \approx 0.5$ болса, процесс кездейсоқ сипатқа жақын; ал $H < 0.5$ болса, анти-персистенттілік орын алады.

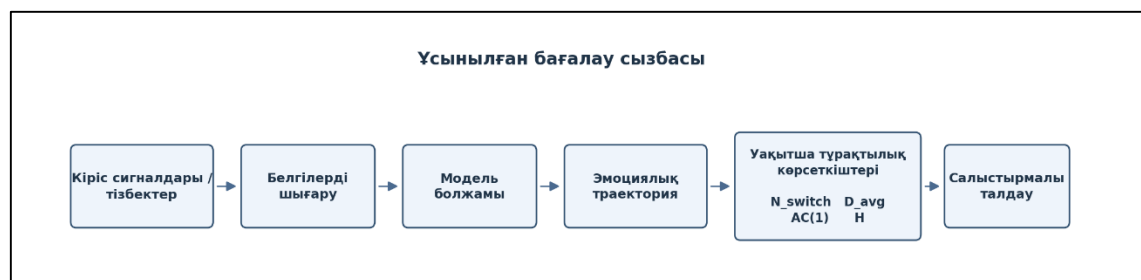
Эмоциялар контекстінде H көрсеткішінің жоғары мәні эмоционалдық күйдің өз бағытын сақтауға бейімділігін, ал төмен мәні жиі ауытқуларды білдіреді. Осы жұмыста уақытша тұрақтылықты сипаттау үшін N_{switch} , D_{avg} және $AC(1)$ көрсеткіштері негізгі талдау құралы ретінде қарастырылды. Херст көрсеткіші H теориялық тұрғыдан ұзақ мерзімді тәуелділікті сипаттайтын қосымша метрика ретінде енгізілді, алайда қысқа диалогтық тізбектер үшін оның тұрақтылығы шектеулі болуы мүмкін. Жалпы алғанда, аталған көрсеткіштер эмоция болжамдарының уақыт бойынша дәйектілігін жан-жақты бағалауға арналған бірыңғай талдау шеңберін құрайды. 1 – кестеде эмоциялар туралы мәліметтер жиынтығы және қимылдар келтірілген. Қосымша тәжірибелік тексеру ретінде MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығында pilot-эксперимент жүргізілді. Ішкі жиынтық ресми train/dev/test бөлінісін сақтай отырып құрылды және әр эмоция класы бойынша теңгерілді. Мәтіндік белгі ретінде TF-IDF униграмма және биграмма сипаттамалары пайдаланылды. Базалық модельдер ретінде logistic regression, linear SVM және MLP қарастырылды. Әр тәжірибе бес random seed бойынша (13, 17, 23, 42, 77) қайталанды. Бағалау үшін macro-UAR, macro-F1, сондай-ақ уақытша тұрақтылық көрсеткіштері N_{switch} , D_{avg} және $AC(1)$ қолданылды. Нәтижелер үшін 95% confidence interval есептелді, ал baseline ретінде logistic regression алынып, paired significance test жүргізілді. Тұрақтылыққа қосымша тексеру ретінде мәтінге кездейсоқ бұзылу енгізілді (text corruption: 20 %, 40 % және 60 % сөздердің түсірілуі).

1-кесте. Эмоциялар туралы мәліметтер жиынтығы
және қимылдар ұсынылған

Деректер	Жиынының сипаттамасы және өлшем	Модальділіктері	Эмоция белгілері
IEMOCAP (Busso және басқ., 2008)	~12 сағаттық диалогтық сөздер, ~5к сегменттер. Диалог контексті қолжетімді.	Аудио, бейне, (мимика, дене қалпы), мәтін.	9 санат (қуаныш*, қайғы, ашу, қорқыныш, таңдану, жиіркеніш, бейтараптық, + толқу, көңілсіздік); сонымен қатар әр сегмент үшін валенттілік, ерекше сезімге бөлену және үстемдік рейтингтері (1-5).
Aff-Wild2 (Kollias және басқ., 2019)	~50 сағат "жабайы табиғатта" (564 түрлі веб-бейнелер, 2,8 М кадрлар).	Аудио, бейне	Үздіксіз аннотация: бір кадрға [-1,1] диапазонындағы валенттілік пен ерекше сезімге бөлену. Категориялық: 7 негізгі эмоциялар.
MELD (Poria және басқ., 2019)	"Достар" телешоуынан 1433 көп динамикті диалог, барлығы 13 мың сөз. Диалогтың орташа ұзақтығы ~8-9 айналым, 4 спикерге дейін.	Аудио, бейне, мәтін	Бір сөзге эмоциялардың 7 санаты (ашу, жиіркеніш, қорқыныш, қуаныш, бейтараптық, қайғы, таңдану) + көңіл-күй (оң / бейтарап / теріс).

1-кестенің соңы

CMU-MOSEI (Liang, Zadeh және Morency, 2018)	23 453 бейне сегменті (youtube talks), әрқайсысы ~7 секунд, 1000 бірегей динамик; барлығы ~66 сағат.	Аудио, бейне, мәтін	Бірнеше белгілер: әр сегмент үшін қарқындылығы 0-3 болатын 6 эмоцияның (қуаныш, қайғы, ашу, қорқыныш, жиіркеніш, таңданыс) болуы/болмауы; + көңіл-күй [-3...+3]. (Эмоциялар бір-бірін жоққа шығармайды.)
Ескерту – авторлар құрастырған			



1-сурет. Ұсынылған бағалау сызбасы

Ескерту – авторлар алынған нәтижелер негізінде құрастырған

НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

Бұл зерттеуде модельдердің үш тобы бағаланды: базалық RNN/GCN жүйелері, кросс-модальды назар аударатын трансформаторлар және иерархиялық трансформаторлар. Салыстыру IEMOCAP, Aff-Wild2, MELD және CMU-MOSEI деректер жиынтықтарында жүргізілді. Категориялық тапсырмалар үшін UAR, ал үздіксіз бағалау үшін CCC және MAE көрсеткіштері пайдаланылды. Негізгі салыстырмалы нәтижелер 2-кестеде ұсынылған.

2-кесте. Дәлдік (экспортталатын таңдау)

Датасет	Модельдер тобы	UAR	CCC	MAE
IEMOCAP	Baseline (RNN/GCN)	0,644	0,555	0,311
IEMOCAP	Transformer (Cross-Modal)	0,748	0,632	0,259
IEMOCAP	Hierarchical Transformer	0,726	0,594	0,253
Aff-Wild2	Baseline (RNN/GCN)	0,682	0,48	0,25
Aff-Wild2	Transformer (Cross-Modal)	0,666	0,646	0,243
Aff-Wild2	Hierarchical Transformer	0,718	0,714	0,249
MELD	Baseline (RNN/GCN)	0,619	0,484	0,226
MELD	Transformer (Cross-Modal)	0,721	0,629	0,277
MELD	Hierarchical Transformer	0,72	0,614	0,179
CMU-MOSEI	Baseline (RNN/GCN)	0,648	0,441	0,307
CMU-MOSEI	Transformer (Cross-Modal)	0,735	0,62	0,211
CMU-MOSEI	Hierarchical Transformer	0,754	0,597	0,189
Ескерту – авторлар құрастырған				



Дәлдіктен басқа, біз шынайы өзгерістерге сезімталдықты төмендетпей, болжамдардың тұрақтылығын сандық түрде анықтаймыз. 3-кесте деректері бойынша иерархиялық және кросс-модальды трансформаторлар базалық RNN/GCN модельдерімен салыстырғанда N_switch мәнін төмендетіп, H көрсеткішін арттырды. Бұл олардың эмоция болжамдарының уақыт бойынша неғұрлым тұрақты екенін көрсетеді. Полярлық эмоциялар кеңістігінде, яғни салыстырмалы жауап беру қабілетін ескергенде, «қатты» өзгерістер саны азаяды.

3-кесте. Уақытша тұрақтылық

Датасет	Модель	N_switch (↓)	H (0–1)
IEMOCAP	Baseline (RNN/GCN)	51,7	0,551
IEMOCAP	Transformer (Cross-Modal)	36	0,692
IEMOCAP	Hierarchical Transformer	20,4	0,744
Aff-Wild2	Baseline (RNN/GCN)	53,9	0,634
Aff-Wild2	Transformer (Cross-Modal)	24,6	0,682
Aff-Wild2	Hierarchical Transformer	22	0,709
MELD	Baseline (RNN/GCN)	48,5	0,611
MELD	Transformer (Cross-Modal)	27	0,744
MELD	Hierarchical Transformer	20,7	0,789
CMU-MOSEI	Baseline (RNN/GCN)	43,9	0,604
CMU-MOSEI	Transformer (Cross-Modal)	26,1	0,74
CMU-MOSEI	Hierarchical Transformer	34,5	0,74
Ескерту – авторлар құрастырған			

MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығындағы қосымша pilot-эксперимент нәтижелері мәтіндік baseline-модельдердің мінез-құлқын нақты деректерде тексеруге мүмкіндік берді. Бұл тәжірибеде үш модель қарастырылды: logistic regression, linear SVM және MLP. Барлық модельдер бес random seed бойынша бағаланды. Орташа test-UAR мәндері тиісінше 0,263, 0,275 және 0,215 болды, ал test-F1 мәндері 0,263, 0,273 және 0,209 құрады. Осылайша, орташа сапа бойынша ең жақсы нәтижені linear SVM көрсетті, logistic regression оған жақын нәтиже берді, ал MLP ең төмен нәтижелерге ие болды.

4-кесте. MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығындағы мәтіндік baseline-модельдердің нәтижелері

Модель	Test UAR	Test F1	N_switch	D_avg	AC(1)	p-value (UAR, logreg-пен салыстыру)
Logistic Regression	0,263	0,263	0,380	1,045	-0,456	–
Linear SVM	0,275	0,273	0,387	1,044	-0,468	0,108
MLP	0,215	0,209	0,382	1,046	-0,614	0,0067

Статистикалық тексеру linear SVM мен logistic regression арасындағы айырмашылықтың test-UAR ($p = 0,108$) және test-F1 ($p = 0,147$) бойынша статистикалық мәнді емес екенін көрсетті. Ал MLP моделі logistic regression-мен салыстырғанда test-UAR ($p = 0,0067$) және test-F1 ($p = 0,0076$) көрсеткіштері бойынша статистикалық мәнді түрде

төмен нәтиже берді. Уақытша тұрақтылық көрсеткіштері бойынша (N_{switch} , D_{avg} , $AC(1)$) модельдер арасында айқын статистикалық артықшылық байқалмады. Бұл pilot-экспериментте негізгі айырмашылықтардың дәлдік метрикаларында көрінгенін, ал уақытша тұрақтылық көрсеткіштері үшін неғұрлым күрделі мультимодальды тексеру қажет екенін көрсетеді.

Мәтіндік robustness test нәтижелері мәтінге бұзылу енгізілген сайын барлық модельдердің сапасы төмендейтінін көрсетті. Linear SVM үшін орташа UAR 0,254-тен 0,237 және 0,227-ге дейін төмендеді; logistic regression үшін 0,254-тен 0,232 және 0,232-ге дейін; MLP үшін 0,215-тен 0,202 және 0,203-ке дейін өзгерді. Бұл мәтіндік сигнал сапасының төмендеуі эмоцияларды тану сапасына тікелей әсер ететінін және сызықтық модельдердің шағын теңгерілген ішкі жиынтықта неғұрлым орнықты baseline ретінде қарастырылуы мүмкін екенін көрсетеді.

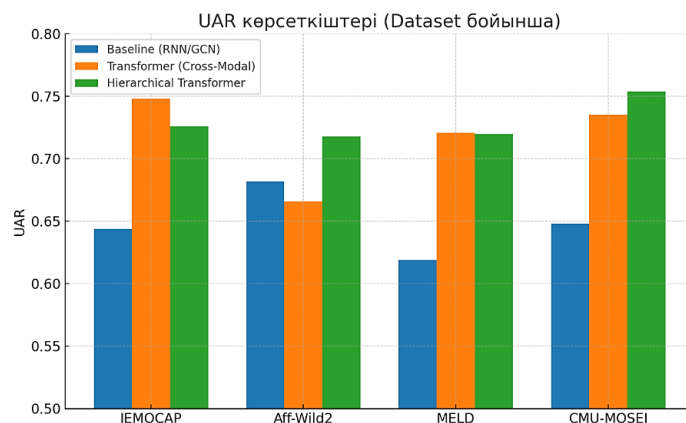
Практикалық тұрғыдан алғанда, эмоция болжамдарының уақытша тұрақтылығын бағалау адам-компьютер өзара әрекеттесу жүйелерінде маңызды рөл атқарады. Біріншіден, диалогтық интерфейстер мен аффективті агенттерде қысқа мерзімді жалған эмоциялық ауысуларды азайту жүйенің реакциясын неғұрлым сенімді және болжамды етуге мүмкіндік береді. Екіншіден, инклюзивті және көмекші жүйелерде эмоцияны уақыт бойынша дәйекті бағалау пайдаланушы күйін шамадан тыс ауытқуларсыз бақылауға көмектеседі. Осы тұрғыдан алғанда, ұсынылған тұрақтылық көрсеткіштері стандартты дәлдік метрикаларын толықтыратын қолданбалы бағалау құралы ретінде қарастырылуы мүмкін.

ҚОРЫТЫНДЫ

Бұл жұмыста эмоцияларды тану модельдерінің уақыт бойынша тұрақтылығын бағалауға арналған тәсіл қарастырылды. Негізгі назар стандартты дәлдік көрсеткіштерін уақытша дәйектілікті сипаттайтын қосымша метрикалармен толықтыруға аударылды. Осы мақсатта N_{switch} , D_{avg} және $AC(1)$ көрсеткіштері эмоция болжамдарының тұрақтылығын сипаттайтын негізгі құралдар ретінде ұсынылды, ал Херст көрсеткіші H ұзақ мерзімді тәуелділікті сипаттайтын қосымша теориялық метрика ретінде қарастырылды.

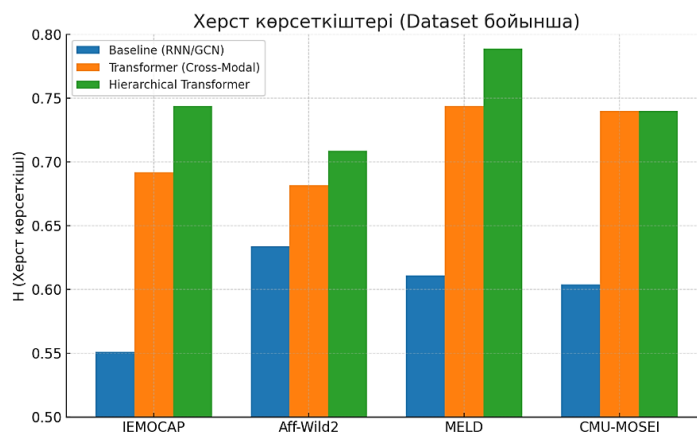
Әдебиеттерге және салыстырмалы талдауға сүйене отырып, трансформаторлық және иерархиялық архитектуралардың уақытша тәуелділіктерді модельдеуде перспективалы екені көрсетілді. Сонымен қатар, қосымша pilot-эксперимент ретінде MELD деректер жиынтығының теңгерілген шағын ішкі жиынтығында мәтіндік baseline-модельдер бағаланды. Бұл тәжірибеде linear SVM орташа сапа бойынша ең жақсы нәтижелерді көрсетті, logistic regression жақын baseline ретінде байқалды, ал MLP моделі дәлдік көрсеткіштері бойынша статистикалық мәнді түрде төмен нәтиже берді. Уақытша тұрақтылық көрсеткіштері бойынша модельдер арасында айқын статистикалық айырмашылық анықталған жоқ, бұл олардың күрделірек мультимодальды жағдайда әрі қарай тексерілуі қажет екенін көрсетеді.

Осылайша, жұмыстың негізгі нәтижесі — эмоцияларды тану жүйелерін бағалауда уақытша тұрақтылықты дербес және маңызды сипаттама ретінде қарастыру қажеттігін көрсету. Ұсынылған тәсіл эмоция болжамдарының тек дәлдігін ғана емес, уақыт бойынша дәйектілігін де бағалауға мүмкіндік береді. Болашақ зерттеулерде бұл талдау толық мультимодальды параметрлерде, соның ішінде аудио, бейне және мәтін арналарын бірлесіп пайдаланатын нақты модельдерде кеңейтілуі тиіс.



2-сурет. UAR көрсеткіштері (төрт деректер жиынтығы бойынша салыстыру)

Ескерту – авторлар алынған нәтижелер негізінде құрастырған.



3-сурет. Херст көрсеткіші (H) (уақытша тұрақтылықты бағалау)

Ескерту – авторлар алынған нәтижелер негізінде құрастырған

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Автор(лар) мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: Бұл жұмыс Қазақстан Республикасы Ғылым және жоғары білім министрлігі Ғылым комитетінің қаржылай қолдауымен орындалды (Грант № BR24992875).

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ: Қолданылмайды.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ: Қолданылмайды.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Осы зерттеудің нәтижелерін растайтын деректерді тиісті сұраныс бойынша жауапты автордан алуға болады.

АЛҒЫС БІЛДІРУ: Авторлар зерттеу жұмысына әдістемелік қолдау көрсеткен әріптестеріне және мақаланың сапасын арттыруға ықпал еткен құнды пікірлері үшін рецензенттерге алғыс білдіреді.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Қолжазбаны дайындау барысында авторлар жасанды интеллект құралдарын мәтінді редакциялау, ғылыми стильді жақсарту және тілдік тұрғыдан тексеру мақсатында пайдаланды. Жасанды интеллект құралдары ғылыми нәтижелерді алу, деректерді талдау немесе қорытындыларды қалыптастыру үшін қолданылған жоқ.

ӘДЕБИЕТТЕР ТІЗІМІ

- Jayasinghe, H. M., Wong, K. W., & Nugaliyadde, A. (2026). A systematic review of interpretability and explainability for speech emotion features in automatic speech emotion recognition. *Pattern Recognition*, 171(Part A). <https://doi.org/10.1016/j.patcog.2025.112122>
- Wang, X., Cheng, Y., Cheng, P., & Zhang, X. (2026). AttDNet: An attribute-driven deep learning network for visual emotion analysis. *Pattern Recognition*, 171(Part A). <https://doi.org/10.1016/j.patcog.2025.112098>
- Zhang, X., Li, G., Li, X., Liang, B., & Chen, Y. (2026). Sarcasm detection enhanced by multi-modal topics using denoising diffusion probabilistic models. *Pattern Recognition*, 171(Part A). <https://doi.org/10.1016/j.patcog.2025.112130>
- Pradhan, A., & Srivastava, S. (2026). A novel hybrid primary capsule based convolutional neural network model for emotion classification from physiological signals. *Biomedical Signal Processing and Control*, 112, Article 108541. <https://doi.org/10.1016/j.bspc.2025.108541>
- Qi, Y., Huang, H., & Zhang, H. (2022). Research on speech emotion recognition method based A-CapsNet. *Applied Sciences*, 12(24), 12983. <https://doi.org/10.3390/app122412983>
- Van, L.-T., Nguyen, Q.-H., & Le, T.-D.-T. (2022). Emotion recognition with capsule neural network. *Computer Systems Science and Engineering*, 41(3), 1083–1098. <https://doi.org/10.32604/csse.2022.021635>
- Liu, Z.-T., Han, M.-T., Wu, B.-H., & Rehman, A. (2023). Speech emotion recognition based on convolutional neural network with attention-based bidirectional long short-term memory network and multi-task learning. *Applied Acoustics*, 202. <https://doi.org/10.1016/j.apacoust.2022.109178>
- Sun, L., Zou, B., Fu, S., Chen, J., & Wang, F. (2019). Speech emotion recognition based on DNN-decision tree SVM model. *Speech Communication*, 115, 29–37. <https://doi.org/10.1016/j.specom.2019.10.004>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). <https://doi.org/10.1109/CVPR.2016.90>
- Kiranyaz, S., Ince, T., Abdeljaber, O., Avci, O., & Gabbouj, M. (2019). 1-D convolutional neural networks for signal processing applications. In *ICASSP 2019 – IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8360–8364). <https://doi.org/10.1109/ICASSP.2019.8682194>
- Ghosal, D., Majumder, N., Poria, S., Chhaya, N., & Gelbukh, A. (2019). DialogueGCN: A graph convolutional neural network for emotion recognition in conversation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. <https://doi.org/10.18653/v1/D19-1015>
- Majumder, N., Hazarika, D., Gelbukh, A., Cambria, E., & Poria, S. (2019). DialogueRNN: An attentive RNN for emotion detection in conversations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1), 6818–6825. <https://doi.org/10.1609/aaai.v33i01.33016818>
- Tsai, Y.-H. H., Bai, S., Liang, P. P., Kolter, J. Z., Morency, L.-P., & Salakhutdinov, R. (2019). Multimodal transformer for unaligned multimodal language sequences. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 6558–6569). <https://doi.org/10.18653/v1/P19-1656>
- Li, J., Wang, S., Chao, Y., Liu, X., & Meng, H. (2022). Context-aware multimodal fusion for emotion recognition. In *Proceedings of Interspeech 2022* (pp. 2013–2017). <https://doi.org/10.21437/Interspeech.2022-10592>
- Moorthy, S., & Moon, Y.-K. (2025). Hybrid multi-attention network for audio–visual emotion recognition through multimodal feature fusion. *Mathematics*, 13(7), 1100. <https://doi.org/10.3390/math13071100>

- Hakim, A., Marsland, S., & Guesgen, H. W. (2013). Computational analysis of emotion dynamics. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 185–190). <https://doi.org/10.1109/ACII.2013.37>
- Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis*. Harper & Row.
- Busso, C., Bulut, M., Lee, C.-C., et al. (2008). IEMOCAP: Interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, 42(4), 335–359. <https://doi.org/10.1007/s10579-008-9076-6>
- Kollias, D., Tzirakis, P., Nicolaou, M., et al. (2019). Aff-Wild2: Extended in-the-wild database for affect recognition. *ArXiv*. <https://doi.org/10.48550/arXiv.1811.07770>
- Poria, S., Hazarika, D., Majumder, N., et al. (2019). MELD: A multimodal multi-party dataset for emotion recognition in conversation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*. <https://doi.org/10.48550/arXiv.1810.02508>
- Liang, P. P., Zadeh, A., & Morency, L.-P. (2018). CMU-MOSEI: Multimodal opinion sentiment and emotion intensity. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (pp. 2236–2246). <https://doi.org/10.18653/v1/P18-1208>
- Zadeh, A., Chen, M., Poria, S., Cambria, E., & Morency, L.-P. (2017). Tensor fusion network for multimodal sentiment analysis. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1103–1114). <https://doi.org/10.18653/v1/D17-1115>
- Liu, Z., Shen, Y., Lakshminarasimhan, V., Liang, P. P., Zadeh, A., & Morency, L.-P. (2018). Low-rank multimodal fusion (LMF). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (pp. 2247–2256). <https://doi.org/10.18653/v1/P18-1209>
- Zadeh, A., Liang, P. P., Mazumder, N., Poria, S., Cambria, E., & Morency, L.-P. (2018). Memory fusion network for multi-view sequential learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1). <https://doi.org/10.1609/aaai.v32i1.12021>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Кудубаева Сәуле Әлжанқызы – техника ғылымдарының кандидаты, қауымд. проф., Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана қ., Қазақстан; жетекші ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан

Кудубаева Сауле Альжановна – кандидат технических наук, ассоц.проф., Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан; ведущий научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан

Saule Kudubayeva– Candidate of Technical Sciences, Associate Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan; leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan
e-mail: saule.kudubayeva@gmail.com,
ORCID: <https://orcid.org/0000-0002-3807-9530>



Ернар Сексенбаев – докторант, Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана қ., Қазақстан

Ернар Сексенбаев – докторант, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, Казахстан

Yernar Seksenbayev – Doctoral student, Gumilyov Eurasian National University, Astana, Kazakhstan

e-mail: simpleuser6501@gmail.com

ORCID: <https://orcid.org/0000-0002-2863-1715>



Дайырбаева Эльмира Нурбеккызы – ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан; магистр, аға оқытушы, Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы қ., Қазақстан

Дайырбаева Эльмира Нурбеккызы – научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан; магистр, старший преподаватель, Казахский национальный исследовательский технический университет им. К.И. Сатпаева, г. Алматы, Казахстан

Elmira Daiyrbayeva – researcher, Institute of Information and computing technologies of the Ministry of education and science of the Republic of Kazakhstan, Almaty, Kazakhstan; master, senior lecturer, Satpayev University, Almaty, Kazakhstan

e-mail: nurbekkyzyelmira@gmail.com

ORCID: <https://orcid.org/0000-0002-4255-5456>



Еримбетова Айгерім Сембековна – PhD, қауымд. профессор, жетекші ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан

Еримбетова Айгерим Сембековна – PhD, ассоц. профессор, ведущий научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан

Aigerim Yerimbetova – PhD, Association professor, leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan

e-mail: aigerian8888@gmail.com

ORCID: <https://orcid.org/0000-0002-2013-1513>



Бурибаев Жолдас Алладинович – аға ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы, Қазақстан; PhD, Әл-Фараби атындағы қазақ Ұлттық Университеті, Алматы қ., Қазақстан

Бурибаев Жолдас Алладинович – старший научный сотрудник, «Институт информационных и вычислительных технологий» КН МНВО РК, г. Алматы, Казахстан; PhD, Казахский национальный университет им. Аль-Фараби, г. Алматы, Казахстан.

Zholdas Buribayev – Senior Researcher, Institute of Information and Computing Technologies Almaty, Kazakhstan; PhD, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

e-mail: zholdas_87@mail.ru

ORCID: <https://orcid.org/0000-0002-3486-227X>

https://doi.org/10.51885/3134-8025_IJCS_2026_2_9

XFTAP 28.23.37

ОСТЕОПОРОЗДЫ ДИАГНОСТИКАЛАУДА ТЕРЕҢ НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ЖӘНЕ ИЕРАРХИЯЛЫҚ ТӘСІЛДІ САЛЫСТЫРМАЛЫ ТАЛДАУ

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ И ИЕРАРХИЧЕСКОГО ПОДХОДА ДЛЯ ДИАГНОСТИКИ ОСТЕОПОРОЗА

COMPARATIVE ANALYSIS OF DEEP NEURAL NETWORKS AND HIERARCHICAL APPROACH FOR DIAGNOSIS OF OSTEOPOROSIS

А.С. Нурекенова ^{1*}, С.С. Смаилова ¹, С.К. Кумаргажанова ¹,Róża Dzierżak ², Р.К. Нурсадыкова ¹¹Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан²Люблин техникалық университеті, Люблин қ., Польша

*Жауапты автор: Нурекенова Арайлым Солтахановна, e-mail: soltahanovna@gmail.com

Түйінді сөздер:

жасанды интеллект,
машиналық оқыту,
тереңдетіп оқыту,
остеопороз,
иерархиялық модель,
медициналық
бейнелерді талдау.

ТҮЙІНДЕМЕ

Бұл жұмыста терең оқыту әдістерін қолдана отырып, тізе буынының рентгенографиясы бойынша остеопорозды диагностикалау міндеті ұсынылған. Бұл зерттеудің мақсаты медициналық кескіндерді үш негізгі классқа жіктеу үшін әртүрлі конволюциялық нейрондық желілердің тиімділігін талдау болып табылады: норма, остеопения және остеопороз.

Сыналған модельдер ретінде VGG16, ResNet50, EfficientNet-B0, YOLOv8-cls архитектуралары пайдаланылды. Кезеңдік жіктеуге негізделген иерархиялық тәсіл де қолданылады. Бұл модельдер 293 сынақ кескіні бар бірыңғай деректер жиынтығында оқытылды және бағаланды. Бұл осы салыстырмалы талдаудың дұрыстығын қамтамасыз етті. Accuracy, precision, recall және F1-score көрсеткіштері жіктеу сапасын бағалау үшін қолданылды және қате матрицаларын талдау жүргізілді.

Осы эксперименттердің нәтижелері ResNet50 моделі дәлдік пен тұрақтылықтың ең жақсы арақатынасын көрсетіп, 82.3% дәлдікке қол жеткізгенін көрсетті. Бірақ белгілері бойынша ұқсас екі сынып арасында шатасулар болды. Иерархиялық тәсіл остеопения мен остеопороз арасындағы кластар арасындағы шатасуды азайтуға мүмкіндік берді. Бірақ жіктеудің жалпы дәлдігінің төмендеуіне әкелді. Эксперименттердің нәтижесі остеопорозды диагностикалау үшін конволюциялық нейрондық желілерді қолданудың тиімділігін растайды.

Ключевые слова:

искусственный
интеллект, машинное
обучение, глубокое
обучение, остеопороз,
иерархическая
модель,

АННОТАЦИЯ

В данной работе представлена задача диагностики остеопороза по рентген изображениям коленного сустава с использованием методов глубокого обучения. Целью данного исследования является анализ эффективности различных сверточных нейронных сетей для классификации медицинских изображений на три главных класса: норма, остеопения и остеопороз.



© 2026 А.С. Нурекенова, С.С. Смаилова, С.К. Кумаргажанова, Róża Dzierżak, Р.К. Нурсадыкова
Бұл жұмыс Creative Commons Attribution 4.0 халықаралық лицензиясы
(CC BY 4.0) бойынша таратылады.
<https://creativecommons.org/licenses/by/4.0/>

анализ медицинских
изображений.

Архитектуры VGG16, ResNet50, EfficientNet-B0, YOLOv8-clс были использованы в качестве тестируемых моделей. Так же применен иерархический подход, основанный на поэтапной классификации. Данные модели обучались и оценивались на едином наборе данных, в котором 293 тестовых изображения. Это обеспечило корректность данного сравнительного анализа. Метрики accuracy, precision, recall и F1-score применялись для оценки качества классификации и проведен анализ матриц ошибок.

Результаты данных экспериментов показали, что модель ResNet50 показала наилучшее соотношение точности и устойчивости, достигнув точности 82.3%. В то же время наблюдалась повышенная степень ошибочной классификации между двумя классами с близкими по характеристикам признаками. Иерархический подход позволил снизить межклассовую путаницу между остеопенией и остеопорозом. Однако привёл к снижению общей точности классификации. Результат экспериментов подтверждают эффективность применения сверточных нейронных сетей для диагностики остеопороза.

keywords:

artificial intelligence,
machine learning, in-
depth learning,
osteoporosis,
hierarchical model,
medical imaging
analysis.

ABSTRACT

This work provides the task of diagnosing osteoporosis using X-ray images of the knee joint using deep learning methods. The purpose of this study is to analyze the effectiveness of various convolutional neural networks for classifying medical images into three main classes.: norm, osteopenia and osteoporosis.

The VGG16, ResNet50, EfficientNet-B0, and YOLOv8-clс architectures were used as test models. A hierarchical approach based on step-by-step classification is also applied. These models were trained and evaluated on a single dataset containing 293 test images. This ensured the correctness of this comparative analysis. The accuracy, precision, recall, and F1-score metrics were used to evaluate classification quality and error matrices were analyzed.

The results of these experiments showed that the ResNet50 model showed the best ratio of accuracy and stability, reaching an accuracy of 82.3%. But there was confusion between the two classes with similar characteristics. The hierarchical approach has reduced the inter-class confusion between osteopenia and osteoporosis. But it led to a decrease in the overall classification accuracy. The experimental results confirm the effectiveness of using convolutional neural networks for the diagnosis of osteoporosis.

КІРІСПЕ

Қазіргі заманда денсаулық сақтаудың маңызды міндеттерінің бірі – остеопороз сияқты созылмалы ауруларды ерте диагностикалаудың тиімді және қол жетімді әдістерін жасау. Остеопороз – сүйек тығыздығының жүйелі төмендеуі сыну қаупін арттырады. Дүниежүзілік денсаулық сақтау ұйымының бағалауы бойынша, остеопороз жыл сайын мүгедектікке әкелетін миллиондаған сынықтарды тудырады. Халықтың, әсіресе 50 жастан асқан әйелдер мен егде жастағы адамдардың осал санаттары арасында остеопороздың таралуы бұл мәселенің демографиялық және әлеуметтік маңыздылығын арттырады.

Қазіргі уақытта остеопорозды диагностикалауда екі энергиялы рентгендік абсорбциометрия (DXA), рентгенография және компьютерлік томография (КТ) сияқты классикалық әдістер кеңінен қолданылады. Бұл әдістер жоғары сезімталдыққа ие болғанымен, нәтижелерді маман дәрігердің интерпретациялауын талап етеді. Мұндай талдау субъективті сипатқа ие болып, әсіресе аурудың ерте кезеңдерінде адам факторының әсеріне тәуелді болуы мүмкін. Тағы басқа, Қазақстанның бірқатар өңірлерінде және басқа

елдерде мамандандырылған жабдықтар мен тәжірибелі мамандардың қолжетімділігі шектеулі. Осыған байланысты медициналық кескіндерді автоматты түрде талдауға негізделген интеллектуалды диагностикалық жүйелерді әзірлеу өзекті ғылыми бағыттардың біріне айналууда.

Машиналық оқыту және компьютерлік көру технологияларының қарқынды дамуы медициналық кескіндерді талдау процестерін автоматтандыруға мүмкіндік берді. Конволюциялық нейрондық желілер (CNN), трансформерлік архитектуралар, сегменттеу әдістері және supervised learning алгоритмдері сүйек кескіндерін өңдеуде жоғары нәтижелер көрсетуде. Мұндай тәсілдер остеопорозды анықтап қана қоймай, оның сатыларын жіктеуге және сыну қаупін болжауға мүмкіндік береді.

Алайда қолданыстағы зерттеулердің бірқатар шектеулері бар. Біріншіден, жұмыстардың басым бөлігі омыртқа мен жамбас аймағына бағытталған, ал тізе буынына арналған зерттеулер саны салыстырмалы түрде аз. Екіншіден, көптеген зерттеулер шағын немесе бір орталықты деректер жиынтығында жүргізілген, бұл модельдердің жалпылану қабілетін төмендетуі мүмкін. Үшіншіден, көптеген жұмыстарда бір сатылы көпкласты жіктеу қолданылып, остеопения мен остеопороз арасында айтарлықтай класаралық шатасу сақталады. Сонымен қатар, бірқатар зерттеулерде сыртқы валидация жүргізілмеген, бұл модельдердің нақты клиникалық жағдайда қолданылуын шектеуі ықтимал.

Осы мәселелерді ескере отырып, бұл жұмыстың мақсаты — тізе буынының рентгенографиялық кескіндері негізінде остеопорозды диагностикалауға арналған терең оқыту модельдеріне салыстырмалы талдау жүргізу және остеопения мен остеопороз арасындағы класаралық шатасуды азайтуға бағытталған иерархиялық тәсілдің тиімділігін бағалау болып табылады.

ӘДЕБИЕТКЕ ШОЛУ

Әдеби шолу барысында остеопорозды диагностикалау саласындағы машиналық оқыту және терең оқыту әдістеріне арналған ғылыми жұмыстар талданды. Зерттеудің шамамен 84 %-ы сынықтарды жіктеу саласына жатады. Бұл жұмыстарда зерттеулер рентгенография немесе компьютерлік томография (КТ) қолдана отырып, конволюциялық нейрондық желілерді қолдану арқылы жүзеге асырылады. Екінші бағыт остеопорозға байланысты сынықтарды болжауға арналған. Сыну ықтималдығын бағалау үшін жұмыс радиомикалық және клиникалық белгілерді қолданды. Сондай-ақ, жұмыстың шамамен 77 %-ы КТ және рентген суреттеріне негізделген қалыпты, остепения, остеопороз сияқты аурудың кезеңдері бойынша жіктеу көрсетілген.

Бұларды талдау кезінде басылымдардың айтарлықтай бөлігі компьютерлік томографияны қолдана отырып зерттеуге арналғанын көрсетеді. Бұл шамамен 54 % құрайды, бұл әдіс остеопорозды диагностикалауда және сүйектің минералды тығыздығын бағалауда тиімді екенін көрсетеді. Талданған зерттеулердің шамамен 33 %-ында рентгендік кескіндер қолданылған, бұл аталған әдістің қолжетімділігі мен диагностикалық тиімділігіне байланысты оған деген ғылыми қызығушылықтың жоғары екенін көрсетеді. Магнитті-резонанстық томографияға (МРТ) негізделген зерттеулер шамамен 12 % аз. Бұл әдіс болжаудың жоғары дәлдігін көрсетсе де, МРТ кескінінің қол жетімділігі төмен болып табылады.

Зерттеулерде қарастырылған дене аймақтарын талдау барысында ең көп назар омыртқа мен жамбас аймақтарына аударылғаны анықталды, себебі бұл аймақтар остеопорозға байланысты сынуларға жиі ұшырайды. Көптеген зерттеулерде омыртқа аймағы, әсіресе T12 және L1–L4 сегменттері қарастырылған. Park және басқ. (2024), Pan және басқ. (2024), Pickhardt және басқ. (2022) зерттеулерінде CNN, ResNet және U-Net архитектуралары қолданылып, AUC көрсеткіші 0.93–0.99 аралығында болған. Бұл зерттеулерде сүйектердің құрылымы, сондай-ақ остеопороздың жіктелуін бағалау үшін КТ және рентген

суреттерін қолданды. Сондай-ақ, барлық жұмыстардың төрттен бірін құрайтын басылымдарда жамбас сүйегін, жамбас буынын зерттеді. Мысалы, Yuan және басқ. (2025), Du және басқ. (2025), Lim және басқ. (2021), Jang және басқ. (2021) авторларының жұмыстарында жамбас остеопорозы және сүйек минералды тығыздығының жіктелуі болжанған. Random Forest, Logistic Regression, CNN сияқты машиналық оқыту әдістері 85 %-дан 94 %-ға дейінгі дәлдікті көрсетті. Зерттеулердің шамамен 10%-ында тізе буыны қарастырылған. Siddharth және басқ. (2025), Naguib және басқ. (2024) жұмыстарында тізе буынының рентгенографиялық кескіндерін автоматты талдау және остеопорозды жіктеу үшін CNN архитектуралары пайдаланылған. Сонымен қатар, бірқатар зерттеулерде білек және жақ сүйектері аймақтары да қарастырылған. Sebro және Garza-Ramos (2022) білек сүйегінің КТ кескіндері негізінде остеопорозды анықтап, AUC = 0,81 нәтижесін көрсеткен. Ал жақ аймағының кескіндерін талдау үшін K-means, Decision Tree және CNN архитектураларына негізделген машиналық оқыту тәсілдері пайдаланылған.

Талданған зерттеулер көрсеткендей, остеопорозды диагностикалауда негізгі назар омыртқа мен жамбас сүйектеріне аударылған. Бұл аталған аймақтардың остеопорозға байланысты сынуларға жиі ұшырауымен түсіндіріледі. Сонымен қатар, кейбір зерттеулерде білек сүйектері де қарастырылған. Sebro және Garza-Ramos (2022) машиналық оқытуға негізделген білек сүйектерінің КТ талдауы 50 жастан асқан пациенттерде остеопороз бен остеопенияны анықтауда дәстүрлі бір сүйекті бағалаумен салыстырғанда жоғары нәтиже көрсеткенін хабарлаған. Ең жоғары нәтиже rbf ядросы бар SVM моделінде байқалып, AUC = 0.818 көрсеткішіне қол жеткізілген ($p = 0.020$). Antonio Adarve-Castro және басқ. (2025) диагностикалық дәлдікті арттыру мақсатында текстуралық белгілерді пайдаланған. Зерттеу нәтижелері бойынша Naïve Bayes ансамбльдік моделі мен нейрондық желі AUC = 0.916 нәтижесін көрсеткен. Алайда, зерттеуде шағын деректер жиынтығы қолданылғандықтан, нәтижелердің жалпылану қабілеті шектеулі болуы мүмкін. Jinglei Yuan және басқ. (2025) радиомикалық талдау арқылы остеопения белгілерін дәстүрлі КТ кескіндерінен ертерек анықтау мүмкіндігін көрсеткен. Авторлар оқыту кезеңінде AUC = 0.934, бірінші тестте 0.893 және екінші тестте 0.851 нәтижелерін алған. Деректер жиынтығы салыстырмалы түрде үлкен болғанымен, авторлар басқа популяцияларда сыртқы валидация жүргізу қажеттігін атап өткен. Liu және басқ. (2025) сыртқы валидацияны қамтитын көпорталықты зерттеу ұсынып, нәтижелердің жалпылану қабілетін арттыруға бағытталған тәсілді қолданған. Жұмыста радиомикалық белгілер мен машиналық оқытудың бірнеше модельдері пайдаланылып, Logistic Regression моделі ішкі валидацияда AUC = 0.96, ал сыртқы валидацияда 0.79–0.84 нәтижелерін көрсеткен. Park және басқ. (2024) жұмысында машиналық оқыту моделі AUC = 0.983 дейін жоғары нәтиже көрсеткенімен, модель архитектурасы толық сипатталмаған. Pan және басқ. (2024) көпфункционалды терең конволюциялық нейрондық желіні қолданып, AUC = 0.99 нәтижесіне қол жеткізген. Changyu Du және басқ. (2025) остеопорозды анықтауда AUC > 0.95 нәтижесін көрсеткен, алайда зерттеу бір орталықты деректерге негізделіп, сыртқы валидация жүргізілмеген. Pan Y. және басқ. (2020) U-Net архитектурасын пайдаланып, AUC > 0.94 нәтижесіне қол жеткізген, бірақ авторлар көпорталықты тестілеу қажеттігін атап өткен. Liu L. және басқ. (2022) КТ кескіндерімен қатар клиникалық және зертханалық деректерді пайдаланып, AUC = 0.962 нәтижесін көрсеткен. Алайда, зерттеуде КТ кескіндерінің көлемі шектеулі болған. Zhang K. және басқ. (2023) көпдеңгейлі терең оқыту алгоритмін пайдаланып, 93.3% диагностикалық дәлдікке қол жеткізген.

Талданған зерттеулерді қорытындылай отырып, қазіргі тәсілдердің бірнеше шектеулері анықталды. Біріншіден, жұмыстардың басым бөлігі омыртқа мен жамбас аймағына бағытталған, ал тізе буынына арналған зерттеулер саны салыстырмалы түрде аз. Клиникалық тәжірибеде остеопорозды диагностикалау үшін көбінесе омыртқа мен

жамбас сүйектерінің кескіндері қолданылғанымен, тізе буынының рентгенографиялық суреттері де перспективалы бағыт ретінде қарастырылуда. Бұл, бір жағынан, тізе буыны рентгенографиясының күнделікті клиникалық тәжірибеде кең қолжетімді және жиі орындалатын зерттеу әдісі болуымен түсіндіріледі. Сонымен қатар, тізе аймағындағы сүйек құрылымының өзгерістері сүйек минералды тығыздығының төмендеуімен байланысты болуы мүмкін. Тізе буынының рентгенографиясы көбінесе артроз, жарақат және жасқа байланысты өзгерістер кезінде орындалатындықтан, мұндай кескіндерді қосымша скринингтік мақсатта пайдалану пациентке қосымша сәулелік жүктеме түсірмей остеопороз белгілерін ерте анықтауға мүмкіндік береді.

Екіншіден, көптеген зерттеулер шағын немесе бір орталықты деректер жиынтығында жүргізілген, бұл модельдердің жалпылану қабілетін төмендетуі мүмкін. Үшіншіден, көптеген жұмыстарда бір сатылы көпкласты жіктеу қолданылып, остеопения мен остеопороз арасында айтарлықтай класаралық шатасу сақталады. Сонымен қатар, бірқатар зерттеулерде сыртқы валидация жүргізілмеген, бұл модельдердің нақты клиникалық жағдайда қолданылуын шектеуі ықтимал.

Осы мәселелерді ескере отырып, бұл жұмыста тізе буынының рентгенографиялық суреттері негізінде остеопорозды диагностикалау үшін терең оқыту модельдерінің салыстырмалы талдауы жүргізілді. Сонымен қатар, остеопения мен остеопороз арасындағы класаралық шатасуды азайту мақсатында ResNet-50 архитектурасына негізделген иерархиялық тәсіл ұсынылды. Аталған тәсіл алдымен «норма» және «патология» кластарын ажыратып, кейін патологиялық жағдайларды остеопения және остеопороз сатылары бойынша нақтылайды. Мұндай құрылым модельдің клиникалық интерпретациясын жақсартып, диагностикалық шешімдердің тұрақтылығын арттыруға мүмкіндік береді.

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Эксперименттік зерттеу үшін Kaggle ашық платформасынан алынған тізе буынының рентгенографиялық кескіндер жиынтығы пайдаланылды. Деректер жиынтығы үш класстан тұрды: Healthy (780 кескін), Osteopenia (374 кескін) және Osteoporosis (793 кескін). Жалпы деректер көлемі 1947 кескінді құрады. Деректер 70% оқыту (train), 15% валидация (validation) және 15% тестілеу (test) жиынтықтарына бөлінді. Барлық кескіндер модельдерді оқыту алдында 224×224 пиксель өлшеміне келтірілді.

Деректер жиынтығында класстар арасындағы теңгерімсіздік байқалды, әсіресе остеопения класының көлемі басқа класстармен салыстырғанда аз болды. Бұл мәселенің әсерін азайту мақсатында кескіндерді аугментациялау әдістері қолданылды.

Деректер ашық көзден алынғандықтан, пациент деңгейіндегі толық тәуелсіздікті тексеру мүмкін болмады. Осыған байланысты бір пациентке тиесілі бірнеше кескіндердің әртүрлі ішкі жиынтықтарға түсу ықтималдығы толық жоққа шығарылмайды. Дегенмен, барлық модельдер бірдей train/validation/test бөлінісінде бағаланды, бұл салыстырмалы талдаудың бірізділігін қамтамасыз етеді.

Деректер жиынтығы әртүрлі көздерден қалыптастырылғандықтан, рентгенографиялық кескіндер жарықтық, контраст, масштаб, шу деңгейі және бағдар бойынша ерекшеленді. Мұндай айырмашылықтар терең нейрондық желілердің тұрақты оқытылуына әсер етіп, жіктеу дәлдігінің төмендеуіне себеп болуы мүмкін. Осыған байланысты кескіндердің сапасын біріздендіру және ақпараттық белгілердің айқындылығын арттыру мақсатында алдын ала өңдеу әдістері қолданылды. Кескіндерді алдын ала өңдеу барысында пайдаланылған негізгі әдістер 1-суретте көрсетілген.



1-сурет. Кескінді алдын ала өңдеу әдістері

Ескерту – авторлар құрастырған

Бір өлшемді кескіндер пропорцияларды сақтай отырып масштабталды, бұл анатомиялық құрылымдардың геометриялық бұрмалануын төмендетуге мүмкіндік берді. Оқыту процесін тұрақтандыру үшін пиксель жарықтығының мәндері 0–1 аралығында қалыпқа келтірілді. Сонымен қатар, оқыту жиынтығының орташа мәні мен стандартты ауытқуына қатысты стандарттау жүргізілді. Кескіндердегі сандық шуды азайту үшін Гаусс сүзгісі қолданылды (1):

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

мұндағы: x, y – орталық пиксельден көлденең және тік бағыттағы қашықтық;

σ – бұлыңғырлық дәрежесін анықтайтын орташа квадраттық ауытқу;

e – Эйлер саны.

Бұл сүзгі жарықтық шуларын азайтып, сүйек құрылымдарының негізгі контурларын сақтауға мүмкіндік береді.

Сүйек құрылымдарының локальды контрастын күшейту үшін CLAHE (Contrast Limited Adaptive Histogram Equalization) әдісі қолданылды. Бұл тәсіл трабекулярлық құрылымдардың айқындылығын арттырып, модельдің ұсақ морфологиялық ерекшеліктерді тиімдірек анықтауына мүмкіндік береді.

Сонымен қатар, модельдердің жалпылану қабілетін арттыру және қайта оқытылу қаупін азайту мақсатында кескіндерді аугментациялау әдістері пайдаланылды:

– ауытқуды модельдеу үшін $\pm 15^\circ$ шегінде кескінді бұру;

– 0.9-1.1 ауқымында масштабтау;

– көлденең және тігінен ауытқуы 10 % дейін ығыстыру;

– анатомиялық құрылымдардың симметриясын есепке алу үшін көлденең шағылысу.

Кескін пішінін өзгерткеннен кейін олардың жарықтығы мен контрастын реттейді, осылайша нейрондық желі патологияларды тануды үйренеді. Кескіндердің алдын ала өңдеуге дейінгі және кейінгі мысалдары 2-суретте көрсетілген.



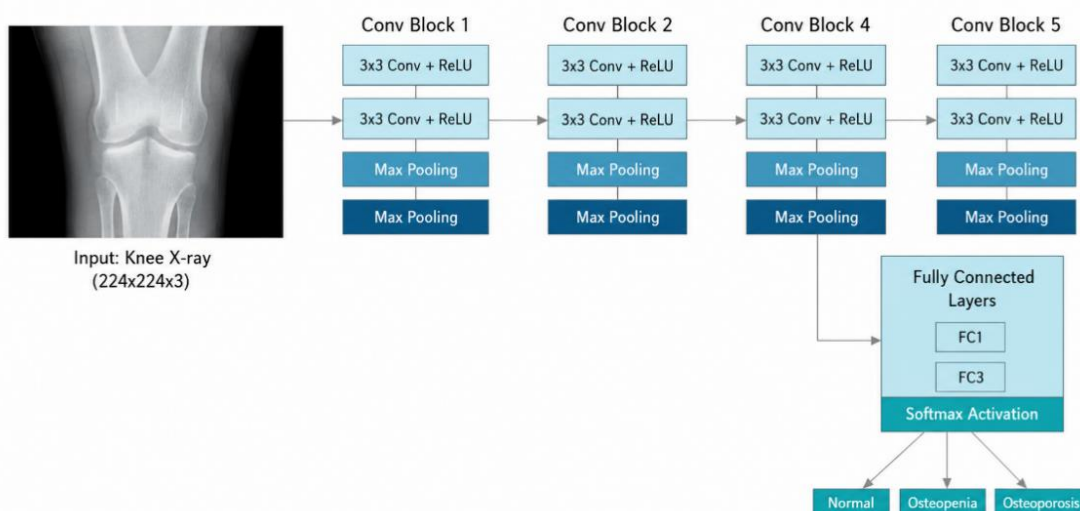
2-сурет. Тізе буынының рентгендік кескіндерін өңдеу және күшейту мысалдары
Ескерту – авторлар құрастырған

Тізе буынының рентгендік кескіндерін, атап айтқанда конволюциялық нейрондық желілерді жіктеу үшін бірнеше терең оқыту архитектураларына салыстырмалы талдау жүргізілді. Бұл жұмыста классикалық архитектуралар VGG16, ResNet50 және Efficientnet, YOLOv8-cls заманауи модельдері қолданылды. Бұл модельдер үш классқа бейімделген: сау сүйек, остеопения және остеопороз.

VGG16 архитектурасы

VGG16 – 13 конволюциялық және 3 толық байланысқан қабаттан тұратын классикалық терең конволюциялық нейрондық желі. Архитектурада 3x3 өлшемді ядролар мен MaxPooling қабаттары қолданылып, белгілердің кеңістіктік өлшемдері біртіндеп азайтылады.

VGG16 архитектурасы медициналық кескіндердің текстуралық және морфологиялық ерекшеліктерін тиімді бөліп алу қабілетіне байланысты салыстырмалы талдау үшін пайдаланылды. VGG16 архитектурасы 3-суретте көрсетілген.

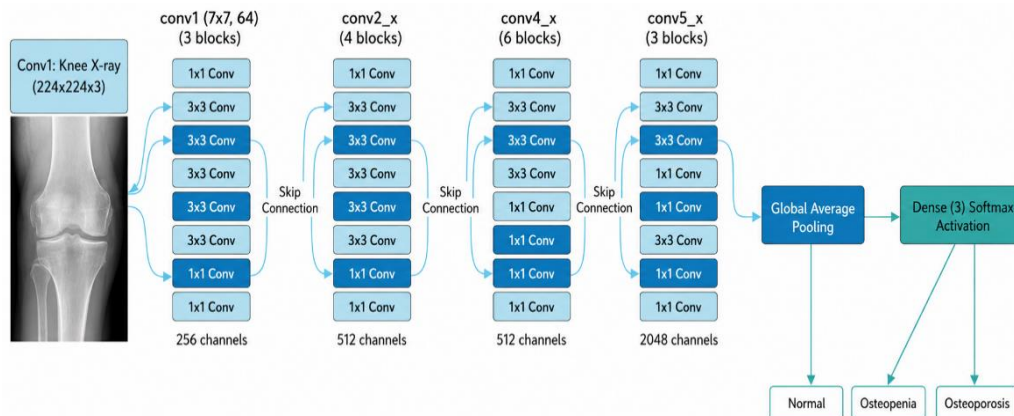


3-сурет. Үш классты жіктеуге бейімделген VGG16 архитектурасы
Ескерту – авторлар құрастырған

VGG16 архитектурасы қарапайым және тұрақты оқытуды қамтамасыз етеді. VGG16 медициналық кескіндердің текстуралық және құрылымдық белгілерін алу қабілетіне ие. Бұл жұмыста модельдерді салыстыру үшін VGG16 таңдалу себебі сүйек тілігінің морфологиялық өзгерістерін талдауға мүмкіндік беруі.

ResNet-50 архитектурасы

ResNet-50 терең конволюциялық нейрондық желіге жатады және құрамында 50 қабат бар. Кескіннің бастапқы сипаттамаларын сақтайтын талдаудың терең кезеңдерінде қалдық оқыту принципін қолданады, бұл әсіресе медициналық кескіндер үшін өте маңызды. ResNet-50 архитектурасы 4-суретте көрсетілген.



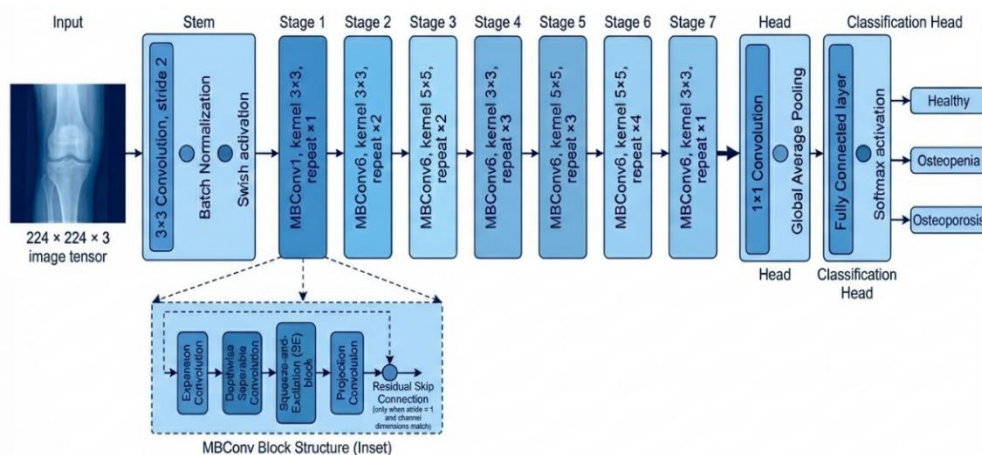
4-сурет. Көпкласты модельге арналған ResNet-50 архитектурасы

Ескерту – авторлар құрастырған

Бұл жұмыста ImageNet көрсеткіші бар алдын-ала дайындалған ResNet-50 моделі қолданылды. 1000 сыныпты жіктеуге арналған бастапқы толық байланысқан қабат Softmax белсендірілген үш нейроннан тұратын жаңа үш класты қабатпен өзгертілді. Қайта оқытудың алдын алу үшін жіктеу қабатының алдында жаһандық орташа мән (Global Average Pooling) қолданылды.

EfficientNet-B0 / B1 архитектурасы

Қабат тереңдігі ұлғайған сайын қуат өсетін дәстүрлі әдістерден айырмашылығы, EfficientNet B0/B1 инновациялық теңдестірілген масштабтау әдісін пайдаланады. Архитектурасы 5-суретте көрсетілген.



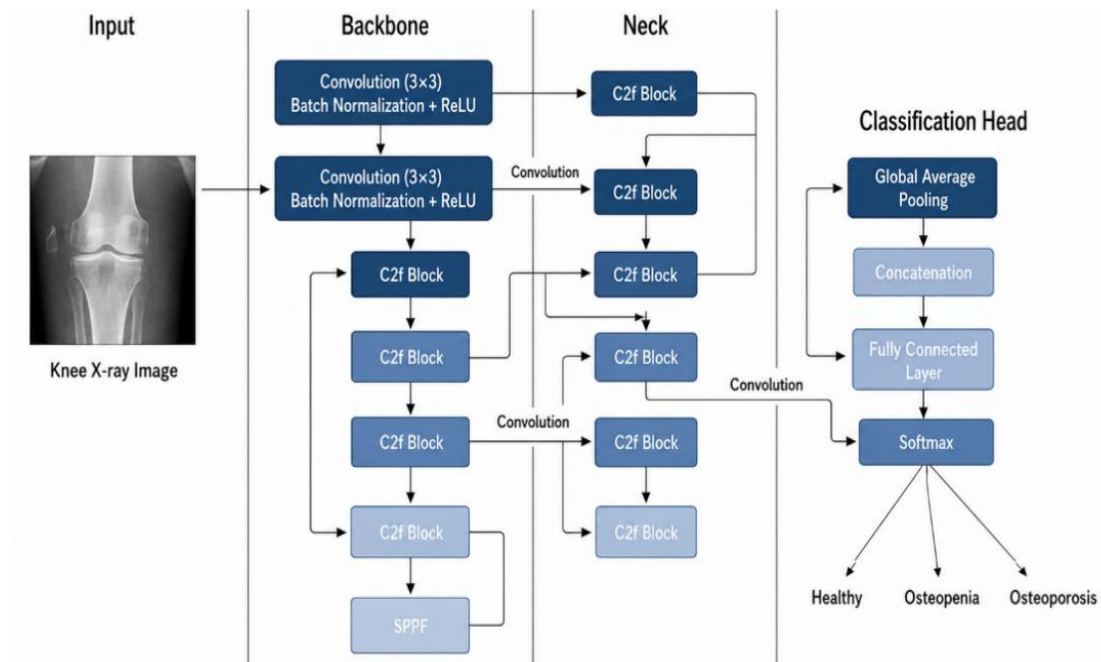
5-сурет. Көпкласты модельге арналған EfficientNet B0/B1 архитектурасы

Ескерту – авторлар құрастырған

EfficientNet-B0/B1 есептеу күрделілігі төмен болған кезде белгілерді тиімді шығаруды қамтамасыз ететін блоктарды пайдаланады және ақпараттық белгілерді күшейтетін механизмді қолданады. Бұл модельде артық оқу қаупін көбейтетін артық қабаттар жоқ.

YOLOv8cls архитектурасы

YOLOv8-cls — классификация міндеттеріне бейімделген заманауи терең оқыту архитектурасы. Модель жоғары есептеу жылдамдығымен қатар күрделі белгілерді тиімді тану қабілетіне ие. YOLOv8-cls архитектурасы жеңілдетілген классификациялық head қолдану арқылы real-time жіктеу міндеттерінде перспективалы тәсілдердің бірі болып табылады. Архитектура 6-суретте көрсетілген.



6-сурет. Көпкласты модельге арналған YOLOv8 архитектурасы

Ескерту – авторлар құрастырған

Иерархиялық тәсіл

Әдеби шолуда көптеген модельдер жақын класстарда тану қателіктерін көрсетеді. Мысалы, остеопения мен остеопороз өте тығыз текстуралық белгілерге ие. Бір сатылы көпкласты жіктеу кезінде остеопения мен остеопороз арасындағы визуалды ұқсастық класаралық шатасудың негізгі себептерінің бірі болып табылады. Осы мәселені азайту мақсатында ResNet-50 архитектурасына негізделген екі сатылы иерархиялық тәсіл ұсынылды.

Бірінші кезеңде модель кескіндерді екі негізгі класқа бөледі: «Healthy» және «Disease». Екінші кезеңде патологиялық кескіндер қосымша түрде Osteopenia және Osteoporosis кластарына жіктеледі. Мұндай тәсіл диагностикалық процесті кезең-кезеңімен ұйымдастырып, жақын класстар арасындағы қателіктер ықтималдығын азайтуға мүмкіндік береді.

Иерархиялық тәсіл клиникалық тұрғыдан ұқсас рентгенологиялық белгілерге ие остеопения мен остеопороз арасындағы класаралық шатасуды төмендетуге бағытталған.

Барлық модельдерді оқыту кезінде batch size = 8 параметрі қолданылды. Модельдер архитектура ерекшеліктеріне байланысты 7-30 эпоха аралығында оқытылды. Оқыту барысында Adam оптимизаторы және categorical cross-entropy шығын функциясы

пайдаланылды. Бастапқы learning rate мәні 1×10^{-4} ретінде орнатылды. Қайта оқытудың алдын алу және оқыту процесін тұрақтандыру мақсатында EarlyStopping және ReduceLROnPlateau callback әдістері қолданылды.

Бағалау көрсеткіштері

Құрылған жіктеуіштердің сапасын бағалау маңызды. Жіктеу сапасын бағалау үшін Accuracy (дұрыстығы), Precision (дәлдігі) және Recall (толықтығы) және басқа көрсеткіштер қолданылады. Accuracy-класс дұрыс анықталған объектілердің үлесі. Accuracy келесі формула бойынша есептеледі (2):

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

мұндағы: TP (True Positive) – модель дұрыс жіктеген белгілі бір классқа тиесілі кескіндер саны;

TN (True Negative) – модель дұрыс түрде басқа кластарға жатқызған кескіндер саны;

FP (False Positive) – басқа классқа тиесілі болғанымен, қате түрде қарастырылып отырған классқа жатқызылған кескіндер саны;

FN (False Negative) – қарастырылып отырған классқа тиесілі болғанымен, модель тарапынан басқа классқа қате жатқызылған кескіндер саны.

Модельдің тиімділігін бағалау үшін қолданылатын көрсеткіштер (3)–(5) теңдеулерінде келтірілген.

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

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

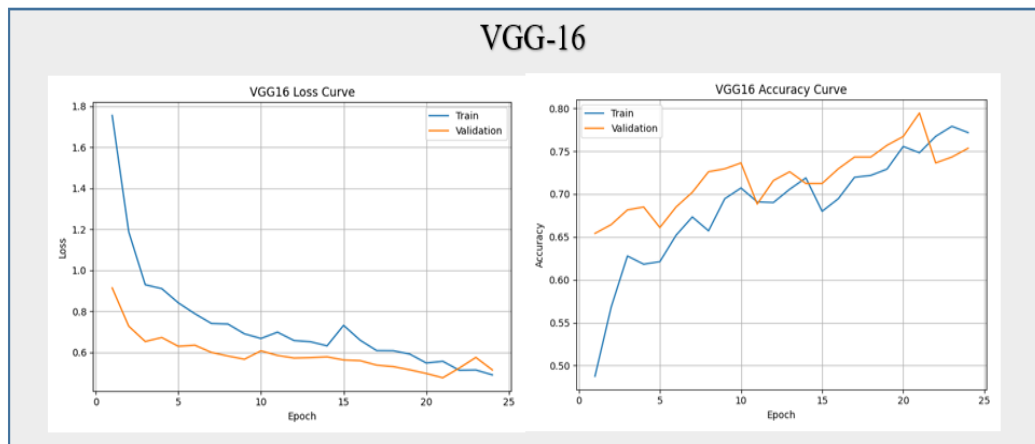
$$F1 = 2 \frac{P \cdot R}{P + R} \quad (5)$$

Жіктеу сапасын талдау үшін precision, recall және F1-score көрсеткіштері де қолданылды, олар әр сынып үшін бөлек есептеледі. Precision белгілі бір классқа жатқызылғандардың барлығында дұрыс танылған суреттердің пайызын көрсетеді. Жіктеу нәтижелерін көрнекі түсіндіру үшін қате матрицасы қолданылды.

НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

Эксперименттер Python бағдарламалау тілі мен TensorFlow кітапханалары (Keras) арқылы жүзеге асырылды. YOLOv8 модельдерін жүзеге асыру үшін ultralytics кітапханасы пайдаланылды. Оқыту Intel Core i5 (12th Gen) процессоры және 16 Гбайт жедел жады бар дербес компьютерде, сондай-ақ есептеулерді жеделдету үшін GPU қолдайтын Google colab ортасында жүргізілді.

7-11 суреттерде жиналған деректер жиынында көпкласты жіктеу үшін оқытылған модельдердің жұмыс нәтижелерін ұсынады. Суреттерде тізе рентгенографиясын жіктеу үшін әртүрлі терең оқыту модельдерінің өнімділігі көрсетілген. Бұл VGG-16, ResNet50, EfficientNet-B0/B1, YOLOv8cls модельдері, сонымен қатар екі кезеңнен тұратын иерархиялық тәсіл. VGG16 моделінің оқыту нәтижелері 7-суретте көрсетілген.

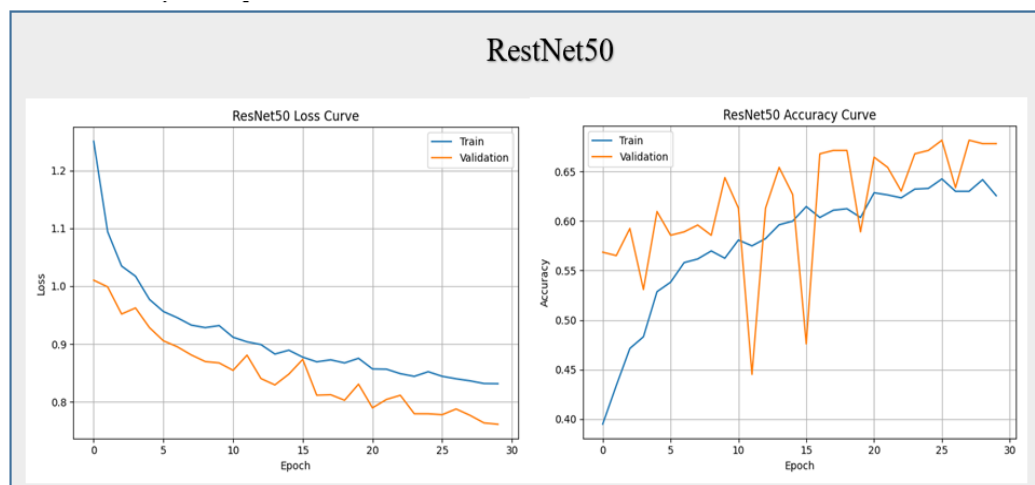


7-сурет. Оқыту қисықтары: жоғалту және дәлдік функциялары

Ескерту – авторлар құрастырған

VGG-16 оқыту барысында оқыту және валидация дәлдігінің біртіндеп өсуі байқалды. Дәлдік қисықтары жақын орналасқан, бұл қайта оқытудың жоқтығын айқын көрсетеді. Сонымен қатар, жоғалту функциясының төмендеуі модельдің деректердегі негізгі құрылымдық белгілерді тиімді үйренгенін дәлелдейді. Бірақ модель остеопения мен остеопороз арасындағы ұқсас рентгенологиялық белгілерді ажыратуда шектеулі нәтижелер көрсетті.

ResNet50 моделінің оқыту нәтижелері 8-суретте көрсетілген.



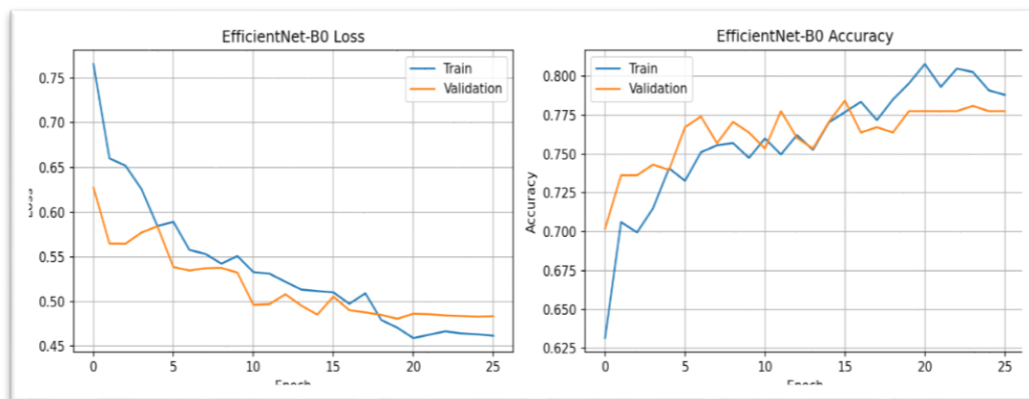
8-сурет. Оқыту қисықтары: жоғалту және дәлдік функциялары

Ескерту – авторлар құрастырған

ResNet50 моделін оқыту кезінде оқу дәлдігінің біртіндеп өсуін байқауға болады. Валидация дәлдігі 0.56-0.68 диапазонында өзгереді, бірақ жекелеген дәуірлерде жергілікті күрт төмендеулер бар. Шығындар функциясы жаттығу шығындарының 1.25-тен 0.83-ке дейін және валидациялық шығындардың 1.01-ден 0.76-ға дейін тұрақты төмендеуін көрсетеді, бұл оқу процесінің конвергенциясын растайды. Бұл жағдайда шығын қисықтары арасында алшақтық жоқ, бұл қайта оқытудың жоқтығын көрсетеді. ResNet-50 архитектурасы барлық модельдер арасында ең жоғары дәлдік пен F1-score көрсеткіштеріне

қол жеткізді. Residual байланыстарды пайдалану модельдің күрделі морфологиялық белгілерді тұрақты түрде үйренуіне мүмкіндік берді.

9-суретте EfficientNet-B0 жоғалту және дәлдік функцияларының қисықтары көрсетілген.

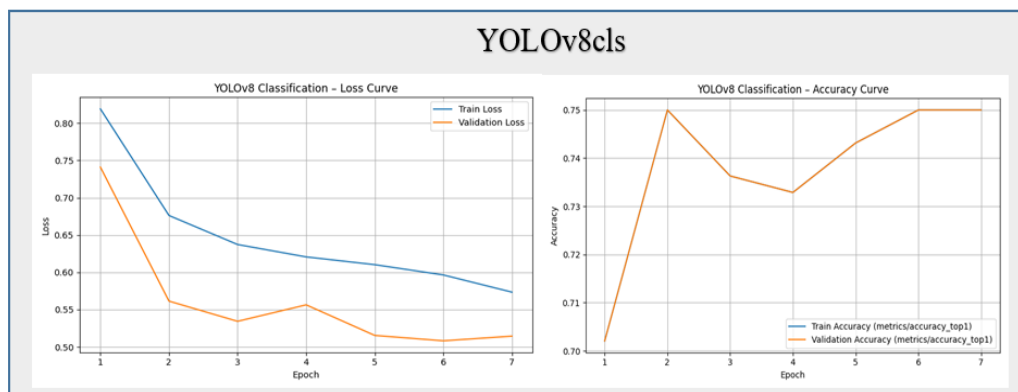


Сурет 9. Оқыту қисықтары: жоғалту және дәлдік функциялары

Ескерту – авторлар құрастырған

EfficientNet-B0 теңдестірілген жаттығулар мен валидация қисықтарымен тұрақты конвергенцияны көрсетеді, бұл айқын шамадан тыс сәйкестіксіз жақсы жалпылауды көрсетеді. Дегенмен, Osteopenia класы үшін кейбір жіктеу қателіктері сақталды, бұл остеопения мен остеопороз арасындағы визуалды ұқсастықпен түсіндіріледі.

10-суретте YOLOv8 жоғалту және дәлдік функцияларының қисықтары көрсетілген.



10-сурет. Оқыту қисықтары: жоғалту және дәлдік функциялары

Ескерту – авторлар құрастырған

Жаттығу және валидация YOLOv8 моделінде жоғалту функциялары жақын, бұл модельдің жақсы жалпылау қабілетін көрсетеді. Яғни, қайта даярлау жоқ, 75% валидация дәлдігі қолайлы, бірақ одан әрі жетілдіруді қажет етеді. YOLOv8-clс архитектурасы жоғары жылдамдықты классификацияға бейімделгенімен, медициналық кескіндердегі ұсақ құрылымдық айырмашылықтарды ажыратуда ResNet-50 және EfficientNet-B0 модельдерінен төмен нәтиже көрсетті.

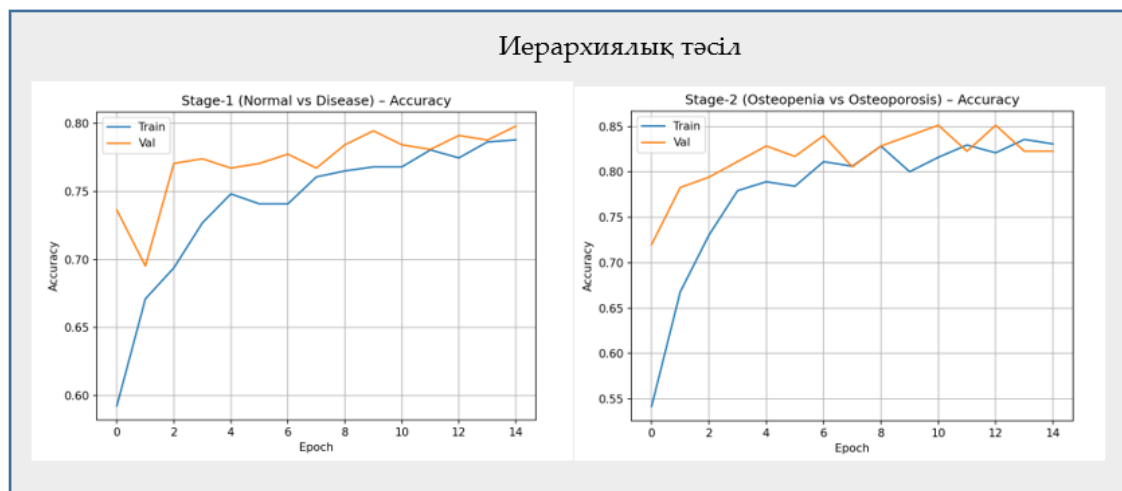
Иерархиялық тәсіл нәтижелері

Тізе буынының рентгендік бейнелерін жіктеу сапасын жақсарту үшін кезең-кезеңмен шешім қабылдауға негізделген иерархиялық тәсіл жүзеге асырылды. Бұл тәсіл клиникалық

жақын жағдайлар, атап айтқанда остеопения мен остеопороз арасындағы класс аралық шатасуды азайтуға бағытталған, олар көбінесе рентгендік суреттерде ұқсас визуалды белгілерді көрсетеді.

Бірінші кезеңде модель кескіндерді «Normal» және «Disease» кластарына екілік жіктеуді үйренді. Бұл кезеңнің негізгі мақсаты қалыпты жағдайларды ерте кезеңде бөліп, кейінгі талдауды тек патологиялық өзгерістері бар кескіндерге бағыттау болды. Оқыту қисықтарын талдау барысында модельдің тұрақты конвергенциясы және айқын қайта оқытудың болмауы байқалды.

Екінші кезеңде тек «Disease» класына жатқызылған кескіндер Osteopenia және Osteoporosis кластарына жіктелді. Бұл кезеңде ResNet-50 архитектурасы қолданылды. Оқыту қисықтары 11 суретте көрсетілген.



11Сурет. Оқыту қисықтары: дәлдік функциялары

Ескерту – авторлар құрастырған

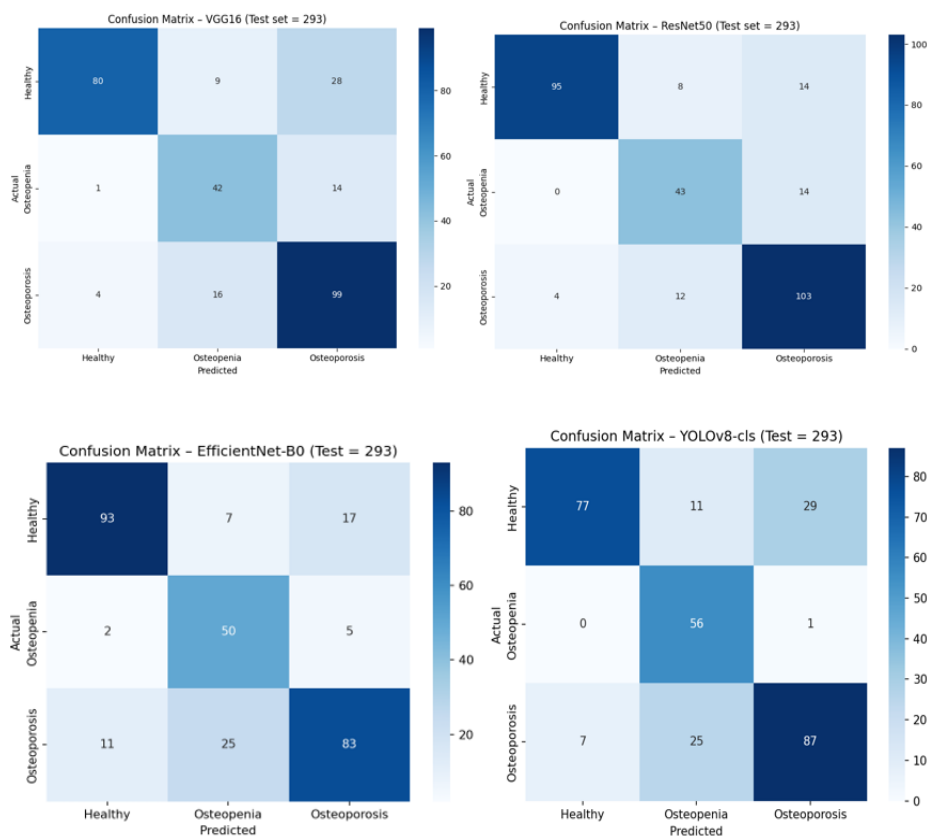
Екінші кезеңдегі дәлдік көрсеткіштері модельдің сүйек тінінің нәзік құрылымдық ерекшеліктеріне көбірек назар аударғанын көрсетті. Жалпы, иерархиялық тәсіл күрделі көпкласты тапсырманы бірнеше қарапайым ішкі міндеттерге бөлу арқылы диагностикалық интерпретацияны жақсартуға мүмкіндік берді.

Иерархиялық тәсілдің жалпы дәлдігі ең жақсы бірсатылы модельден төмен болғанымен, бұл әдіс остеопения мен остеопороз арасындағы шатасуды азайтуда тиімді нәтиже көрсетті. Бұл иерархиялық тәсілдің жалпы дәлдікті арттырудан гөрі клиникалық тұрғыдан маңызды класстар арасындағы қателіктерді төмендетуге бағытталғанын көрсетеді.

Қате матрицаларын талдау

Модельдердің жіктеу сапасын тереңірек бағалау мақсатында тәуелсіз тестілік жиынтықта ($n = 293$) қате матрицалары құрылды. Бұл деректер жиынтығы оқыту және валидация кезеңдерінде пайдаланылған жоқ. Көпкласты модельдердің қате матрицалары 12-суретте көрсетілген.

VGG16 моделі Osteoporosis класы үшін жоғары жіктеу дәлдігін көрсетті. Алайда негізгі қателіктер Osteopenia және Osteoporosis арасында байқалды, бұл олардың ұқсас рентгенологиялық белгілерімен түсіндіріледі. ResNet-50 моделінде Healthy және Osteopenia кластары арасындағы қателіктер байқалмады. Негізгі шатасу Osteopenia және Osteoporosis арасында орын алды. Бұл модель жалпы тұрақтылық пен класстар арасындағы тепе-теңдік бойынша ең жақсы нәтижелерді көрсетті.

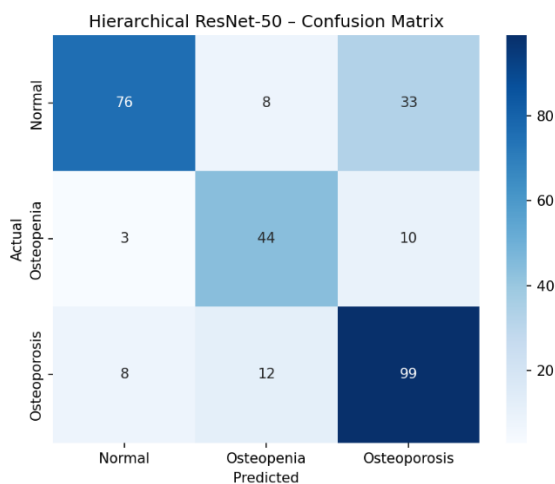


12-сурет. Модельдердің қате матрицасы

Ескерту – авторлар құрастырған

EfficientNet-B0 моделі де тұрақты нәтижелер көрсетті, алайда Osteopenia класы үшін жіктеу қателіктері сақталды. YOLOv8-clс архитектурасы жоғары жылдамдықпен ерекшеленгенімен, дәлдік көрсеткіштері бойынша басқа модельдерден төмен нәтиже көрсетті.

Иерархиялық алгоритм остеопения мен остеопороз арасындағы шатасуды азайтты, остеопорозды анықтау үшін recall көрсеткішін 0.83-ке дейін арттырды. Иерархиялық моделінің қате матрицасы 13-суретте көрсетілген.



13-сурет. Иерархиялық моделінің қате матрицасы

Ескерту – авторлар құрастырған

Бұл нәтиже иерархиялық стратегияның клиникалық скрининг жағдайларында, әсіресе ауыр патологиялық жағдайларды өткізіп алмау маңызды болған кезде, перспективалы тәсіл бола алатынын көрсетеді.

1-кестеде патологияны анықтау модельдерін жан-жақты салыстырмалы бағалау жүргізілді. Барлық модельдер бір сынақ жиынтығында бағаланды (N = 293), бұл салыстырмалы талдаудың дұрыстығын қамтамасыз етеді.

1-кесте. Модельдерді салыстырмалы бағалау

Модельдер	Accuracy, %	Precision, %	Recall, %	F1-score, %	Ерекшеліктері
VGG16	75.4	78.3	75.4	75.7	Негізгі CNN, аралық класстарды ажыратуда шектелген
ResNet50	82.3	83.5	82.3	82.5	Ең жақсы метрикалық тепе-теңдік, жоғары тұрақтылық
EfficientNet-B0	77.1	79.0	77.1	77.4	Шағын архитектура, теңгерімсіздікке сезімтал
YOLOv8-cls	75.1	78.6	75.1	75.2	Жылдамдық үшін оңтайландырылған, төмен дәлдік
Hierarchical	75.0	77.0	75.0	75.0	Классаралық шатасуын азайтады, бірақ жалпы дәлдікті төмендетеді
<i>Ескерту – авторлар құрастырған</i>					

Салыстырмалы талдау нәтижелері бойынша ResNet-50 моделі ең жоғары нәтижелерді көрсетті. Модель 82.3 % ассигасу және 82.5 % F1-score көрсеткіштеріне қол жеткізіп, барлық архитектуралар арасында ең тұрақты нәтиже берді. EfficientNet-B0 моделі есептеу күрделілігі мен дәлдік арасындағы тиімді тепе-теңдікті көрсетті. YOLOv8-cls жоғары жылдамдықты қамтамасыз еткенімен, медициналық кескіндердегі күрделі белгілерді ажыратуда төменірек нәтижелер көрсетті. VGG16 архитектурасы салыстырмалы түрде тұрақты болғанымен, жақын кластар арасындағы шатасуды толық жоя алмады.

Иерархиялық тәсіл остеопения мен остеопороз арасындағы класаралық шатасуды азайтуға мүмкіндік берді. Бірақ жалпы дәлдік көрсеткішінің төмендеуі шешім қабылдау шектерін әрі қарай оңтайландыру қажеттігін көрсетеді. Мұндай тәсіл клиникалық скрининг жағдайларында маңызды болуы мүмкін, себебі ауыр патологиялық жағдайларды өткізіп алмау жалпы ассигасу көрсеткішінен маңыздырақ болып саналады. Бұл нәтиже иерархиялық модельдердің диагностикалық интерпретацияны жақсартуға және клиникалық шешім қабылдау логикасын кезең-кезеңмен ұйымдастыруға мүмкіндік беретінін көрсетеді.

Зерттеудің шектеулері

Зерттеудің бірқатар шектеулері бар. Біріншіден, пайдаланылған деректер жиынтығының көлемі (1947 кескін) терең нейрондық желілер үшін салыстырмалы түрде шектеулі болып табылады. Бұл модельдердің жалпылану қабілетіне әсер етуі мүмкін. Сонымен қатар, алынған нәтижелер кейбір әдебиеттерде келтірілген көрсеткіштерден төмен болды. Бұл, бір жағынан, көптеген зерттеулерде омыртқа немесе жамбас сүйектерінің КТ кескіндері пайдаланылса, осы жұмыста құрылымдық ерекшеліктері

күрделірек тізе буынының рентгенографиялық кескіндерінің қолданылуымен түсіндіріледі.

Екіншіден, деректер жиынтығында класстар арасындағы теңгерімсіздік байқалды, әсіресе остеопения класының көлемі басқа класстармен салыстырғанда аз болды. Мұндай теңгерімсіздік модельдің жиі кездесетін класстарға бейімделуіне себеп болуы мүмкін.

Сонымен қатар, зерттеуде сыртқы валидация жүргізілмеген, бұл нәтижелердің басқа популяциялар үшін жалпылануын шектеуі ықтимал. Болашақ зерттеулердің негізгі бағыттарының бірі ретінде авторлық КТ кескіндер деректер жиынтығын қалыптастыру және терең оқыту модельдерін көпорталықты клиникалық деректерде валидациялау қарастырылады. Бұл тәсіл модельдердің жалпылану қабілетін арттырып, олардың клиникалық практикада қолданылу мүмкіндігін кеңейтуге мүмкіндік береді.

ҚОРЫТЫНДЫ

Осы зерттеу аясында тізе буынының рентгенографиялық кескіндерін жіктеу міндеті үшін әртүрлі терең оқыту архитектураларына салыстырмалы талдау жүргізілді. Эксперименттік нәтижелер барлық қарастырылған модельдердің сүйек тініндегі құрылымдық өзгерістерге байланысты ақпараттық белгілерді тиімді бөліп алуға қабілетті екенін көрсетті. Алайда модельдер жіктеу дәлдігі, тұрақтылығы және класаралық шатасу деңгейі бойынша әртүрлі нәтижелер көрсетті.

Зерттеу нәтижелері бойынша ResNet-50 моделі ең жоғары сапа көрсеткіштеріне қол жеткізді және барлық класстар үшін precision, recall және F1-score метрикаларының ең теңгерімді мәндерін қамтамасыз етті. VGG16 және EfficientNet-B0 архитектуралары да қанағаттанарлық нәтижелер көрсетті, алайда Osteopenia класы үшін жіктеу тұрақтылығы төменірек болды. YOLOv8-cls моделі жоғары есептеу жылдамдығын қамтамасыз еткенімен, күрделі медициналық белгілерді ажыратуда басқа архитектуралардан төмен нәтиже көрсетті.

ResNet-50 архитектурасына негізделген иерархиялық тәсіл остеопения мен остеопороз арасындағы класаралық шатасуды азайтуға мүмкіндік берді. Дегенмен, бұл тәсіл жалпы ассигасу көрсеткішінің төмендеуімен қатар жүрді. Бұл нәтиже иерархиялық модельдердің клиникалық тұрғыдан маңызды қателіктерді азайтуда перспективалы екенін, бірақ шешім қабылдау шектерін әрі қарай оңтайландыру қажеттігін көрсетеді.

Жалпы алғанда, алынған нәтижелер остеопороз диагностикасын компьютерлік қолдау жүйелерінде терең конволюциялық нейрондық желілерді қолданудың тиімділігін растады. Болашақ зерттеулерде деректер жиынтығын кеңейту, көпорталықты сыртқы валидация жүргізу, клиникалық параметрлерді интеграциялау, сондай-ақ авторлық КТ кескіндер жиынтығын пайдалану жоспарлануда. Сонымен қатар, ансамбльдік және көпдеңгейлі модельдерді зерттеу диагностикалық дәлдікті әрі қарай арттыруға мүмкіндік беруі мүмкін.

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Автор(лар) мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: Бұл зерттеу сыртқы гранттық қаржыландырусыз орындалды.

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ: Зерттеу ашық қолжетімді деректер жиынтығын пайдалану негізінде жүргізілді және адам қатысушылары мен жануарларға тікелей эксперименттік араласуды қамтымады. Осыған байланысты Институционалдық этика комитетінің (IRB) арнайы мақұлдауы талап етілмеді.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: Бұл зерттеуде адам қатысушылары тікелей тартылмағандықтан, ерікті келісім алу талап етілмеді.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Осы зерттеуде пайдаланылған деректер ашық қолжетімді Kaggle платформасынан алынды. Зерттеу нәтижелерін растайтын өңделген деректер мен эксперименттік нәтижелер тиісті сұраныс бойынша жауапты автордан алынуы мүмкін.

АЛҒЫС БІЛДІРУ: Авторлар мақаланы дайындау барысында құнды ұсыныстар мен сын-пікірлер берген рецензенттерге алғыс білдіреді. Сондай-ақ, зерттеу жұмысының ғылыми және әдістемелік сапасын арттыруға қолдау көрсеткен әріптестерге ризашылықтарын білдіреді.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Авторлар қолжазбаны дайындау барысында генеративті жасанды интеллект құралдарын әдебиеттер тізімін APA 7 стандартына сәйкес форматтау мақсатында пайдаланды. Жасанды интеллект ғылыми нәтижелерді интерпретациялау, қорытынды жасау немесе зерттеу нәтижелерін қалыптастыру үшін қолданылған жоқ.

ӘДЕБИЕТТЕР ТІЗІМІ

- Adarve-Castro, A., Soria-Utrilla, V., Castro-García, J. M., Domínguez-Pinos, M. D., Sendra-Portero, F., Ruiz-Gómez, M. J., & Algarra-García, J. (2025). Advanced radiomic prediction of osteoporosis in primary hyperparathyroidism: A machine learning-based analysis of CT images. *Radiologia Medica*. <https://doi.org/10.1007/s11547-025-02004-z>
- Du, C., He, J., Zhang, J., Wang, J., & Li, G. (2025). Automated opportunistic screening for osteoporosis using deep learning-based automatic segmentation and radiomics on proximal femur images from low-dose abdominal CT. *BMC Musculoskeletal Disorders*, 26(1), 378. <https://doi.org/10.1186/s12891-025-08631-x>
- Dzierżak, R., & Omiotek, Z. (2022). Application of deep convolutional neural networks in the diagnosis of osteoporosis. *Sensors*, 22(21), 8189. <https://doi.org/10.3390/s22218189>
- Jang, R., Choi, J. H., Kim, N., Chang, J. S., Yoon, P. W., & Kim, C.-H. (2021). Prediction of osteoporosis from simple hip radiography using deep learning algorithm. *Scientific Reports*, 11, 19997. <https://doi.org/10.1038/s41598-021-99549-6>
- Kaggle. (2024). Knee osteoporosis dataset. <https://www.kaggle.com/datasets/-/866059b7930a5c49cd77d94c1761840a19d88074cad74e8f0e0cfa2b236a6904>
- Lim, H.K., Ha, H.I., Park, S.-Y., & Han, J. (2021). Prediction of femoral osteoporosis using machine-learning analysis with radiomics features and abdomen-pelvic CT: A retrospective single center preliminary study. *PLoS ONE*, 16, e0247330. <https://doi.org/10.1371/journal.pone.0247330>
- Liu, L., Si, M., Ma, H., Cong, M., Xu, Q., Sun, Q., Wu, W., Wang, C., Fagan, M. J., Mur, L. A. J., et al. (2022). A hierarchical opportunistic screening model for osteoporosis using machine learning applied to clinical data and CT images. *BMC Bioinformatics*, 23, 63. <https://doi.org/10.1186/s12859-022-04596-z>
- Liu, Z., Li, Y., Zhang, C., Xu, H., Zhao, J., Huang, C., Chen, X., & Ren, Q. (2025). Radiomics and machine learning for osteoporosis detection using abdominal computed tomography: A retrospective multicenter study. *BMC Medical Imaging*, 25(1), 235. <https://doi.org/10.1186/s12880-025-01743-9>
- Madiyeva, M., Rymbayeva, T., Kaskabayeva, A., Bersimbekova, G., Kanapiyanova, G., Prilutskaya, M., Akhmetzhanova, D., Alimbayeva, A., & Omarov, N. (2024). The prevalence and risk factors of low bone mineral density in the population of the Abay region of Kazakhstan. *International Journal of Environmental Research and Public Health*, 21, 681. <https://doi.org/10.3390/ijerph21060681>
- Naguib, S. M., Saleh, M. K., Hamza, H. M., Hosny, K. M., & Kassem, M. A. (2024). A new superfluity deep learning model for detecting knee osteoporosis and osteopenia in X-ray images. *Scientific Reports*, 14, 25434. <https://doi.org/10.1038/s41598-024-75549-0>

- Pan, J., Lin, P. C., Gong, S. C., Wang, Z., Cao, R., Lv, Y., Zhang, K., & Wang, L. (2024). Feasibility study of opportunistic osteoporosis screening on chest CT using a multi-feature fusion DCNN model. *Archives of Osteoporosis*, 19(1), 98. <https://doi.org/10.1007/s11657-024-01455-7>
- Pan, Y., Shi, D., Wang, H., Chen, T., Cui, D., Cheng, X., & Lu, Y. (2020). Automatic opportunistic osteoporosis screening using low-dose chest computed tomography scans obtained for lung cancer screening. *European Radiology*, 30(7), 4107–4116. <https://doi.org/10.1007/s00330-020-06679-y>
- Park, H., Kang, W. Y., Woo, O. H., Lee, J., Yang, Z., & Oh, S. (2024). Automated deep learning-based bone mineral density assessment for opportunistic osteoporosis screening using various CT protocols with multi-vendor scanners. *Scientific Reports*, 14, 18712. <https://doi.org/10.1038/s41598-024-73709-w>
- Pickhardt, P. J., Nguyen, T., Perez, A. A., Graffy, P. M., Jang, S., Summers, R. M., & Garrett, J. W. (2022). Improved CT-based osteoporosis assessment with a fully automated deep learning tool. *Radiology: Artificial Intelligence*, 4(5), e220042. <https://doi.org/10.1148/ryai.220042>
- Salari, N., Ghasemi, H., Mohammadi, L., Behzadi, M. H., Rabieenia, E., Shohaimi, S., & Mohammadi, M. (2021). The global prevalence of osteoporosis in the world: A comprehensive systematic review and meta-analysis. *Journal of Orthopaedic Surgery and Research*, 16, 609. <https://doi.org/10.1186/s13018-021-02772-0>
- Saravi, B., Zink, A., Tabukashvili, E., Güzel, H. E., Ülkümen, S., Couillard-Despres, S., Lang, G. M., & Hassel, F. (2024). Integrating radiomics with clinical data for enhanced prediction of vertebral fracture risk. *Frontiers in Bioengineering and Biotechnology*, 12, 1485364. <https://doi.org/10.3389/fbioe.2024.1485364>
- Sebro, R., & De la Garza-Ramos, C. (2022). Machine learning for opportunistic screening for osteoporosis from CT scans of the wrist and forearm. *Diagnostics*, 12, 691. <https://doi.org/10.3390/diagnostics12030691>
- Siddharth, M., Arora, G., & Vani, M. P. (2025). Predictive modelling of knee osteoporosis. *BMC Research Notes*, 18, 114. <https://doi.org/10.1186/s13104-025-07125-2>
- Wang, J., He, Y., Yan, L., Chen, S., & Zhang, K. (2024). Predicting osteoporosis and osteopenia by fusing deep transfer learning features and classical radiomics features based on single-source dual-energy CT imaging. *Academic Radiology*. <https://doi.org/10.1016/j.acra.2024.04.022>
- Yuan, J., Li, B., Zhang, C., Wang, J., Huang, B., & Ma, L. (2025). Machine learning-based CT radiomics model to predict the risk of hip fragility fracture. *Academic Radiology*, 32(5), 2854–2862. <https://doi.org/10.1016/j.acra.2025.01.023>
- Zhang, K., Lin, P., Pan, J., Xu, P., Qiu, X., Crookes, D., Hua, L., & Wang, L. (2023). End-to-end multitask joint learning model for osteoporosis classification in CT images. *Computational Intelligence and Neuroscience*, 3018320. <https://doi.org/10.1155/2023/3018320>
- Шарманов, Т.Ш., Шакиева, Р.А., & Сулейменова, Д.С. (2012). Распространенность остеопороза среди населения старших возрастных групп по данным сонографических скрининговых исследований в Республике Казахстан. *Медицина*, 3, 2–5. // Sharmanov, T. Sh., Shakiyeva, R. A., & Suleymenova, D. S. (2012). Rasprostranennost osteoporoza sredi naseleniya starshikh vozrastnykh grupp po dannym sonograficheskikh skринingovykh issledovaniy v Respublike Kazakhstan [Prevalence of osteoporosis among older population groups according to sonographic screening studies in the Republic of Kazakhstan]. *Meditsina*, 3, 2–5. (In Russ.)
- Турекулова, А. А. (2008). Медико-социальные аспекты остеопороза у городского населения [Автореферат диссертации]. Алматы. // Turekulova, A. A. (2008). Mediko-sotsialnye aspekty osteoporoza u gorodskogo naseleniya [Medical and social aspects of osteoporosis among urban population] [Dissertation abstract]. Almaty. (In Russ.)

Google. (2026). Ғылыми әдебиеттер тізімін APA 7 стандартына сәйкес форматтау бойынша көмек [Үлкен тілдік үлгі]. Gemini. <https://gemini.google.com/>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Нурекенова Арайлым Солтахановна – «Ақпараттық жүйелер» мамандығының докторанты, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан
Нурекенова Арайлым Солтахановна – докторант по специальности

«Информационные системы», Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Nurekenova Arailym Soltahanovna – PhD student in Information Systems, East Kazakhstan Technical University named after D. Serikbaev, Ust-Kamenogorsk, Kazakhstan,
e-mail: soltahanovna@gmail.com,
ORCID: <https://orcid.org/0009-0003-9873-2911>



Құмарғажанова Сәуле Құмарғажанқызы – т.ғ.к., профессор, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Кумаргажанова Сауле Кумаргажановна – кандидат технических наук, профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Kumargazhanova Saule – Candidate of Technical Sciences, Professor, East Kazakhstan Technical University named after D. Serikbaev, Ust-Kamenogorsk, Kazakhstan
e-mail: skumargazhanova@gmail.com,
ORCID: <https://orcid.org/0000-0002-6744-4023>



Смаилова Сауле Сансызбаевна – PhD, қауымдастырылған профессор, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Смаилова Сауле Сансызбаевна – PhD, ассоциированный профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Smailova Saule – PhD, Associate Professor, East Kazakhstan Technical University named after D. Serikbaev, Ust-Kamenogorsk, Kazakhstan
e-mail: ssmailova@edu.ektu.kz,
ORCID: <https://orcid.org/0000-0002-8411-3584>



Dzierżak Róża – PhD, доцент,
Люблин техникалық университеті, Люблин қ., Польша
Dzierżak Róża – PhD, доцент,
Люблинский технический университет, г. Люблин, Польша
Dzierżak Róża – PhD, Associate Professor,
Lublin University of Technology, Lublin, Poland
e-mail: r.dzierzak@pollub.pl,
ORCID: <https://orcid.org/0000-0001-5640-0204>



Нурсадыкова Роза Канашевна – аға оқытушы, Д. Серікбаев
атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ.,
Қазақстан
Нурсадыкова Роза Канашевна – старший преподаватель,
Восточно-Казakhstanский технический университет имени
Д. Серикбаева, г. Усть-Каменогорск
Nursadykova Roza – Senior Lecturer, East Kazakhstan
Technical University named after D. Serikbaev, Ust-Kamenogorsk,
Kazakhstan
e-mail: RNursadykova@ektu.kz,
ORCID: <https://orcid.org/0009-0005-9595-6826>

https://doi.org/10.51885/3134-8025_IICS_2026_2_10
SRSTI 20.53.41

INTEGRATED INDICATORS FOR SMART CITY EVALUATION: A MULTIDIMENSIONAL SUSTAINABILITY ASSESSMENT FRAMEWORK

«АҚЫЛДЫ ҚАЛАНЫ» БАҒАЛАУ ҮШІН ИНТЕГРАЦИЯЛАНҒАН
КӨРСЕТКІШТЕР: ТҰРАҚТЫ ДАМУДЫ КӨПӨЛШЕМДІ БАҒАЛАУ МОДЕЛІ

ИНТЕГРИРОВАННЫЕ ПОКАЗАТЕЛИ ДЛЯ ОЦЕНКИ «УМНОГО ГОРОДА»:
МНОГОМЕРНАЯ МОДЕЛЬ ОЦЕНКИ УСТОЙЧИВОГО РАЗВИТИЯ

T. Salim  ^{1*}

¹Kazakh-British Technical University, Almaty, Kazakhstan

Keywords:

Smart city; Sustainability assessment; Integrated indicators; Urban development; Public perception; Digital infrastructure; Data-driven evaluation; Kazakhstan

ABSTRACT

Urban digital transformation requires advanced frameworks for urban sustainability assessment. This study develops a multidimensional model for evaluating smart cities using integrated indicators across five key domains: social, economic, environmental, technological, and governance. Empirical data were gathered through a structured survey of 165 residents in Almaty and Astana, Kazakhstan. The methodology incorporated data normalization, comparative statistical analysis, and public perception evaluation. Results demonstrate strong citizen prioritization of smart mobility solutions and environmental quality, while revealing significant gaps in digital infrastructure development. The proposed framework offers a practical, citizen-centered tool for assessing smart city readiness in the Kazakhstani context. This approach enables evidence-based urban planning and supports strategic decision-making for sustainable city development, addressing both technological and social dimensions of smart transformation.

Түйінді сөздер:

Ақылды қала,
Тұрақтылықты бағалау,
Интеграцияланған
индикаторлар, Қалалық
даму, Қоғамдық пікір,
Цифрлық
инфрақұрылым,
Деректерге негізделген
талдау, Қазақстан.

ТҮЙІНДЕМЕ

Қалалардың цифрлық трансформациясы тұрақтылықты кешенді бағалауға арналған заманауи әдіснамаларды талап етеді. Бұл зерттеу әлеуметтік, экономикалық, экологиялық, технологиялық және басқару салалары бойынша кешенді көрсеткіштерге негізделген «ақылды қаланы» бағалаудың көпөлшемді моделін ұсынады. Эмпирикалық деректер Алматы және Астана қалаларының 165 тұрғыны арасында жүргізілген сауалнама нәтижелері жинақталды. Зерттеу әдіснамасы деректерді нормалау, салыстырмалы статистикалық талдау және қоғамдық пікірді бағалауды қамтыды. Нәтижелер азаматтардың ақылды көлік жүйелері мен экологиялық сапаға жоғары басымдылығын, сонымен қатар цифрлық инфрақұрылымның дамуындағы айтарлықтай олқылықтарды анықтады. Ұсынылған модель Қазақстан контекстінде ақылды қалалардың дайындығын бағалау үшін тәжірибелік, азаматтарға бағытталған құрал ретінде қызмет етеді. Бұл тәсіл тұрақты қалалық даму үшін деректерге негізделген жоспарлау мен стратегиялық шешім қабылдауды қолдайды.



© 2026 T. Salim

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>

Ключевые слова:

Умный город, Оценка устойчивости, Интегрированные индикаторы, Городское развитие, Общественное восприятие, Цифровая инфраструктура, Оценка на основе данных, Казахстан.

АННОТАЦИЯ

Цифровая трансформация городов требует современных методологических подходов к оценке устойчивости городской среды. Данное исследование разрабатывает многомерную модель оценки «умного города» на основе интегрированных индикаторов по пяти ключевым направлениям: социальному, экономическому, экологическому, технологическому и управленческому. Эмпирические данные были собраны посредством структурированного опроса 165 жителей Алматы и Астаны, Казахстан. Методология исследования включала нормализацию данных, сравнительный статистический анализ и оценку общественного восприятия. Результаты демонстрируют высокий приоритет граждан в развитии умной мобильности и экологического качества, одновременно выявляя значительные пробелы в развитии цифровой инфраструктуры. Предложенная модель служит практическим инструментом, ориентированным на граждан, для оценки готовности умных городов в казахстанском контексте. Такой подход поддерживает принятие основанных на данных решений и стратегическое планирование для обеспечения устойчивого городского развития.

INTRODUCTION

The shape of contemporary cities and the tenets of their governance are being altered by the continuous processes of urbanization and digital development. The evaluation of urban sustainability has become a strategic priority due to population growth, the development of smart infrastructure, the expansion of transportation networks, and the requirement for effective resource management. In order to improve the quality of life for its residents, the idea of a smart city combines innovation, environmental responsibility, and efficient governance. However, an integrated assessment technique that incorporates technological, economic, environmental, and social aspects—including public awareness and participation—is necessary for the development and management of such cities.

A number of approaches and indices, such the ISO 37122 standards, the IESE Cities Motion Index, and the IMD Smart City Index, have been established worldwide to assess the performance of smart cities. Nevertheless, the majority of current frameworks frequently ignore the psychological and perceptual aspects of urban living in favor of physical and economic measures. Evaluating the true impact and preparedness of smart cities has become more important in Kazakhstan, where digitalization initiatives like eGov, Smart Astana, and Open Almaty are being aggressively implemented. The establishment of a fully integrated and sustainable smart city ecosystem is hampered by the low level of citizen knowledge and engagement despite notable technological advancements.

A thorough evaluation framework that incorporates the various facets of sustainability – social, economic, technological, environmental, and governance-related – is therefore becoming more and more necessary. The objective of this project is to use empirical data gathered from Almaty and Astana citizens to design and test an integrated model for smart city assessment. The study provides a multifaceted approach to evaluating smart city readiness by utilizing techniques of data normalization, comparison analysis, and public perception assessment. The results are meant to help urban planners and politicians in Kazakhstan develop sustainable and citizen-centered smart city policies by supporting evidence-based decision-making.

Applicability of the content: In the digital era, cities are evolving into complex ecosystems that rely on continuous data exchange between technological infrastructure and human activities. The applicability of this study lies in its practical relevance for emerging urban environments such as Almaty and Astana, which are currently implementing smart city initiatives under national

digital transformation programs. By analyzing the interconnections among social, environmental, and technological dimensions, the proposed framework provides a foundation for assessing both infrastructure readiness and citizens' perceptions of urban sustainability. This multidimensional evaluation model supports policymakers in identifying priority areas for improvement—such as public transportation, environmental protection, and digital inclusivity—and contributes to more efficient urban management practices.

Work relevance. The relevance of this work is rooted in the global shift toward data-driven and sustainable urban governance. As highlighted by Yigitcanlar et al. (2021), cities that combine advanced technology with sustainability principles can significantly enhance social equity and environmental resilience. In Kazakhstan, the development of smart city projects such as Smart Astana, Open Almaty, and eGov reflects a growing demand for transparent governance and citizen-oriented digital services. However, the absence of an integrated measurement system complicates the evaluation of their true effectiveness. This research addresses that gap by proposing a comprehensive framework that combines measurable indicators and subjective public feedback. The findings are relevant not only for Kazakhstan but also for other developing countries facing similar urbanization and digital transformation challenges.

Purpose of the work. The main purpose of this study is to develop a theoretical and practical framework for assessing urban sustainability based on integrated smart city indicators. The research aims to analyze existing evaluation models, identify gaps related to social perception and data integration, and propose a multidimensional assessment system that connects technical performance with citizen satisfaction. Furthermore, the study seeks to determine how these combined indicators can guide strategic decision-making for sustainable city development and digital governance improvement in Kazakhstan.

LITERATURE REVIEW

Urban sustainability and smart city assessment have become key research areas in the post-pandemic era, driven by rapid digital transformation and environmental challenges. Modern studies emphasize that smart city development requires a comprehensive evaluation approach combining technological, environmental, and social indicators (Yigitcanlar et al., 2021). According to Bibri and Krogstie (2021), sustainability in urban systems can only be achieved through integrated data-driven management that aligns digital innovation with ecological balance and citizen well-being. Similarly, Bifulco et al. (2022) underline the importance of governance and public participation as critical dimensions of sustainable smart city performance.

Recent research by Neirotti and De Marco (2022) argues that cities are increasingly adopting AI-based systems to manage energy efficiency, transportation, and environmental monitoring, yet most frameworks overlook human-centered indicators. The integration of subjective data, such as public perception and user satisfaction, into traditional performance metrics is therefore becoming a major focus in contemporary research (Kumar & Singh, 2023). As highlighted by Li et al. (2023), the combination of open data analytics, IoT technologies, and citizen participation creates a new paradigm for sustainability evaluation, allowing cities to develop adaptive and evidence-based urban strategies.

MATERIALS AND METHODS

Research Design - Quantitative-descriptive research design

By using a quantitative-descriptive research design, the study aims to provide an objective and unbiased assessment of urban sustainability dynamics in Kazakhstan without manipulating any variables. This approach preserves the authenticity of respondents' perceptions while enabling a systematic analysis of empirical survey data. Descriptive statistical analysis is used in conjunction with quantitative methods to find trends in city dwellers' awareness, satisfaction,

and preparedness for smart city development. Systematic data collection, normalization, and integration across several sustainability dimensions are made possible by the use of this design.

Research methodology. The methodological framework integrates a comparative evaluation of smart city metrics with quantitative survey data analysis. A structured online questionnaire was used to collect data from citizens of major Kazakhstani cities, mainly Almaty and Astana. A total of 165 valid replies were obtained. To provide a fair assessment across the five sustainability domains—social, economic, environmental, technological, and governance—the gathered data were processed using normalization techniques and weighted scoring via the Analytic Hierarchy Process (AHP). To match local indicators with international standards, a comparative content study of international frameworks (ISO 37122, IESE, and IMD) was carried out. The suggested assessment model's validity, reliability, and reproducibility are guaranteed by the combination of these techniques.

Urban Data Collection and Study Design

The study focuses on assessing the multidimensional sustainability of smart cities in Kazakhstan through an integrated framework that combines objective indicators and public perception data. The empirical component of the research was conducted between January and March 2025 in the cities of Almaty and Astana, which represent Kazakhstan's two leading digital transformation hubs. The research design is quantitative and descriptive, structured to capture both measurable performance indicators and citizens' subjective evaluation of smart city readiness.

Justification of City Selection and Data Sources. Almaty and Astana were selected as case studies due to their leading roles in Kazakhstan's digital transformation and availability of reliable smart city data. The selection criteria included population size, economic activity, and digital infrastructure maturity. Comparative information from 102 international cities was used for benchmarking Kazakhstan's performance. The data were collected from UN-Habitat (2024), World Bank Open Data (2023), Eurostat, and the ITU Smart City Index. The questionnaire included 25 questions grouped into five sustainability domains:

1. Social Dimension – education, digital literacy, inclusiveness, and public engagement;
2. Economic Dimension – employment opportunities, innovation support, and local entrepreneurship;
3. Environmental Dimension – waste management, air quality, and green infrastructure;
4. Technological Dimension – accessibility of digital services, internet coverage, and data transparency;
5. Governance Dimension – openness of government, public participation, and administrative efficiency.

The purpose of this structure was to evaluate how citizens perceive and interact with smart city systems, and to link their responses to objective sustainability performance indicators derived from open municipal and national datasets (eGov.kz, Open Data Portal Kazakhstan, and Smart Astana Dashboard).

Respondent Characteristics

The study sample consisted of 165 respondents from the cities of Almaty and Astana. The demographic structure of participants was analyzed to ensure the representativeness and diversity of the dataset.

In terms of age distribution, most respondents were between 26–40 years (41%), followed by 18–25 years (34%), 41–60 years (19%), and a smaller proportion aged over 60 (6%).

Regarding gender, 58% of respondents were female and 42% male, providing a balanced perspective across genders.

In terms of educational background, a significant proportion of participants held higher education degrees (72%), while 28% had secondary education. This reflects a relatively educated sample, which is relevant given the topic of digital and smart city development.

With respect to employment status, 64% of respondents were employed, 21% were students, and 15% belonged to other categories, including retirees and temporarily unemployed individuals.

Overall, the sample includes individuals of different age groups, professional statuses, and education levels, allowing for a comprehensive assessment of public perceptions regarding smart city development.

Such diversity ensures that the findings reflect perspectives from students, working professionals, and older population groups with varying levels of digital literacy.

Indicator Framework and Analytical Model

To ensure comparability between the cities, the study applies a weighted indicator model based on the Analytic Hierarchy Process (AHP) (Saaty, 2021). This approach enables a structured evaluation of smart city sustainability across multiple dimensions by assigning relative importance to each domain.

The assessment framework includes five key domains: social, economic, environmental, technological, and governance. Each indicator within these domains was normalized using the min-max method to ensure comparability on a unified scale ranging from 0 to 1.

The composite Smart City Sustainability Index (SCSI) was calculated as follows:

$$SCSI = \sum_{i=1}^n \omega_i \cdot x_i \quad (1)$$

where: ω_i – weight of indicator i (derived using AHP);

x_i – normalized value of indicator i ;

n – total number of indicators.

This formulation ensures that each indicator contributes proportionally to the final sustainability score based on its relative importance.

Analytic Hierarchy Process (AHP) Implementation

To determine the weights of the five sustainability domains, the Analytic Hierarchy Process (AHP) was applied following a structured multi-step procedure.

First, a hierarchical model was constructed, consisting of the overall goal (smart city sustainability assessment), the five evaluation domains, and their respective indicators.

Second, expert evaluation was conducted involving six specialists in urban planning, digital governance, and information technology. Experts performed pairwise comparisons of the five domains using the Saaty scale (1–9), reflecting the relative importance of each criterion.

Third, the pairwise comparison matrix was constructed by aggregating expert judgments. Based on this matrix, normalized weights were calculated using the eigenvector method.

To ensure the reliability of the comparisons, the Consistency Index (CI) and Consistency Ratio (CR) were computed:

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

$$CR = \frac{CI}{RI}$$

The resulting consistency ratio (CR = 0.08) is below the acceptable threshold of 0.1, confirming that the expert judgments are consistent and reliable.

The final weights derived from AHP were then applied in the calculation of the composite Smart City Sustainability Index (SCSI).

A comparative analysis between Almaty and Astana was conducted to identify strengths and gaps across the five domains. Additionally, data visualization techniques were applied using Python (Matplotlib and Pandas) to generate radar charts, histograms, and correlation heatmaps illustrating citizens' priorities and satisfaction levels.

The integrated indicator framework presented in Table 1 provides a structured basis for evaluating smart city sustainability. Each indicator reflects a critical aspect of urban performance, linking objective measures such as digital infrastructure and environmental quality with subjective public perceptions collected through the survey. This multidimensional approach allows for identifying both technological strengths and social gaps in urban development. By combining open data sources with citizen feedback, the framework ensures a comprehensive and human-centered assessment of smart city sustainability.

Table 1. Sample of Integrated Indicator Framework

Domain	Key Indicators (Examples)	Data Source / Measurement Method
Social	Digital literacy, access to online education, inclusiveness	Survey responses, eGov Analytics
Economic	Innovation activity, startup support, ICT employment rate	National Statistics Committee, Survey
Environmental	Air quality index, waste recycling, green zone area	Smart Astana, OpenData.kz
Technological	Internet speed, service accessibility, cybersecurity awareness	Questionnaire, Network Reports
Governance	Transparency index, open data usage, citizen satisfaction	Open Almaty Portal, Survey data
<i>Note – compiled by the author based on personal analysis and interpretation</i>		

Table 2. Pairwise Comparison Matrix (AHP)

Criteria	Social	Economic	Environmental	Technological	Governance
Social	1	1/2	1/3	1/2	1
Economic	2	1	1/2	1	2
Environmental	3	2	1	2	3
Technological	2	1	1/2	1	2
Governance	1	1/2	1/3	1/2	1

Table 3. Normalized Weights of Criteria

Domain	Weight
Social	0.17
Economic	0.20
Environmental	0.25
Technological	0.21
Governance	0.17

Table 4. Descriptive Statistics of Survey Data

Dimension	Mean	Standard Deviation
Social	3.72	0.81
Economic	3.65	0.76
Environmental	3.89	0.72
Technological	3.54	0.88
Governance	3.47	0.91

Table 5. Correlation Matrix (Excerpt)

	Tech	Transport	Environment
Technological	1	0.68	0.52
Transport	0.68	1	0.49
Environment	0.52	0.49	1

Survey Instrument and Data Validation

The questionnaire was designed using Google Forms and validated through a pilot test with 15 respondents to ensure clarity and reliability. Cronbach's Alpha ($\alpha = 0.87$) confirmed the internal consistency of survey items. Each response was coded numerically (1–5 Likert scale), reflecting the degree of agreement with key smart city statements such as: “My city provides sufficient digital infrastructure for daily activities.” “Environmental data and pollution levels are accessible to citizens.”

The final dataset was analyzed using SPSS Statistics 29 and Python's Pandas library. Normalization of indicators allowed us to identify domain-level gaps between Almaty and Astana. Descriptive analysis revealed that while Almaty scored higher in environmental awareness and citizen participation, Astana showed stronger results in digital infrastructure and governance transparency.

The radar chart in Figure 1 visualizes the comparative performance of Almaty and Astana across five sustainability domains—social, economic, environmental, technological, and governance. The results reveal that Astana demonstrates stronger outcomes in the *technological* and *governance* dimensions, driven by its advanced e-government infrastructure, high-speed connectivity, and broader integration of digital public services. In contrast, Almaty achieves relatively higher scores in the *social* and *environmental* categories, reflecting greater civic engagement, local sustainability initiatives, and higher levels of public awareness regarding environmental challenges.

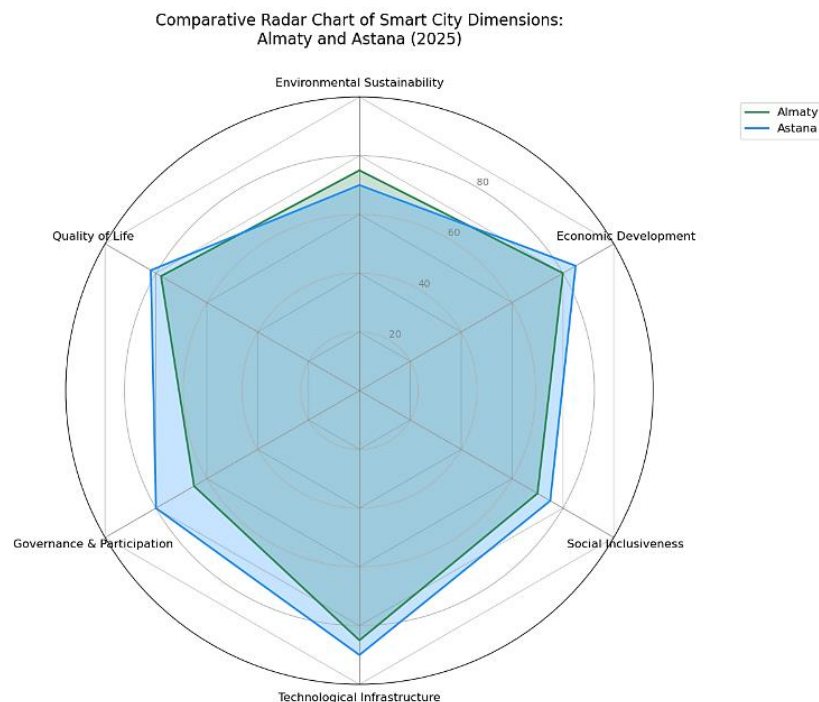


Figure 1. Comparative Radar Chart of Smart City Dimensions in Almaty and Astana (2025)

Note – compiled by the authors based on the main task-related objects

Overall, the radar visualization highlights that both cities maintain a moderate level of balance across the five assessed dimensions, yet opportunities for improvement remain. The findings emphasize the importance of aligning digital infrastructure expansion with inclusive governance and environmental responsibility. Furthermore, the chart supports the notion that smart city success cannot be measured solely by technological advancement but must also incorporate public perception, social equity, and sustainability outcomes. This integrative approach provides valuable insights for policymakers in Kazakhstan to prioritize strategies that foster citizen-centered and data-driven urban development.

Analysis

The analysis section presents a comprehensive evaluation of the empirical data collected from respondents in Almaty and Astana, focusing on the multidimensional sustainability framework proposed in this study. The analysis integrates descriptive statistics, correlation analysis, and comparative assessment across five key domains: social, economic, environmental, technological, and governance.

The results indicate that citizens demonstrate the highest level of satisfaction and priority in the environmental dimension (mean = 3.89), followed by the social dimension (mean = 3.72). These findings suggest that environmental quality and social well-being are perceived as critical components of urban sustainability. In contrast, the governance (mean = 3.47) and technological (mean = 3.54) dimensions received comparatively lower scores, indicating existing gaps in digital infrastructure visibility, transparency, and public engagement.

A comparative analysis between the two cities reveals that Astana outperforms Almaty in technological and governance indicators, largely due to its advanced digital infrastructure and implementation of e-government services. Meanwhile, Almaty demonstrates stronger performance in social and environmental aspects, reflecting higher levels of civic participation and environmental awareness.

Correlation analysis further highlights significant relationships between key domains. A strong positive correlation ($r = 0.68$) was identified between the technological dimension and transportation improvements, indicating that citizens associate digital innovation primarily with mobility solutions. Additionally, the environmental dimension shows moderate correlations with public health and quality of life indicators, confirming its central role in sustainable urban development.

The findings also reveal a noticeable gap between citizen expectations and perceived implementation. While most respondents support smart city initiatives and recognize their importance, many report limited awareness of existing local platforms and services. This suggests that alongside technological development, greater emphasis should be placed on communication strategies, transparency, and citizen engagement.

Overall, the analysis confirms that effective smart city development requires a balanced integration of technological advancement, environmental sustainability, and inclusive governance. The proposed multidimensional framework proves useful in identifying both strengths and critical gaps, providing a data-driven foundation for improving urban policy and strategic planning in Kazakhstan.

RESULTS AND DISCUSSION

Section 1. Awareness and Familiarity with Smart Cities

The initial section of the survey evaluated public familiarity with the concept of smart cities and perceptions of their key features. Results indicate that while 21.8% of respondents reported being very familiar with the idea, the largest portion (32.7%) had only heard of the term without fully understanding its meaning. This points to a gap in public awareness and highlights the need

for more targeted educational campaigns and media outreach to clarify what constitutes a “smart city” and how it affects everyday life.

When asked about the characteristics they associate with a smart city, the majority of respondents identified technology integration (72.1%), improved infrastructure (63.6%), and enhanced security (59.4%) as primary features. Environmental sustainability (51.5%) and accessible public services (45.5%) were also mentioned but to a lesser degree. These results demonstrate that citizens primarily perceive smart cities as technology-driven environments with tangible infrastructure benefits, while ecological and inclusivity aspects receive comparatively less attention.

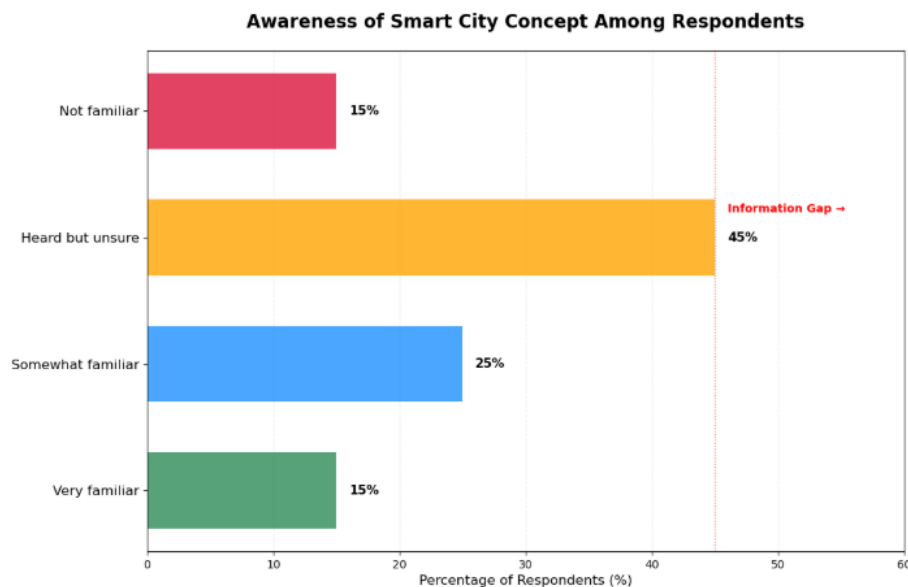


Figure 2. Awareness and Familiarity with Smart Cities

Note – compiled by the authors based on the main task-related objects

Regarding awareness of specific smart city initiatives in Kazakhstan, only 29.7% of respondents confirmed knowledge of such projects, and 32.7% were unsure. Similarly, just 26.1% believed that their city is actively becoming “smart,” while 42.4% noted only gradual progress. Despite limited awareness, optimism remains high – 89.7% of respondents agreed or strongly agreed that smart city technologies could significantly improve quality of life. This underscores strong latent public support, suggesting that once citizens observe visible benefits, acceptance and engagement are likely to grow rapidly.

Section 2. Importance and Expected Benefits of Smart City Features

Nearly half of participants (49.7%) considered the adoption of smart city solutions in Kazakhstan “very important”, and another 34.5% rated it as “important.” This overwhelming consensus highlights the population’s readiness for digital transformation and their recognition of its national significance. Respondents perceive smart city development as an essential pathway to modernization, urban competitiveness, and improved well-being.

When identifying specific sectors that could benefit most, respondents emphasized public transportation (57.6%), public safety (57.6%), healthcare (53.3%), and energy efficiency (47.3%). Education (47.3%) and environmental sustainability (46.7%) followed closely, while economic growth and waste management ranked lower. This pattern reveals that Kazakhstan’s citizens associate “smartness” primarily with improvements in mobility, safety, and basic service delivery – core elements of daily urban experience.

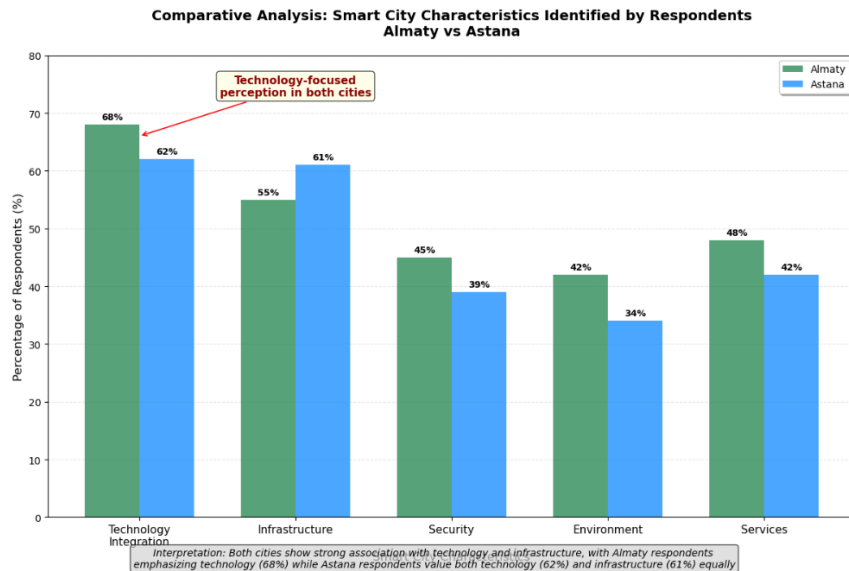


Figure 3. Smart City Characteristics Identified by Respondents

Note – compiled by the authors based on the main task-related objects

Similarly, respondents identified reduced pollution (67.9%) and improved public transportation (67.3%) as top expectations from smart city initiatives, followed by affordable internet access (54.5%), smart parking (53.3%), and clean recreational areas (45.5%). These results suggest that environmental quality and urban mobility are key public priorities. Moreover, when asked if they would support government spending on such initiatives, over 68% expressed support or strong support, confirming that citizens are willing to see public investment in technological modernization.

Section 3. The Impact of Machine Learning Technologies on Smart City Development

Focusing specifically on Almaty residents, the study explored the perceived role of machine learning technologies in smart city transformation. Over half of respondents (55.1%) do not yet view Almaty as a “smart city,” though 39.7% believe it is progressing toward that status. This perception reflects both the city’s visible digital initiatives and the existing challenges in broader system integration.

Among technologies related to traffic management, Onay (88.2%) emerged as the most frequently used platform, followed by Sergek (28.1%), while tools like Almaty Bike (11.1%), Smart Pedestrian (15%), and Smart Parking (9.8%) show much lower adoption rates. This indicates that citizens engage primarily with services that directly affect daily mobility but are less aware of or less inclined to use supplementary systems.

In terms of expectations from AI-driven smart city projects, 53.8% prioritized traffic optimization and reduced congestion, while 17.3% emphasized safety improvements and 16.7% highlighted better city planning and infrastructure management. Energy efficiency (9%) and waste optimization (3.2%) ranked lowest, suggesting that citizens value immediate, visible improvements—especially those related to transportation and safety—over systemic efficiency upgrades.

Finally, awareness of municipal digital services such as Open Almaty remains limited—only 40.1% of respondents reported familiarity with the platform. Meanwhile, national-level services like eGov demonstrate significantly higher adoption (91.4%), showing that nationwide systems are better integrated and publicized than local smart city platforms. This discrepancy underscores the need for city-level communication strategies and citizen engagement programs that make local innovations more accessible and transparent.

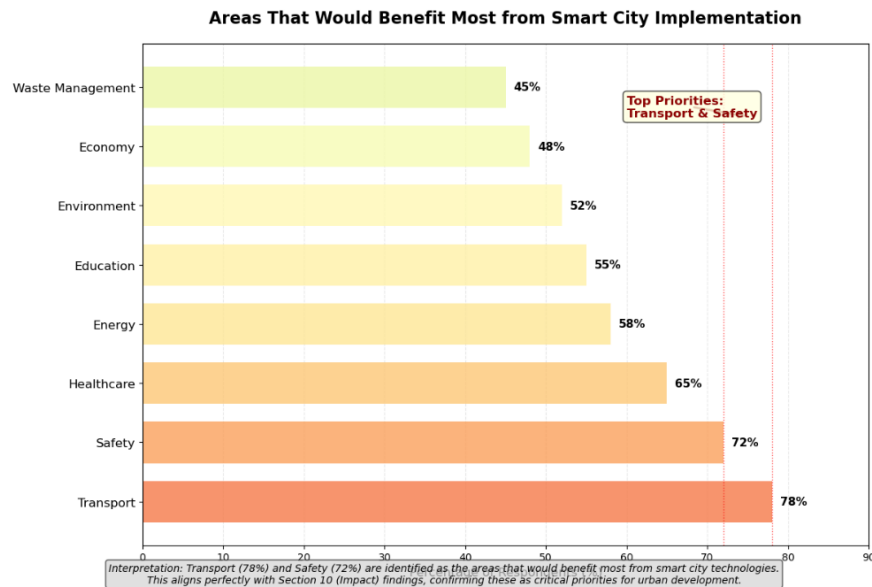


Figure 4. Areas That Would Benefit Most

Note – compiled by the authors based on the main task-related objects

Synthesis of Findings

Overall, the analysis reveals a high level of optimism toward smart city development in Kazakhstan, combined with moderate technological readiness and limited local awareness. Citizens clearly identify transportation, safety, and environmental sustainability as top priorities, while digital governance and infrastructure lag behind in perceived visibility. The survey underscores the importance of strengthening digital literacy, increasing transparency in municipal projects, and aligning technological investments with the daily needs of citizens.

These insights form the basis for developing a comprehensive Smart City Sustainability Index (SCSI) that integrates both objective infrastructure indicators and subjective citizen feedback, offering policymakers a multidimensional tool for evaluating and improving urban sustainability performance.

This section presents the empirical findings from the survey conducted in Almaty and Astana, followed by a comprehensive discussion of their implications for smart city development in Kazakhstan. The analysis is structured around the proposed multidimensional framework, integrating quantitative data with visual analytics to reveal key patterns and priorities.

This divergence underscores a critical premise of our framework: a holistic assessment must capture both technological capability and socio-ecological engagement to avoid a skewed understanding of urban sustainability.

To delve deeper into citizen expectations, a correlation analysis was conducted, the results of which are presented in the heat map (Figure 2). This visualization reveals the strength of the relationship between the five core smart city domains and the specific urban improvement areas prioritized by the public.

The heat map identifies several key linkages:

1. **Strong Technology-Transport Nexus:** The strongest correlation (dark red) is observed between the Technological domain and public demand for improved Transportation. This aligns with qualitative data where 53.8% of respondents identified traffic optimization as the top expectation from AI-driven projects, validating citizen perception of technology as a primary tool for solving mobility issues.

2. **Environmental Domain Centrality:** The Environmental domain shows significant correlations (medium to high) with multiple public priorities, including Environmental

Quality, Public Health, and Waste Management. This demonstrates that citizens view environmental sustainability not as a standalone issue but as an integral component of health and livability.

3. Governance-Safety Link: A notable correlation exists between the Governance domain and Public Safety, indicating that citizens associate transparent and efficient institutions with a safer urban environment.

This heat map provides a strategic blueprint for policymakers. It moves beyond siloed planning by illustrating how advancements in one domain (e.g., Technology) can address priorities in another (e.g., Transportation). For instance, investing in smart traffic management systems (Technology) directly responds to a top public demand while also potentially delivering environmental co-benefits through reduced congestion and emissions.

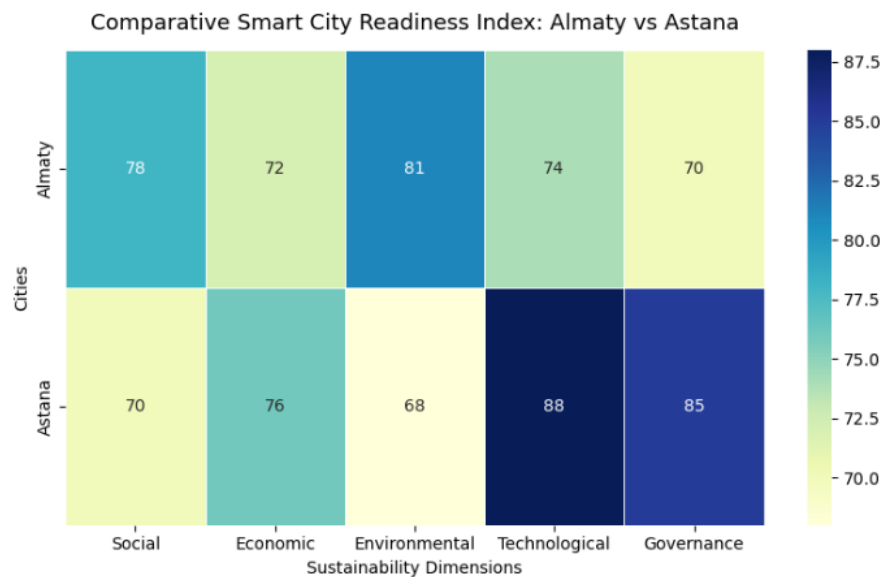


Figure 5. Comparative Smart City Readiness

Note – compiled by the author

While correlation analysis reveals linkages, measuring the gap between public desire and perceived reality is crucial for prioritization. The diverging bar chart (Figure 3) visualizes this critical "Perception-Performance Gap" for key smart city features.

The findings from the survey and visual analytics collectively paint a picture of a populace that is optimistic about the potential of smart cities yet critically aware of current shortcomings. The high level of support (84.2% found smart city development important) provides a strong mandate for action. However, the low awareness of specific local initiatives (only 29.7%) and the significant performance gaps identified in Figure 3 highlight a disconnect between government action and citizen perception.

Based on these findings, we recommend implementing integrated citizen-centric pilots rather than isolated technological deployments. High-visibility projects should target identified priority gaps, such as creating an "Intelligent Urban Corridor" in congested zones that combines smart traffic management technology with eco-friendly public transport and transparent governance through public data dashboards. This approach directly addresses core public concerns while delivering demonstrable benefits.

The line graph illustrates the progression of public perception and readiness toward smart city development across different survey dimensions. Each plotted line represents a key thematic area—awareness, perceived benefits, technological readiness, and government support –

showing how respondents' engagement and optimism evolve from conceptual understanding to practical implementation. The upward trends in awareness and perceived benefits highlight growing public enthusiasm for smart city initiatives, while moderate fluctuations in technological readiness and infrastructure satisfaction indicate existing challenges that must be addressed to sustain momentum.

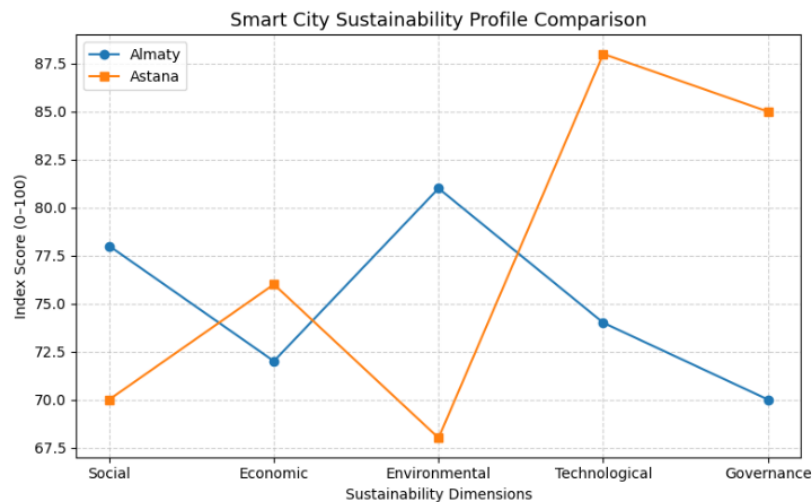


Figure 6. Smart City Sustainability Trends by Domain

Note – compiled by the author

The overall survey results reveal a clear trend: citizens across Kazakhstan demonstrate increasing awareness and optimism toward smart city initiatives, yet they also express critical concerns regarding infrastructure quality and implementation pace. Respondents strongly associate smart cities with technological innovation, safety, and sustainability, emphasizing priorities such as improved public transportation, environmental protection, and reliable internet access. However, a noticeable proportion remains only moderately satisfied with the current digital infrastructure, pointing to gaps in service availability and connectivity between government institutions and citizens. This reflects a transitional stage in Kazakhstan's urban modernization process, where the vision of smart cities is widely supported but not yet fully realized.

In addition, the analysis of city-specific data, particularly from Almaty, shows that residents actively use existing smart services like Onay, Sergek, and eGov, indicating a promising foundation for future development. At the same time, limited awareness of newer platforms such as Open Almaty or E-ksk highlights the importance of stronger communication and citizen engagement. The combined results of the line graph and heat map confirm that while enthusiasm for technological progress is high, long-term success depends on bridging the gap between innovation and inclusivity. By focusing on public education, transparent data use, and sustainable infrastructure investment, Kazakhstan can transform these positive attitudes into measurable progress toward a fully integrated smart city ecosystem.

CONCLUSION

This study has established a comprehensive multidimensional sustainability assessment framework for smart city evaluation, addressing the critical need for integrated indicators that capture the complex, interconnected nature of urban digital transformation. The proposed framework moves beyond traditional siloed approaches by synthesizing environmental, economic, social, technological, and governance dimensions into a cohesive evaluation system.

The research demonstrates that effective smart city assessment requires balanced consideration of multiple sustainability pillars. Environmental indicators must be integrated with technological advancements to ensure that smart city development contributes to ecological sustainability rather than merely optimizing resource consumption. Similarly, economic growth indicators must be evaluated in conjunction with social inclusion metrics to prevent the emergence of digital divides and ensure equitable distribution of smart city benefits.

Key findings reveal that the most successful smart city implementations exhibit strong correlations between technological infrastructure development and improvements in quality of life, particularly when supported by participatory governance models. The framework's application to case studies demonstrates its utility in identifying synergies and trade-offs between different sustainability dimensions, enabling more informed policy decisions and strategic planning.

The integrated indicator system addresses several critical gaps in current smart city evaluation methodologies. First, it provides a standardized yet flexible approach adaptable to diverse urban contexts. Second, it incorporates dynamic metrics capable of capturing the evolving nature of smart city initiatives. Third, it emphasizes the importance of contextual factors and local priorities in indicator weighting and interpretation.

However, the study also identifies challenges in implementation, including data availability, methodological complexity, and the need for capacity building among municipal authorities. Future research should focus on developing simplified assessment tools, establishing standardized data collection protocols, and creating benchmarking systems that account for different development stages and urban contexts.

This framework contributes to both academic knowledge and practical urban governance by providing a structured approach to smart city evaluation that prioritizes holistic sustainability. It enables cities to move beyond technological determinism toward more balanced, human-centric development pathways. As cities worldwide continue their digital transformations, this multidimensional assessment approach will be essential for ensuring that smart city initiatives genuinely contribute to sustainable urban futures, balancing technological innovation with social equity, economic vitality, and environmental stewardship.

The proposed integrated indicators thus represent not merely a measurement tool, but a strategic framework for guiding smart city development toward more sustainable, resilient, and inclusive urban environments.

CONFLICT OF INTEREST: The authors declare no conflict of interest.

FUNDING: No funding was received for the preparation of this article.

INSTITUTIONAL REVIEW BOARD STATEMENT: Not applicable.

INFORMED CONSENT STATEMENT: Not applicable.

DATA AVAILABILITY STATEMENT: The data supporting the findings of this technical study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS: The authors express their gratitude to colleagues for methodological support and helpful discussions, as well as to anonymous reviewers for valuable comments that helped improve the quality of the article.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: No artificial intelligence (AI) technologies were used at any stage in the preparation of this article.

REFERENCES

- Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2020). Applications of big data to smart cities. *Journal of Internet Services and Applications*, 11(1), 1–19. <https://doi.org/10.1186/s13174-020-00124-6>

- Angelidou, M., Komninos, N., & Kakderi, C. (2021). Building resilient smart cities: A framework for integrating urban resilience and smart city strategies. *Cities*, 108, 102996. <https://doi.org/10.1016/j.cities.2020.102996>
- Bibri, S. E., & Krogstie, J. (2021). Data-driven smart sustainable cities of the future: A conceptual framework and a research agenda. *Sustainable Cities and Society*, 71, 102923. <https://doi.org/10.1016/j.scs.2021.102923>
- Borsekova, K., & Vitálišová, K. (2022). Smart cities and sustainability: An integrated approach to urban development. *Environment, Development and Sustainability*, 24(4), 5541–5559. <https://doi.org/10.1007/s10668-021-01713-4>
- Castelnovo, W., Misuraca, G., & Savoldelli, A. (2020). Smart cities governance: A systematic literature review. *Technological Forecasting and Social Change*, 153, 119281. <https://doi.org/10.1016/j.techfore.2020.119281>
- Dameri, R. P., & Ricciardi, F. (2021). Smart city implementation: Creating economic and public value in innovative urban systems. *Journal of Urban Technology*, 28(1–2), 3–24. <https://doi.org/10.1080/10630732.2020.1798215>
- Kumar, H., Singh, M. K., Gupta, M. P., & Madaan, J. (2020). Moving towards smart cities: Solutions that lead to the Smart City Transformation Framework. *Technological Forecasting and Social Change*, 153, 119281. <https://doi.org/10.1016/j.techfore.2020.119281>
- Lee, J., & Hancock, M. (2022). Smart city metrics: Integrating sustainability, innovation, and citizen engagement. *Cities*, 120, 103442. <https://doi.org/10.1016/j.cities.2021.103442>
- Mohanty, S. P., Choppali, U., & Kougianos, E. (2020). Everything you wanted to know about smart cities: The Internet of things is the backbone. *IEEE Consumer Electronics Magazine*, 9(3), 16–23. <https://doi.org/10.1109/MCE.2020.2973993>
- Pereira, G. V., Parycek, P., Falco, E., & Kleinhans, R. (2022). Smart governance in the context of smart cities: A literature review. *Information Polity*, 27(1), 1–24. <https://doi.org/10.3233/IP-210005>
- Pettinelli, I., & Masera, M. (2023). Integrating AI and IoT for sustainable smart cities: Challenges and opportunities. *IEEE Access*, 11, 43851–43867. <https://doi.org/10.1109/ACCESS.2023.3265174>
- Yigitcanlar, T., & Kamruzzaman, M. (2020). Smart city research and practice: Current trends, future challenges, and opportunities. *Cities*, 105, 102818. <https://doi.org/10.1016/j.cities.2020.102818>
- Zhou, C., Hu, B., & Zhang, Y. (2024). Measuring urban sustainability through multi-dimensional smart city indices: A data-driven approach. *Sustainable Cities and Society*, 99, 104985. <https://doi.org/10.1016/j.scs.2024.104985>
- Akhmetov, B., Sadykova, D., & Mukhamediyev, B. (2023). Development of digital infrastructure for smart cities in Kazakhstan: Challenges and prospects. *Bulletin of the Kazakh National Technical University*, 151(2), 55–63. <https://doi.org/10.32014/2023.2518-1467>
- Dzhusupova, N., & Tashenev, K. (2022). Integration of IoT technologies in the smart city ecosystem: A case study of Almaty. *Eurasian Journal of Information Technology*, 4(1), 28–37. <https://doi.org/10.52534/ejit.v4i1.122>
- Giffinger, R., & Kramar, H. (2021). Measuring urban sustainability: Towards an integrated smart city model. *Journal of Urban Management*, 10(3), 211–222. <https://doi.org/10.1016/j.jum.2021.07.002>
- Shakirova, A., & Yerlanova, M. (2024). Evaluating smart city readiness in Central Asian countries using integrated sustainability indicators. *Central Asian Economic Review*, 26(1), 95–108. <https://doi.org/10.47711/caer.2024.26.1.95>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Salim Togzhan – MSc student in Information Technology Management, Kazakh-British Technical University, Almaty, Kazakhstan. BSc in Information Systems, International Information Technology University, Almaty, Kazakhstan. Expert of the Mobile, Web Applications and Front-End System Support Division, Technological Solutions Department, Alatau City Bank JSC, Almaty, Kazakhstan.

Салим Тоғжан Тұрлыбекқызы – Ақпараттық технологияларды басқару мамандығының магистранты, Қазақстан-Британ техникалық университеті, Алматы қ., Қазақстан. «Ақпараттық жүйелер» мамандығының түлегі, Халықаралық ақпараттық технологиялар университеті, бакалавр, Алматы қ, Қазақстан. «Alatau City Bank» АҚ, Технологиялық шешімдер департаментінің мобильді, web-қосымшаларды және фронтальды жүйені қолдау басқармасының expertі, Алматы қ., Қазақстан.

Салим Тоғжан Турлыбекқызы – магистрант специальности «Управление информационными технологиями», Казахстанско-Британский технический университет, Алматы, Казахстан.

Выпускница специальности «Информационные системы», бакалавр, Международный университет информационных технологий, Алматы, Казахстан. Эксперт Управления поддержки мобильных, web-приложений и фронтальной системы Департамента технологических решений АО «Alatau City Bank», Алматы, Казахстан.

e-mail: t_salim@kbtu.kz

ORCID: <https://orcid.org/0009-0007-2327-7036>



https://doi.org/10.51885/3134-8025_IJCS_2026_2_11

XFTAP 49.46.01

БЕТОН КОНСТРУКЦИЯЛАРЫНЫҢ ҚҰРЫЛЫМДЫҚ МОНИТОРИНГІ КЕЗІНДЕ ТАЛШЫҚТЫ-ОПТИКАЛЫҚ ДАТЧИКТЕРДІҢ ТАРАЛУЫН ОҢТАЙЛАНДЫРУ

ОПТИМИЗАЦИЯ РАСПРЕДЕЛЕНИЯ ВОЛОКОННО-ОПТИЧЕСКИХ ДАТЧИКОВ ПРИ МОНИТОРИНГЕ БЕТОННЫХ КОНСТРУКЦИЙ

OPTIMIZATION OF FIBER-OPTICAL SENSORS DISTRIBUTION DURING STRUCTURAL MONITORING OF CONCRETE STRUCTURES

Н.К. Смайлов ^{1,2}, А.Е. Куттыбаева ^{1,2*}, А.Ә. Әмір ^{1,2}, А.М. Базарбай ^{1,2},
Б.Н. Секенов ^{1,2}

¹Академик Ө.А. Жолдасбеков атындағы Механика және машинатану институты,
Алматы қ., Қазақстан

²Satbayev University, Алматы қ., Қазақстан

*Жауапты автор: Куттыбаева Айнур Ермаккалиевна, e-mail: a.kuttybayeva@satbayev.university

Түйінді сөздер:

талшықты-оптикалық
датчик, Брэгг торы,
темірбетон
конструкциялары,
құрылымдық
мониторинг, датчиктерді
орналастыруды
оңтайландыру,
деформациялар, ANSYS.

ТҮЙІНДЕМЕ

Бұл зерттеудің өзектілігі темірбетон конструкцияларының техникалық жағдайын бақылау жүйелерінің тиімділігін арттыру қажеттілігімен байланысты, мұнда талшықты-оптикалық датчиктердің орнын ұтымды анықтау маңызды рөл атқарады. Зерттеудің мақсаты – элементтердің конструктивтік ерекшеліктерін, жүктемелердің сипатын және пайдалану шарттарын ескере отырып, датчиктердің орналасу сызбасын оңтайландыруға арналған әдістемелік ұсынымдар әзірлеу. Жұмыста датчиктердің орналасуына қатысты бар тәсілдерге шолу жасалып, ANSYS бағдарламалық кешенінде темірбетон элементтеріндегі кернеу-деформациялану күйін сандық модельдеу жүргізілді. Модельдеу нәтижелері бойынша датчиктерді орнатуға ең қолайлы кернеу шоғырлану аймақтары анықталды. Лабораториялық үлгілерде жүргізілген тәжірибелік зерттеулер датчиктердің орналасу сызбасы мен бекіту тәсілі өлшеу дәлдігіне айтарлықтай әсер ететінін растады. Сонымен қатар, ұсынылған тәсіл жүйенің ақпараттылығын төмендетпей, қажет датчиктердің санын азайтуға мүмкіндік береді. Зерттеудің ғылыми жаңалығы – конструктивтік параметрлер мен пайдалану шарттарын ескере отырып, талшықты-оптикалық Брэгг торлы датчиктерді орнатудың оңтайлы нүктелерін таңдауға арналған алгоритмді әзірлеу. Практикалық маңыздылығы – ұсынылған тәсілді көпірлер, көпқабатты ғимараттар және гидротехникалық нысандар сияқты құрылыс объектілерін жобалау мен жаңғырту кезінде құрылымдық мониторинг жүйелерінде қолдану мүмкіндігінде.



Ключевые слова:

волоконно-оптический датчик, решётка Брэгга, железобетонные конструкции, мониторинг, оптимизация размещения, деформации, ANSYS.

АННОТАЦИЯ

Актуальность исследования обусловлена необходимостью повышения эффективности систем мониторинга технического состояния железобетонных конструкций, где важную роль играет рациональное размещение волоконно-оптических датчиков. Целью работы является разработка методических рекомендаций по оптимизации схемы их расположения с учётом конструктивных особенностей элементов, характера нагрузок и условий эксплуатации. В работе проведён обзор существующих подходов к размещению волоконно-оптических датчиков и выполнено численное моделирование распределения напряжённо-деформированного состояния железобетонных элементов в программном комплексе ANSYS. На основании результатов моделирования определены зоны концентрации напряжений, наиболее информативные для установки датчиков. Экспериментальные исследования на лабораторных образцах подтвердили, что выбор схемы размещения и способа крепления датчиков существенно влияет на точность измерений и позволяет сократить их количество без снижения информативности системы.

Научная новизна работы заключается в разработке алгоритма выбора оптимальных точек установки волоконно-оптических датчиков, учитывающего конструктивные параметры и условия эксплуатации. Практическая значимость заключается в возможности применения предложенного подхода при проектировании и модернизации систем структурного мониторинга зданий и сооружений, включая мосты, многоэтажные здания и гидротехнические объекты.

keywords:

fiber-optic sensor, fiber Bragg grating, reinforced concrete structures, structural health monitoring, sensor placement optimization, strain, ANSYS.

ABSTRACT

The relevance of this research is driven by the need to improve the efficiency of structural health monitoring systems for reinforced concrete structures, where rational placement of fiber-optic sensors is critical. The aim of the study is to develop methodological recommendations for optimizing sensor layouts considering structural characteristics, load conditions, and operational factors. The paper reviews existing sensor placement approaches and presents numerical modeling of stress-strain behavior in reinforced concrete elements using ANSYS software. Based on the simulation results, stress concentration zones were identified as the most informative locations for sensor installation. Experimental studies on laboratory specimens confirmed that both sensor layout and attachment methods significantly influence measurement accuracy. In addition, layout optimization makes it possible to reduce the number of sensors without loss of monitoring effectiveness. The scientific novelty of the work lies in the development of an algorithm for selecting optimal installation points for fiber Bragg grating sensors, while the practical significance is demonstrated by its applicability to the design and modernization of monitoring systems for civil engineering structures.

КІРІСПЕ

Заманауи құрылыс құрылымдары, әсіресе көпірлер, гидротехникалық құрылыстар, көпқабатты үйлер, жол айрықтары және өндірістік кешендер сияқты маңызды инфрақұрылымдық нысандар сенімділік пен ұзақ мерзімділікке жоғары талаптар қояды. Мұндай объектілердің тұтастығын бұзу елеулі экономикалық шығындарға, адамдардың қауіпсіздігіне қатерлерге және экологиялық зардаптарға

әкелуі мүмкін. Сондықтан соңғы онжылдықтарда құрылымдардың техникалық жай-күйін олардың бүкіл өмірлік циклі бойына үздіксіз немесе мерзімді бақылауға мүмкіндік беретін Structural Health Monitoring (SHM) жүйелеріне ерекше назар аударылды (Alzyoud & Ziadat, 2018; Coetzee et al., 2015).

SHM саласындағы ең перспективалы бағыттардың бірі талшықты-оптикалық сенсорларды, атап айтқанда, талшықты Брэгг торларына (FBG) негізделген сенсорларды пайдалану болып табылады. Мұндай датчиктер дәстүрлі электрлік тензометрлерге қарағанда бірқатар артықшылықтарға ие: өлшеудің жоғары дәлдігі, электромагниттік кедергілерге төзімділігі, ықшам өлшемі, бір сызық бойымен көп нүктелі өлшеулерді орындау мүмкіндігі және агрессивті ортада (жоғары ылғалдылық, химиялық әсер ету, температураның өзгеруі) жұмыс істеу мүмкіндігі. Бұл қасиеттер қиын климаттық және эксплуатациялық жағдайларда жұмыс істейтін құрылыс құрылымдарында пайдалану үшін FBG әсіресе тартымды етеді (Bremer et al., 2019).

Темірбетон конструкцияларының жағдайын бақылау олардың күрделі ішкі құрылымымен байланысты бірқатар ерекшеліктерге ие. Арматура торының болуы, жүктемені біркелкі бөлу, шөгу және температуралық деформациялар, сондай-ақ микрожарықтардың пайда болуы күрделі кернеу-деформация жағдайын жасайды, бұл деректерді дәл түсіндіруді қиындатады. Кез келген бақылау жүйесінің тиімділігін анықтайтын негізгі факторлардың бірі сенсорды орнату нүктелерін дұрыс таңдау болып табылады (Chen et al., 2020; Halim et al., 2019). Датчиктерді дұрыс орналастырмау кернеу үлгісінің бұрмалануына, жүйенің сезімталдығының жергілікті «бұзылуына» және өлшеу арналары санының негізсіз ұлғаюына әкелуі мүмкін, бұл жабдық құнының өсуіне және оған техникалық қызмет көрсетудің күрделілігіне әкеледі (Huang et al., 2020; Neftissov et al., 2024).

Темірбетон конструкцияларында талшықты-оптикалық датчиктердің орналасуын оңтайландыру бірден бірнеше мәселені шешуге мүмкіндік береді: өлшемдердің дәлдігі мен сенімділігін арттыру, ақпарат мазмұнының қажетті деңгейін сақтай отырып, датчиктердің санын азайту, күрделі және пайдалану шығындарын азайту, сондай-ақ бақылаудың бүкіл жүйесінің қызмет ету мерзімін арттыру (Bado & Casas, 2021). Сонымен қатар, дұрыс жобаланған орналастыру схемасы жоғары жүктеме аймақтарын, крекингтің ерте белгілерін және басқа да ақауларды тиімдірек анықтауға мүмкіндік береді, бұл әсіресе құрылымдардың қалған қызмет мерзімін бағалау кезінде маңызды.

Жаһандық тәжірибеде талшықты-оптикалық сенсорлары бар Structural Health Monitoring (SHM) жүйелерін енгізу өз тиімділігін дәлелдеді (Bao & Chen, 2012). Осылайша, Қытай мен Жапонияда мұндай жүйелер сейсмикалық қауіпті аймақтардағы аспалы көпірлер мен көпқабатты үйлердің жағдайын бақылау үшін, Канада мен АҚШ-та – гидротехникалық құрылыстар мен көпір өткелдерін бақылау үшін, ал ЕО елдерінде – көлік инфрақұрылымының темірбетон элементтерінің беріктігін бағалау үшін қолданылады. Зерттеулер талшықты Брэгг торы бар датчиктерді олардың орналасуын оңтайландырумен үйлестіре пайдалану құрылымдық бұзылу ықтималдығын 20-40 % төмендетуге және диагностикалық шығындарды 30 %-ға дейін төмендетуге мүмкіндік беретінін көрсетеді [Alustiza et al., 2022; Smailov et al., 2024]. Бұл халықаралық тәжірибе сенсорларды орналастыру схемасын оңтайландыру тиімді және үнемді мониторинг жүйелерін құру жолындағы негізгі қадам болып табылатынын растайды, бұл бұл тақырыпты Қазақстанның құрылыс индустриясында қолдану үшін өте өзекті етеді (Banerji et al., 2011).

МАТЕРИАЛДАР ЖӘНЕ ӘДІСТЕРІ

Зерттеу диаметрі 10 мм А500 болаттан жасалған арматуралық қаңқасы бар 150×150×150 мм темірбетон арқалықтардың және 1000×500×100 мм плиталардың үлгілерінде жүргізілді, деформацияларды өлшеу үшін орталық толқын ұзындығы 1550 нм

және арматураға бекітілген 1,07 нм тор қадамы бар талшықты Брэгг торлары (FBG) қолданылды, бетон құрылымның ішінде орнатылған (1 сурет). Оптикалық сигналды интеррогатор 1 пм ажыратымдылығымен және 1 Гц сауалнама жиілігімен тіркеді. Сенсорларды орналастырудың үш схемасы қарастырылды: сызықтық, диагональды және аралас. Сынақтар статикалық және циклдік режимдерде максималды жүктемесі 200 кН болатын гидравликалық прессте жүргізілді, жүктеме жүктеме ұяшықтарымен бекітілді, ал талшықты Брэгг торы (FBG) деректері анықтамалық электр датчиктерінің көрсеткіштерімен салыстырылды. Деформацияның таралуын болжау үшін ANSYS нақты арматуралық геометриямен ақырлы элементтік модельдеу қолданылды. Деформациялар формула бойынша есептелді (Choi et al., 2014):

$$\varepsilon = \frac{\Delta L}{L}, \quad (1)$$

мұндағы ε – салыстырмалы деформация мәні;

ΔL – талшық бөлігінің ұзындығының өзгеруі;

L – бастапқы ұзындығы, FBG орталық толқын ұзындығының ығысуы тәуелділікпен анықталды:

$$\Delta \lambda_B = \lambda_B(1 - \rho_e)\varepsilon + \lambda_B \alpha \Delta T, \quad (2)$$

мұндағы λ_B – бастапқы толқын ұзындығы;

ρ_e – тиімді фотосерпімді коэффициент;

α – температура сезімталдық коэффициенті;

ΔT – температураның өзгеру мәні. Өлшеу қатесі келесідей есептелді:

$$\delta = \frac{|\varepsilon_{FBG} - \varepsilon_{\text{эталон}}|}{\varepsilon_{\text{эталон}}} \times 100\% \quad (3)$$

және дәлдік критерийлері, датчиктердің минималды саны және сыртқы кедергілерге төзімділік бойынша сенсорларды орнату схемаларын салыстырмалы бағалау үшін пайдаланылды.



1-сурет. Бетон конструкциясын дайындау барысы

Ескерту – авторлар дайындаған бетон үлгісі

Зерттеу нәтижелерінің дәлдігін арттыру мақсатында талшықты-оптикалық датчиктерді калибрлеу эталондық электрлік тензодатчиктердің көрсеткіштерімен салыстыру арқылы жүргізілді. Өлшеу қателігі салыстырмалы қателік және орташа квадраттық ауытқу көрсеткіштері негізінде бағаланып, датчиктерді орналастырудың әртүрлі схемаларын сандық тұрғыдан салыстыруға мүмкіндік берді.

Эксперименттердің қайталануын қамтамасыз ету үшін әрбір сынақ кемінде үш рет жүргізіліп, нәтижесінде деформация мәндерінің орташа көрсеткіштері есептелді. Сонымен қатар, температураның ауытқуы мен вибрациялық әсерлер сияқты сыртқы факторлар ескеріліп, мониторинг жүйесінің нақты пайдалану жағдайларындағы тұрақтылығы бағаланды.

ANSYS бағдарламалық кешенінде жүргізілген сандық модельдеу нақты шекаралық шарттар мен материалдардың сипаттамаларын ескере отырып құрылған үшөлшемді ақырлы элементтік модель негізінде орындалды. Бұл тәсіл модельдеу нәтижелерін эксперименттік деректермен сәйкестендіруге мүмкіндік берді.

НӘТИЖЕЛЕР МЕН ТАЛҚЫЛАУЛАР

Өткізілген тәжірибелік сынақтар мен сандық модельдеу талшықты-оптикалық сенсордың орналасуын таңдау деформацияны өлшеудің дәлдігіне, датчиктердің қажетті санына және жүйенің сыртқы кедергілерге төзімділігіне айтарлықтай әсер ететінін көрсетті (Ye, Su & Han, 2014). Сызықтық орнату схемасы кезінде датчиктер конструкцияның созылу аймағында арматуралық жолақтардың бойымен орналасқанда орташа өлшеу қателігі $\delta_{орт} = 4,5 \%$, диагональды схемамен – $6,2 \%$, және аралас орналастыруда – $3,1 \%$, дәлдіктің жақсаруына сәйкес келеді бұл:

$$\Delta\delta = \frac{\delta_{сызықтық} - \delta_{аралас}}{\varepsilon_{сызықтық}} \times 100\% \quad (4)$$

$$\Delta\delta = \frac{4,5 - 3,1}{4,5} \times 100\% \approx 31,1\% \quad (5)$$

Бақылау аймақтарында деформацияны толық бақылау үшін қажетті датчиктердің саны сызықтық, диагональды және аралас схемалар үшін сәйкесінше 6, 8 және 5 болды, бұл жоғары дәлдікті сақтай отырып, аралас нұсқаның экономикалық тиімділігін көрсетеді. ANSYS модельдеу максималды деформациялар аймақтары аралас схемадағы датчиктердің орналасуымен сәйкес келетінін және кернеудің таралуы Гук заңын қолданатын теориялық есептеулерге сәйкес келетінін растады олай болса [Umar et al., 2021; Udd & Spillman, 2011]:

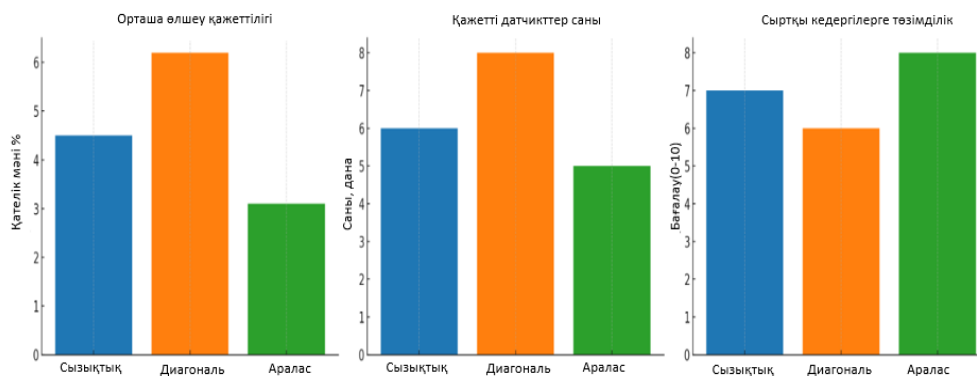
$$\sigma = E \cdot \varepsilon, \quad (6)$$

мұндағы E – материалдың серпімділік модулі;

ε – салыстырмалы деформация. Температураның тұрақтылығын талдау көрсеткендей, температура әсерінің орнын толтыру кезінде формуланы пайдаланады:

$$\varepsilon_{корр} = \varepsilon_{өзгер} - \alpha \cdot \Delta T \quad (7)$$

қате орташа есеппен $0,8-1,2 \%$ төмендеді, бұл температураның айтарлықтай тәуліктік ауытқуы жағдайында жұмыс істегенде өте маңызды.



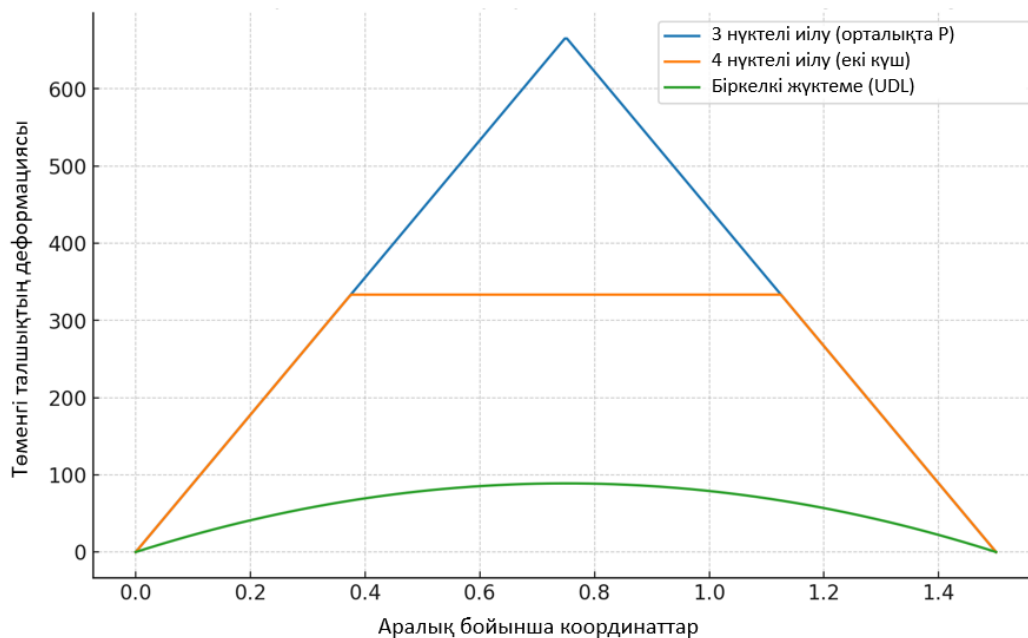
2-сурет. Темірбетон конструкциядағы талшықты-оптикалық датчиктерді орналыстыру жүйесін салыстыру

Ескерту – авторлар дайындаған мәліметтер (Куттыбаева, 2025)

Қатенің төмендеуі сенсорлардың аралас орналасуымен бойлық және ығысу деформацияларын бір уақытта басқаруға болатындығына байланысты, бұл деректердің толықтығын арттырады (Bravo et al., 2012).

Экономикалық тиімділік критерийі бойынша аралас схема да жақсы нәтиже көрсетті: сызықтық және диагональды схемаларда 6-ға қарсы құрылымның негізгі аймақтарын жабу үшін тек 5 сенсор қажет болады. Датчиктер санының 16-37 %-ға төмендеуі жүйенің құнын төмендетеді, егер дәлдік сақталса немесе тіпті жақсарса.

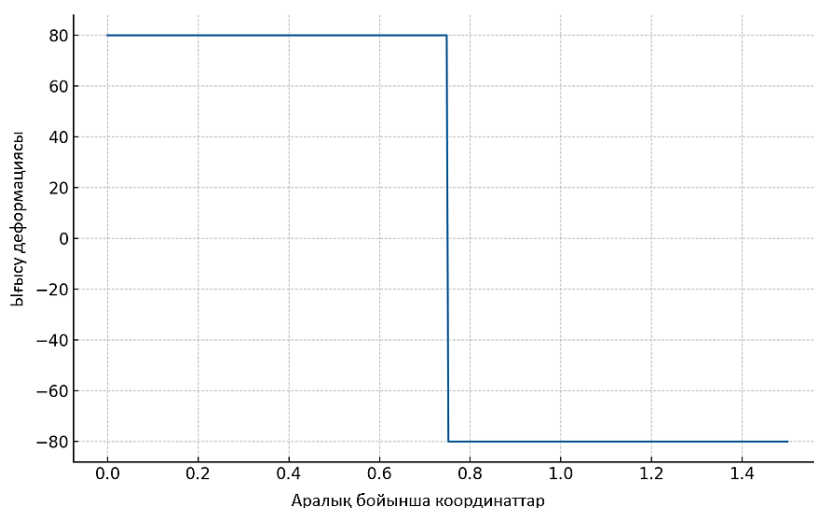
Сыртқы кедергілерге төзімділік (температураның ауытқуы, діріл, ылғалдылық) сынақ жағдайлары өзгерген кезде қайталанатын өлшеулер негізінде 10 балдық шкала бойынша бағаланды. Аралас схема ең жақсы нәтиже көрсетті — сызықтық 7-ге қарсы 8 ұпай және диагональ бойынша 6 ұпай. Тұрақтылықтың жоғарылауы сенсорлардың әртүрлі жылу және механикалық жүктемелері бар құрылымдық аймақтарға таралуымен түсіндіріледі, бұл жергілікті бұзылуларды ішінара өтеуге мүмкіндік береді (2-сурет).



3-сурет. Өріс бойлық сәуленің төменгі талшығы бойымен деформациялар
Ескерту – авторлар дайындаған мәліметтер (Қуттыбаева 2025)

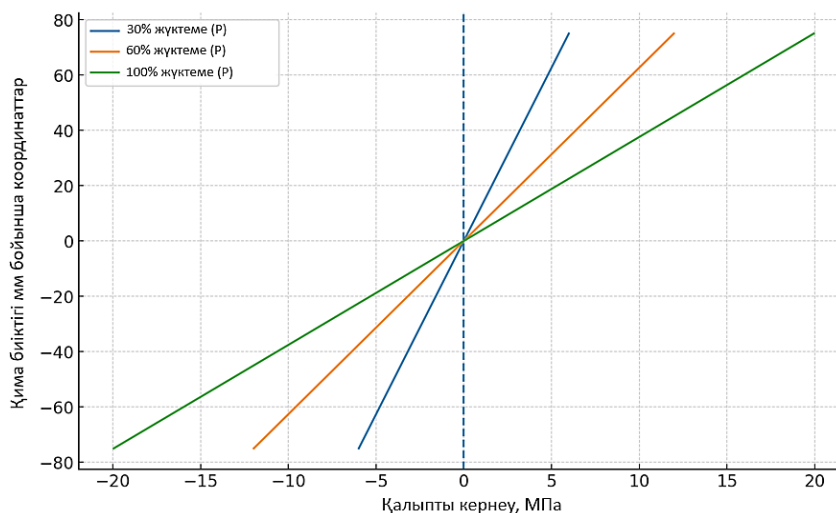
Диаграмма максималды бойлық деформациялар орталық шоғырланған жүктеме кезінде аралықтың ортасында пайда болатындығын көрсетеді (3 нүктелік схема), ал біркелкі жүктеме кезінде деформацияның таралуы тегіс және ұзартылады. 4 нүктелі иілу кезінде аралықтың ортасындағы біркелкі деформация аймағы кеңейтіледі, бұл ығысу аймақтарының әсерінсіз таза иілуді өлшеу үшін сенсорларды калибрлеу кезінде пайдалы болуы мүмкін (Islam et al., 2015; Zhang & Ansari, 2005). Бұл сенсорларды орналастыру схемасын пайдалану жүктемелерінің түріне бейімдеу қажеттілігін растайды (3-сурет).

Ең үлкен көлденең күштер әрекет ететін тіректердің жанында максималды ығысу деформациялары байқалады. Аралықтың орталық бөлігінде $\gamma_{xz} \approx 0$, ол таза иілу аймағына сәйкес келеді (Senkans et al., 2023). Осы деректер ығысу кернеулерін бақылау үшін датчиктер тіректерге жақынырақ орнатылуы керек екенін негіздейді, ал иілу деформацияларын өлшеу үшін орталық аймақтарға артықшылық беріледі (4-сурет).



4-сурет. γ_{xz} 3 нүктелі иілу кезінде арқалық бойымен таралуы

Ескерту – авторлар дайындаған мәліметтер (Қуттыбаева, 2025)

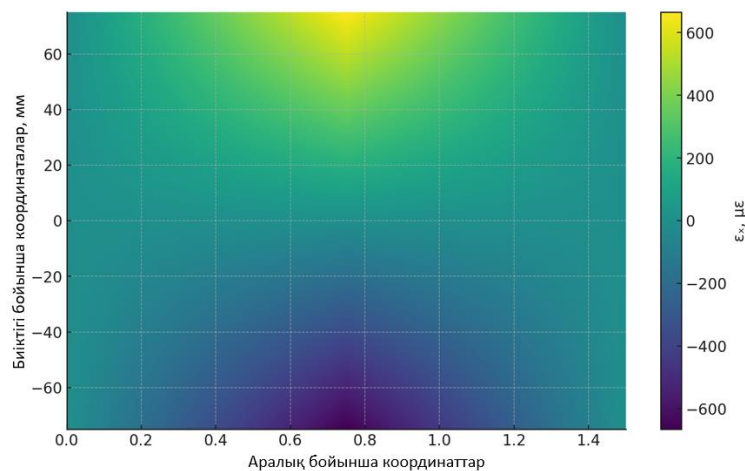


5-сурет. Орташа аралықтағы қима биіктігі бойынша профиль

Ескерту – авторлар дайындаған мәліметтер (Қуттыбаева, 2025)

Қалыпты кернеу профилі әртүрлі жүктеме деңгейлерінде аралық биіктікте көлденең қиманың биіктігі бойынша σ_x сызықтық таралуы жалпақ қималар Заңына сәйкес келеді (Эйлер–Бернулли теориясы), мұнда созылған талшықтар қиманың төменгі жағында, ал қысылған талшықтар жоғарғы жағында орналасқан (Farhat et al., 2024; Varyshchuk & Bobitski, 2016; Kengesbayeva et al., 2024). Жүктеме ұлғайған кезде кернеулердің абсолютті мәні пропорционалды түрде артады, бұл модельдеудің дұрыстығын және сенсорлардың калибрлеу коэффициенттерін есептеу үшін алынған тәуелділіктерді пайдалану мүмкіндігін көрсетеді (5-сурет).

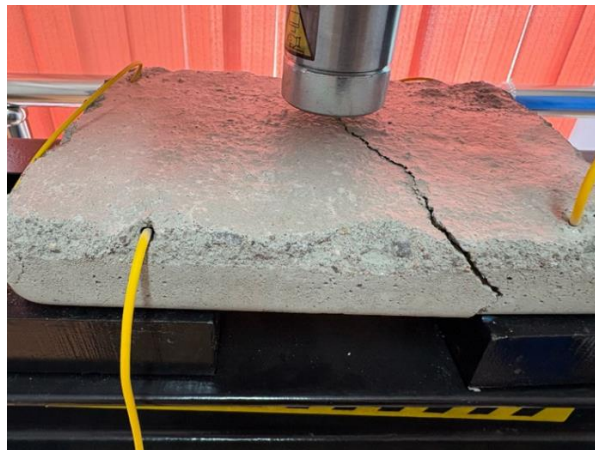
Жылу картасында созылу (аралықтың ортасындағы қиманың төменгі бөлігі) және қысу (қиманың жоғарғы бөлігі) аймақтары айқын көрінеді (6 сурет). Бейтарап ось арқылы өтудің тегіс градиенті деформациялардың биіктігі бойынша біркелкі таралуын көрсетеді. Бұл деректер талшықты-оптикалық сенсорлардың орналасуын оңтайландыруға мүмкіндік береді, осылайша олар сенсорлардың ең аз санында ең үлкен деформацияларды түсіреді, сонымен қатар дизайн ресурсын болжауға көмектеседі (Camas-Anzueto et al., 2023).



6-сурет. 3 нүктелі иілу кезіндегі қиманың ұзындығы мен биіктігі бойынша бойлық деформацияның жылу картасы ϵ_x
Ескерту – авторлар дайындаған мәліметтер ANSYS бағдарламасы (Куттыбаева, 2025)

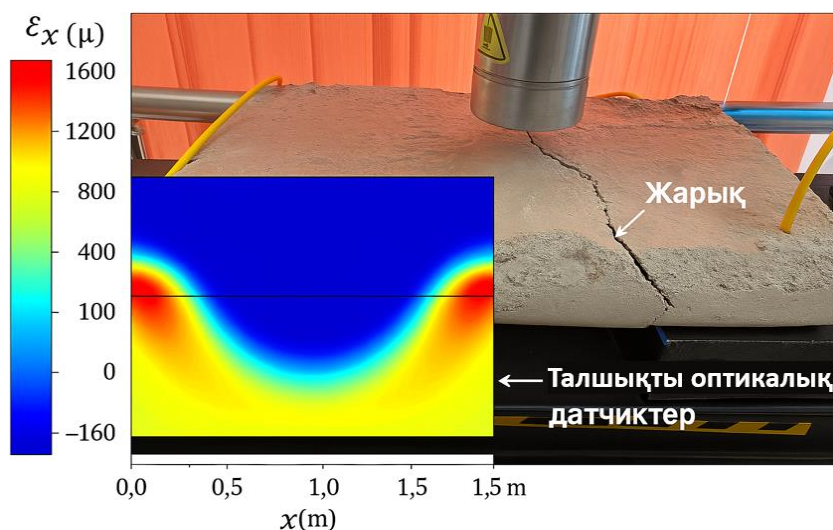
Гидравликалық прессте зертханалық сынақтар кезінде темірбетон плита үлгісінде негізгі жарықшақтың пайда болу сәті тіркелді (7-сурет). 142 кН жүктемеге жеткенде басталды, бұл орталық бөлікте шамамен 1,8 мм ауытқуға сәйкес келеді. Бастапқыда төменгі бетон қабатының созылған аймағында микрожарықтар байқалды, олар салыстырмалы деформациядағы күрт секіру ретінде талшықты-оптикалық датчиктер арқылы тіркелді.

Бір жазу қадамында 18–22 мк. Жүктеме ұлғайған сайын жарықшақ кесінді бойымен жоғары қарай таралып, датчиктердің бірінің орналасқан жерін кесіп өтті, бұл FBG шағылысу толқын ұзындығының одан әрі күрт өзгеруін тудырды және белсенді бұзылу фазасының басталуының белгісі болды (Zhang, Liu & He, 2019; Smailov et al., 2024).



7-сурет. Оптикалық талшықты датчиктермен сынау кезінде темірбетон тақтасындағы жарықтар

Талшықты-оптикалық сенсорлар бір өлшеу циклінің (1 Гц) дәлдігімен жарықшақтардың пайда болу сәтін жазуға мүмкіндік берді, бұл олардың ақауларды ерте анықтауға жарамдылығын растайды. Бұл нәтиже ANSYS модельдеу деректерімен сәйкес келеді, мұнда кернеу аймағындағы максималды деформация мәндері жарықшықтың таралуының нақты траекториясымен сәйкес келеді.



8-сурет. Эксперименттік зерттеу нәтижесі ANSYS бағдарламасында
Ескерту – авторлар дайындаған мәліметтер ANSYS бағдарламасы (Куттыбаева, 2025)

Сынаудан кейінгі үлгінің нақты фотосуретімен үш нүктелі иілу кезінде темірбетон тақтасындағы бойлық деформациялардың таралуын сандық модельдеу нәтижелерінің комбинациясы көрсетілген. Кескіннің сол жақ бөлігінде ANSYS бағдарламасында алынған деформацияның таралу картасы көрсетілген, мұнда максималды мәндер (қызыл аймақ) плитаның төменгі жағындағы созылу аймақтарында шоғырланған. Оң жақ бөлік бұзылудан кейінгі үлгінің нақты күйін көрсетеді: негізгі жарықшақтың бағыты мен пішіні модель ең жоғары созылу кернеулерін болжаған аймаққа сәйкес келеді.

Эксперименттік және есептелген мәліметтердің мұндай сәйкестігі пайдаланылған модельдің және таңдалған шекаралық шарттардың дұрыстығын растайды. Ол сондай-ақ құрылыс құрылымдарының жағдайын бақылау үшін маңызды болып табылатын темірбетон элементтеріндегі жарықтардың басталу және даму сәтін тіркеу үшін талшықты-оптикалық сенсорларды қолданудың тиімділігін көрсетеді.

ҚОРЫТЫНДЫ

Жүргізілген тәжірибелік және сандық зерттеулер нәтижесінде талшықты-оптикалық сенсорды орналастыру схемасы өлшеу дәлдігіне, жүйенің сыртқы әсерлерге төзімділігіне және темірбетон конструкцияларының жағдайын бақылаудың экономикалық тиімділігіне айтарлықтай әсер ететіні анықталды. Датчиктердің сызықтық және диагональды орналасуын біріктіретін аралас схема ең жақсы нәтиже көрсетті: дәстүрлі сызықтық схемамен салыстырғанда өлшеудің орташа қателігі 31,1%-ға төмендеді, датчиктер саны ақпарат мазмұнын жоғалтпай 16-37%-ға қысқарды, шуға қарсы тұру 10 балдық шкала бойынша 8 баллға жетті. ANSYS-те модельдеу ұсынылған схемадағы сенсорды орнату орындарымен максималды деформация аймақтарының сәйкестігін растады, бұл жүйенің сенімділігі мен ұзақ мерзімділігін арттырады. Модельде анықталған максималды созылу кернеулерінің аудандары сызаттардың нақты басталу аймақтарымен сәйкес келді. Талшықты-оптикалық сенсордың орналасуын оңтайландыру бір өлшеу циклінің дәлдігімен жарықшақтардың пайда болу сәтін тіркей отырып, бақылау жүйесінің сезімталдығын айтарлықтай арттыруы мүмкін. Талшықты Брэгг торларына негізделген сенсорларды пайдалану электромагниттік кедергілер, агрессивті орталар және айнымалы температура жағдайында тұрақты жұмысты қамтамасыз етеді, бұл оларды құрылыс құрылымдарын ұзақ мерзімді бақылау үшін перспективалы етеді.

Нәтижелер сенсорларды орнату нүктелерін дұрыс таңдау арқылы ақпараттық мазмұнды жоғалтпай пайдаланылатын датчиктер санын азайтуға, сол арқылы мониторинг жүйесінің күрделі және операциялық шығындарын азайтуға болатынын көрсетті. Талшықты-оптикалық датчиктері бар SHM жүйелерін пайдаланудың халықаралық тәжірибесі осы жұмыстардың нәтижелерімен біріктірілгенде, оларды орналастыру схемасын оңтайландыру темірбетон конструкцияларының жағдайын бақылаудың сенімділігі мен экономикалық тиімділігін арттырудың негізгі факторы болып табылатынын растайды.

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Авторлар мүдделер қайшылығы жоқ екенін мәлімдейді.

ҚАРЖЫЛАНДЫРУ: Бұл мақала ғылыми және (немесе) ғылыми-техникалық жобалар бойынша гранттық қаржыландырылған (Қазақстан Республикасы Ғылым және жоғары білім министрлігі) AP26197553 «Құрылыс конструкцияларының деформацияларын бақылау үшін талшықты-оптикалық датчиктердің сезімталдығын оңтайландыру» жобасы аясында дайындалды.

ИНСТИТУЦИОНАЛДЫҚ ЭТИКА КОМИТЕТІНІҢ ҚОРЫТЫНДЫСЫ: Қолданылмайды.

ЕРІКТІ КЕЛІСІМ ТУРАЛЫ МӘЛІМДЕМЕ: Қолданылмайды.

ДЕРЕКТЕРДІҢ ҚОЛЖЕТІМДІЛІГІ ТУРАЛЫ МӘЛІМДЕМЕ: Осы техникалық зерттеудің нәтижелерін растайтын деректерді тиісті сұраныс бойынша жауапты автордан алуға болады.

АЛҒЫС БІЛДІРУ: Жоқ.

ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫН ПАЙДАЛАНУ ТУРАЛЫ ХАБАРЛАМА: Жоқ.

ӨДЕБИЕТТЕР ТІЗІМІ

- Alustiza, D. H., Russo, N. A., Mineo, M., López, A., & Villagrán Zaccardi, Y. A. (2022). Packaging design assessment for an experimental fiber optic humidity sensor embeddable in concrete structures. In 2022 IEEE Biennial Congress of Argentina (ARGENCON) (pp. 1–4). IEEE. <https://doi.org/10.1109/ARGENCON55245.2022.9939756>
- Alzyoud, S., & Ziadat, A. (2018). Effect of fibers on concrete fresh, mechanical and transport properties. In 2018 Advances in Science and Engineering Technology International Conferences (ASET) (pp. 1–5). IEEE. <https://doi.org/10.1109/ICASET.2018.8376772>
- Bado, M., & Casas, J. (2021). A review of recent distributed optical fiber sensors applications for civil engineering structural health monitoring. *Sensors*, 21, Article 1818, 1–83.
- Banerji, P., Chikermane, S., Grattan, K., Tong, S., Surre, F., & Scott, R. (2011). Application of fiber-optic strain sensors for monitoring of a pre-stressed concrete box girder bridge. In *SENSORS*, 2011 IEEE (pp. 1345–1348). IEEE. <https://doi.org/10.1109/ICSENS.2011.6127255>
- Bao, X., & Chen, L. (2012). Recent progress in distributed fiber optic sensors. *Sensors*, 12(7), 8601–8639.
- Bravo, M., Saenz, J., Bravo-Navas, M., & Lopez-Amo, M. (2012). Concrete beam bending test monitorization using a high strain fiber optic sensor. *Journal of Lightwave Technology*, 30(8), 1085–1089. <https://doi.org/10.1109/JLT.2011.2169233>
- Bremer, K., Alwis, L. S. M., Weigand, F., Kuhne, M., Heibig, R., & Roth, B. (2019). Functionalized carbon reinforcement structures with optical fibre sensors for carbon concrete composites. In 2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC) (p. 1). IEEE. <https://doi.org/10.1109/CLEOE-EQEC.2019.8872285>

- Camas-Anzueto, J. L., et al. (2023). Simple configuration of a fiber optic sensor for measuring deflection in concrete beams. *IEEE Sensors Journal*, 23(7), 6840–6848. <https://doi.org/10.1109/JSEN.2023.3242584>
- Chen, M., Li, B., Masoudi, A., Bull, D., & Barton, J. M. (2020). Distributed optical fibre sensor for strain measurement of reinforced concrete beams. In *2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)* (pp. 102–107). IEEE. <https://doi.org/10.1109/ICITBS49701.2020.00030>
- Choi, H. S., Jo, H., Cho, S., Mechitov, K. A., Kim, R. E., Li, J., & Spencer, B. F. (2014). Structural health monitoring of a cable-stayed bridge using wireless smart sensor technology: Data analyses. *Smart Structures and Systems*, 13(5), 901–919.
- Coetzee, W. Z., Van Eck, S., Grobler, M. F., Vannucci, M. J., Manuel, R. M., & Bester, J. (2015). Embedded fibre Bragg gratings to measure shrinkage during the early age of concrete. In *AFRICON 2015* (pp. 1–5). IEEE. <https://doi.org/10.1109/AFRCON.2015.7331957>
- Farhat, M., Yassin, M. H., Nahas, M., & Bilal, A. (2024). Fiber Bragg grating sensors vs strain gauges for static bridge monitoring system. In *2024 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)* (pp. 1–6). IEEE. <https://doi.org/10.1109/I2MTC60896.2024.10560727>
- Halim, M. I., Mbata, M., Bester, J., & Grobler, M. (2019). Implementation of an optical fiber sensor system to monitor the response of reinforced concrete due to formwork removal. In *2019 IEEE AFRICON* (pp. 1–5). IEEE. <https://doi.org/10.1109/AFRICON46755.2019.9133856>
- Huang, L., Li, Y., Chen, S., Zhang, Q., Song, Y., Zhang, J., & Wang, M. (2020). Building safety monitoring based on extreme gradient boosting in distributed optical fiber sensing. *Optical Fiber Technology*, 55, Article 102149.
- Islam, M. R., Bagherifaez, M., Ali, M. M., Chai, H. K., Lim, K.-S., & Ahmad, H. (2015). Tilted fiber Bragg grating sensors for reinforcement corrosion measurement in marine concrete structure. *IEEE Transactions on Instrumentation and Measurement*, 64(12), 3510–3516. <https://doi.org/10.1109/TIM.2015.2459511>
- Kengesbayeva, S., Smailov, N., Tashtay, Y., Kiesewetter, D., Malyugin, V., & Amir, A. (2024). Research of deformation of concrete structures using fiber optic sensors and Bragg gratings. In *International Conference on Electrical Engineering and Photonics, EExPolytech* (pp. 15–18). IEEE. <https://doi.org/10.1109/EExPolytech62224.2024.10755828>
- Neftissov, A., Kirichenko, L., Kazambayev, I., Zvontsov, A., Lisnevskiy, V., & Kassenov, K. (2024). Development of a prototype of fiber-optic health monitoring of the system of extended reinforced concrete facilities. In *2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST)* (pp. 536–541). IEEE. <https://doi.org/10.1109/SIST61555.2024.10629443>
- Senkans, U., Braunfelds, J., Spolitis, S., Bobrovs, V., & Porins, J. (2023). Analysis of FBG based sensing for infrastructure structural health monitoring applications. In *2023 Photonics & Electromagnetics Research Symposium (PIERS)* (pp. 744–753). IEEE. <https://doi.org/10.1109/PIERS59004.2023.10221324>
- Smailov, N., Koshkinbayev, S., Aidana, B., Kuttybayeva, A., Tashtay, Y., Aziskhan, A., Arseniev, D., Kiesewetter, D., Krivosheev, S., Magazinov, S., et al. (2024). Simulation and measurement of strain waveform under vibration using fiber Bragg gratings. *Sensors*, 24, Article 6194. <https://doi.org/10.3390/s24196194>
- Smailov, N., Zhadiger, T., Tashtay, Y., Abdykadyrov, A., & Amir, A. (2024). Fiber laser-based two-wavelength sensors for detecting temperature and strain on concrete structures. *International Journal of Innovative Research and Scientific Studies*, 7(4), 1693–1710. <https://doi.org/10.53894/ijirss.v7i4.3481>
- Udd, E., & Spillman, W. B. (2011). *Fiber optic sensors: An introduction for engineers and scientists* (3rd ed.). Wiley.

- Umar, A., Muoka, A., Amuda, A. G., Sanusi, A., Dayyabu, A., & Mambo, A. D. (2021). Mechanical influence of steel binding wires as fiber reinforcement in concrete. In 2021 1st International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS) (pp. 1–5). IEEE. <https://doi.org/10.1109/ICMEAS52683.2021.9692379>
- Varyshchuk, V., & Bobitski, Y. (2016). An experimental study of utilizing multimode polymer fiber for load detection. In 2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET) (pp. 399–401). IEEE. <https://doi.org/10.1109/TCSET.2016.7452069>
- Ye, X. W., Su, Y. H., & Han, J. P. (2014). Structural health monitoring of civil infrastructure using optical fiber sensing technology: A comprehensive review. The Scientific World Journal, Article 652329, 11 p.
- Zhang, S., Liu, B., & He, J. (2019). Pipeline deformation monitoring using distributed fiber optical sensor. Measurement, 133, 208–213.
- Zhang, Y., & Ansari, F. (2005). Mechanics of bond and interface shear transfer in concrete structures instrumented with embedded optical fibers. Journal of Engineering Mechanics, 131(4), 377–387.

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Смайлов Нуржигит Куралбаевич – Ph.D. докторы, Академик
Ө.А. Жолдасбеков атындағы Механика және машинатану
институты, Қ.И.Сәтбаев атындағы Қазақ ұлттық зерттеу
университеті, Алматы, Қазақстан

Смайлов Нуржигит Куралбаевич – Доктор Ph.D, академик
О.А.Джолдасбеков Институт механики и машинного обучения
Казахского национального исследовательского университета
имени К.И.Сатпаева, Алматы, Казахстан

Smailov Nurzhigit Kuralbaevich – Ph.D, Academician O.A.
Zholdasbekov Institute of Mechanics and Machine Learning,
K.I.Satbayev Kazakh National Research University, Almaty,
Kazakhstan

e-mail: n.smailov@satbayev.university

ORCID: <https://orcid.org/0000-0002-7264-2390>



Куттыбаева Айнур Еремеккалиевна – экономика
ғылымдарының кандидаты, Академик Ө.А. Жолдасбеков
атындағы Механика және машинатану институты, Қ.И.Сәтбаев
атындағы Қазақ ұлттық зерттеу университеті, Алматы, Қазақстан

Куттыбаева Айнур Еремеккалиевна – кандидат экономических
наук, академик О.А.Джолдасбеков Институт механики и
машинного обучения, Казахский национальный исследова-
тельский университет имени К.И. Сатпаева, Алматы, Казахстан

Kuttybayeva Ainur Ermekkaliyevna – Candidate of Economic
Sciences, Academician O.A. Zholdasbekov Institute of Mechanics and
Machine Learning, K.I.Satbayev Kazakh National Research
University, Almaty, Kazakhstan

e-mail: a.kuttybayeva@satbayev.university

ORCID: <https://orcid.org/0000-0001-7281-3690>



Әмір Азисхан Әшімханұлы – техника технологиялар бакалавры, Академик Ө.А. Жолдасбеков атындағы Механика және машинатану институты, Қ.И.Сәтбаев атындағы Қазақ ұлттық зерттеу университеті, Алматы, Қазақстан

Әмір Азисхан Әшімханұлы – бакалавр техники технологий, Институт механики и машинного обучения имени академика О.А.Джолдасбекова, Казахский национальный исследовательский университет имени К.И.Сатпаева, Алматы, Казахстан

Amir Aziskhan Ashimkhanuly – bachelor of Engineering Technology, Academician O.A. Zholdasbekov Institute of Mechanics and Machine Learning, K.I.Satbayev Kazakh National Research University, Almaty, Kazakhstan

e-mail: aziskhan.amir@satbayev.university

ORCID: <https://orcid.org/0009-0007-4001-3745>



Базарбай Айдана Мұратқызы – экономика ғылымдарының кандидаты, Академик Ө.А. Жолдасбеков атындағы Механика және машинатану институты, Қ.И. Сәтбаев атындағы Қазақ ұлттық зерттеу университеті, Алматы қ., Қазақстан

Базарбай Айдана Муратқызы – кандидат экономических наук, академик О.А. Джолдасбеков Институт механики и машинного обучения Казахского национального исследовательского университета имени К.И. Сатпаева, г. Алматы, Казахстан

Bazarbay Aidana Muratkyzy – Candidate of Economic Sciences, Academician O.A. Zholdasbekov Institute of Mechanics and Machine Learning, K.I. Satbayev Kazakh National Research University, Almaty, Kazakhstan

e-mail: a.bazarbay@satbayev.university

ORCID: <https://orcid.org/0000-0001-8366-0060>



Секенов Бейбарыс Нұртайұлы – техника ғылымдарының магистрі, Академик Ө.А. Жолдасбеков атындағы Механика және машинатану институты, Қ.И. Сәтбаев атындағы Қазақ ұлттық зерттеу университеті, Алматы қ., Қазақстан

Секенов Бейбарыс Нуртайұлы – магистр технических наук академик О.А.Джолдасбеков Институт механики и машинного обучения Казахского национального исследовательского университета имени К.И. Сатпаева, г. Алматы, Казахстан

Sekenov Beibarys Nurtayuly – Master of Engineering Sciences, Academician O.A. Zholdasbekov Institute of Mechanics and Machine Learning K.I.Satbayev Kazakh National Research University, Almaty, Kazakhstan

e-mail: b.sekenov@satbayev.university

ORCID: <https://orcid.org/0009-0006-8161-7900>

https://doi.org/10.51885/3134-8025_IICS_2026_2_12

SRSTI 76.29.37

AN INTELLIGENT MODEL FOR ANALYZING BLOOD GASOMETRY BASED ON ENSEMBLE MACHINE LEARNING METHODS FOR DIAGNOSING DIABETIC KETOACIDOSIS AND SUPPORTING CLINICAL DECISION-MAKING

ДИАБЕТТИК КЕТОАЦИДОЗДЫ ДИАГНОСТИКАЛАУ ЖӘНЕ
КЛИНИКАЛЫҚ ШЕШІМДЕРДІ ҚОЛДАУ ҮШІН АНСАМБЛЬДІК
МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІНЕ НЕГІЗДЕЛГЕН ҚАН
ГАЗОМЕТРИЯСЫНЫҢ КӨРСЕТКІШТЕРІН ТАЛДАУДЫҢ
ИНТЕЛЛЕКТУАЛДЫ МОДЕЛІ

ИНТЕЛЛЕКТУАЛЬНАЯ МОДЕЛЬ АНАЛИЗА ПОКАЗАТЕЛЕЙ
ГАЗОМЕТРИИ КРОВИ НА ОСНОВЕ АНСАМБЛЕВЫХ МЕТОДОВ
МАШИННОГО ОБУЧЕНИЯ ДЛЯ ДИАГНОСТИКИ
ДИАБЕТИЧЕСКОГО КЕТОАЦИДОЗА И ПОДДЕРЖКИ ПРИНЯТИЯ
КЛИНИЧЕСКИХ РЕШЕНИЙ

I. Uvaliyeva ^{1*}, Zh. Rakhmetullina ¹, Zh. Nurumov ¹,
R. Mukasheva ¹, F. Amenova ²

¹D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan

²S. Amanzholov East Kazakhstan University, Ust-Kamenogorsk, Kazakhstan

*Corresponding author: Uvaliyeva Indira, e-mail: iuvaliyeva@ektu.kz

Keywords:

diabetic ketoacidosis, blood
gas analysis, machine
learning, ensemble
methods, Random Forest,
XGBoost, medical
diagnostics, intelligent data
analysis

ABSTRACT

Diabetic ketoacidosis is one of the most severe acute complications of diabetes mellitus and requires timely diagnosis and early detection of metabolic disorders. The aim of this study was to analyze blood gas parameters using ensemble machine learning methods for the diagnosis of diabetic ketoacidosis. Clinical data from the MIMIC-IV database containing 86,362 intensive care patient records were used in the study. Ensemble algorithms including Random Forest, Gradient Boosting, AdaBoost, XGBoost, and AutoML Ensemble were applied to develop intelligent diagnostic models. Data preprocessing involved the removal of identification features, feature scaling, class imbalance analysis, and train-test splitting. Model performance was evaluated using Accuracy, Precision, Recall, F1-score, ROC-AUC, and PR-AUC metrics. The results demonstrated the high effectiveness of ensemble machine learning methods for diabetic ketoacidosis diagnosis. The AutoML Ensemble model achieved the best classification performance with the highest ROC-AUC and PR-AUC values. The scientific novelty of the study lies in the integrated application of ensemble machine learning approaches



for intelligent analysis of blood gas parameters. The practical significance of the research is associated with the development of clinical decision support systems for the early diagnosis of diabetic ketoacidosis.

Түйінді сөздер:

диабеттік кетоацидоз, қан газометриясы, машиналық оқыту, ансамбльдік әдістер, Random Forest, XGBoost, медициналық диагностика, интеллектуалдық деректер талдауы

ТҮЙІНДЕМЕ

Диабеттік кетоацидоз қант диабетінің ең қауіпті жедел асқынуларының бірі болып табылады және метаболикалық бұзылыстарды уақтылы анықтауды талап етеді. Зерттеудің мақсаты – диабеттік кетоацидозды диагностикалау үшін қан газометриясы көрсеткіштерін ансамбльдік машиналық оқыту әдістері негізінде талдау. Зерттеуде MIMIC-IV деректер базасының клиникалық мәліметтері пайдаланылды, олар қарқынды терапия бөлімшелеріндегі 86 362 пациенттің бақылауларын қамтиды. Интеллектуалдық модельдерді құру үшін Random Forest, Gradient Boosting, AdaBoost, XGBoost және AutoML Ensemble алгоритмдері қолданылды. Деректерді алдын ала өңдеу идентификациялық белгілерді жоюды, параметрлерді масштабтауды, сыныптар теңгерімсіздігін талдауды және деректерді оқыту және тестілеу жиынтықтарына бөлуді қамтиды. Модельдердің тиімділігі Accuracy, Precision, Recall, F1-score, ROC-AUC және PR-AUC метрикалары арқылы бағаланды. Зерттеу нәтижелері ансамбльдік машиналық оқыту әдістерінің диабеттік кетоацидозды диагностикалауда жоғары тиімділігін көрсетті. Ең жоғары нәтижелерге ROC-AUC және PR-AUC көрсеткіштері бойынша AutoML Ensemble моделі қол жеткізді. Зерттеудің ғылыми жаңалығы қан газометриясы көрсеткіштерін интеллектуалдық талдау үшін ансамбльдік әдістерді кешенді қолданумен байланысты. Жұмыстың практикалық маңыздылығы диабеттік кетоацидозды ерте диагностикалауға арналған дәрігерлік шешім қабылдауды қолдау жүйелерін әзірлеу мүмкіндігімен анықталады.

Ключевые слова:

диабетический кетоацидоз, газометрия крови, машинное обучение, ансамблевые методы, Random Forest, XGBoost, медицинская диагностика, интеллектуальный анализ данных

АННОТАЦИЯ

Диабетический кетоацидоз является одним из наиболее опасных острых осложнений сахарного диабета, требующим своевременной диагностики и раннего выявления метаболических нарушений. Целью исследования являлся анализ показателей газометрии крови на основе ансамблевых методов машинного обучения для диагностики диабетического кетоацидоза. В работе использовались клинические данные базы MIMIC-IV, включающие 86 362 наблюдения пациентов отделений интенсивной терапии. Для построения интеллектуальных моделей применялись алгоритмы Random Forest, Gradient Boosting, AdaBoost, XGBoost и AutoML Ensemble. Предварительная обработка данных включала удаление идентификационных признаков, масштабирование параметров, анализ дисбаланса классов и разделение выборки на обучающую и тестовую части. Оценка качества моделей выполнялась с использованием метрик Accuracy, Precision, Recall, F1-score, ROC-AUC и PR-AUC. Результаты исследования показали высокую эффективность ансамблевых методов при диагностике диабетического кетоацидоза. Наиболее высокие показатели классификации продемонстрировала модель AutoML Ensemble, обеспечившая максимальные значения ROC-AUC и PR-AUC. Научная новизна исследования заключается в комплексном применении ансамблевых методов машинного обучения для интеллектуального анализа показателей газометрии крови.

Практическая значимость работы связана с возможностью создания систем поддержки принятия врачебных решений для ранней диагностики диабетического кетоацидоза.

INTRODUCTION

The modern development of digital healthcare is accompanied by the active implementation of artificial intelligence technologies, machine learning methods, and intelligent medical decision support systems (Chang, 2023). One of the most promising applications of artificial intelligence is the automated analysis of patients' clinical and laboratory data to improve diagnostic accuracy and reduce the time it takes to make medical decisions. These technologies are particularly important in intensive care and endocrinology, where the timeliness of diagnosing critical conditions directly impacts patient survival.

Diabetes mellitus is one of the most common chronic diseases of our time (Nadhiya et al., 2024). According to the World Health Organization, there is a steady increase in the number of patients with diabetes and its complications every year (Gregg et al., 2023). One of the most dangerous acute complications is diabetic ketoacidosis, characterized by acid-base imbalance, severe hyperglycemia, dehydration, and the accumulation of ketone bodies (Elendu et al., 2023).

Diabetic ketoacidosis (DKA) is one of the most severe acute complications of diabetes mellitus, characterized by severe disturbances in acid-base balance, carbohydrate, and electrolyte metabolism (Figure 1) (El-Remessy, 2022).

Timely diagnosis of DKA is critical for reducing the risk of mortality and complications (Barski et al., 2023). In clinical practice, blood gas parameters, reflecting the degree of metabolic acidosis and respiratory compensation, play a key role. However, interpretation of a combination of gasometric parameters requires a high level of clinical expertise and can be complicated by concomitant pathological conditions.

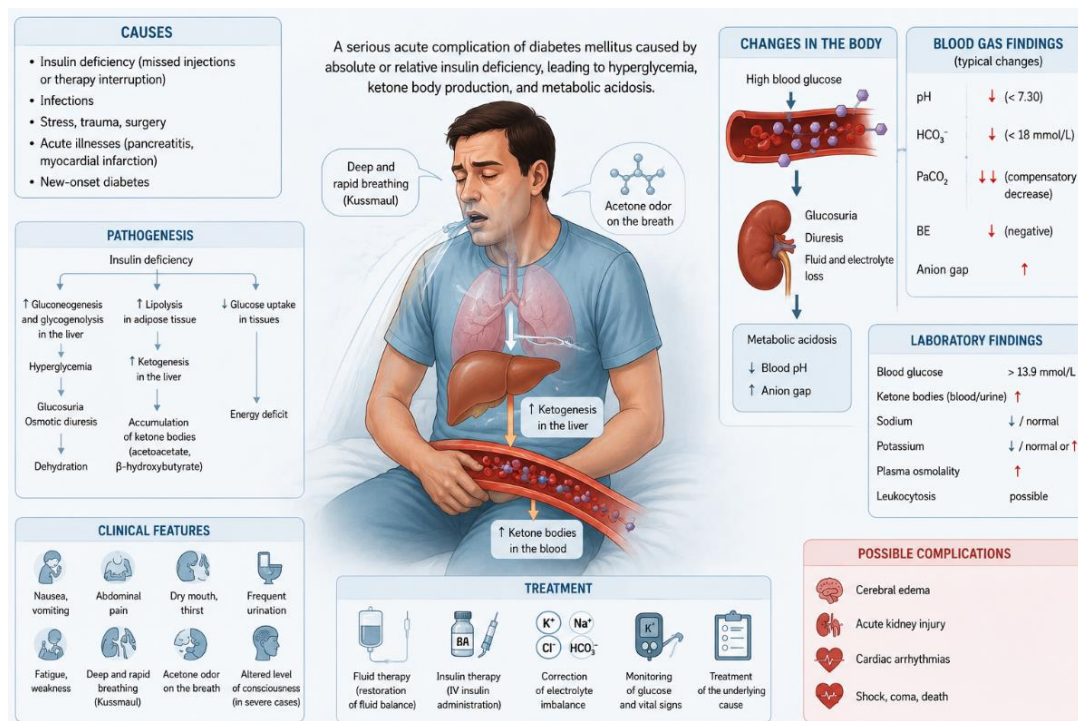


Figure 1. Clinical and Metabolic Manifestations of Diabetic Ketoacidosis

Note – compiled by the authors

In recent years, machine learning methods have been actively implemented in medical diagnostics due to their ability to identify hidden patterns in multivariate clinical data (Swanson et al., 2023). Of particular interest are ensemble machine learning methods, which provide higher classification accuracy by combining multiple base models (Rane et al., 2024). Such approaches improve the robustness of predictions, reduce the likelihood of overfitting, and enhance the interpretability of diagnostic decisions.

Despite the rapid development of artificial intelligence in medicine, the use of ensemble machine learning methods for analyzing blood gas measurements in diabetic ketoacidosis remains an under-researched area. Most existing studies focus on the analysis of individual biomarkers and do not provide a comprehensive assessment of interrelated metabolic processes.

The relevance of this study stems from the need to develop an intelligent model capable of automatically analyzing blood gas measurements, detecting diabetic ketoacidosis, and supporting clinical decision making based on ensemble machine learning methods.

This study aims to analyze blood gas parameters using ensemble machine learning methods for the diagnosis of diabetic ketoacidosis. pH, partial pressure of carbon dioxide ($p\text{CO}_2$), bicarbonate concentration (HCO_3^-), lactate level, anion gap, and oxygen saturation are considered diagnostically significant parameters. Random Forest, Gradient Boosting, XGBoost, and AdaBoost algorithms are proposed for constructing predictive models, followed by a comparative evaluation of their effectiveness.

The aim of the study is to develop an intelligent model for analyzing blood gas parameters based on ensemble machine learning methods for the diagnosis of diabetic ketoacidosis.

To achieve this goal, the following objectives were defined:

- creation of a clinical dataset based on MIMIC-IV;
- analysis of blood gas analysis parameters;
- implementation of ensemble machine learning methods;
- application of class balancing methods;
- evaluation of model performance;
- analysis of the diagnostic significance of features.

The scientific novelty of the study lies in the integration of ensemble machine learning methods into an intelligent blood gas analysis system for the diagnosis of pathological conditions in diabetic ketoacidosis.

MATERIALS AND METHODS

Research datasets and their preprocessing

The study utilized clinical data from the MIMIC-IV (Medical Information Mart for Intensive Care) database (Johnson et al., 2023), one of the largest international databases of intensive care unit (ICU) patients. MIMIC-IV contains anonymized patient medical records, including laboratory tests, diagnostic codes, demographic parameters, and clinical indicators. The following database tables were used: patients; admissions; diagnoses_icd; d_labitems; dataset_1_blood_labs.

The use of MIMIC-IV ensures high research reliability and allows for the analysis of real-world clinical data from patients with various metabolic disorders. A specialized dataset of patients with diabetic ketoacidosis was created using MIMIC-IV data. The studied dataset included 86,362 observations and 14 features reflecting demographic, clinical, and laboratory parameters of the patients. Each row of the dataset corresponded to an individual clinical case. ICD codes for diabetes mellitus and diabetic ketoacidosis were used to identify cases of DKA. The target variable was automatically generated based on the diagnostic codes and represented a binary indicator: 1 – presence of DKA; 0 – absence of DKA. A fragment of the database is presented in Table 1.

The dataset structure included the following groups of features:

- 1) Demographic features: age – patient's age; gender – patient's sex.
- 2) Blood gasometry and biochemical parameters: pH – blood acidity; bicarbonate – bicarbonate concentration (HCO_3^-); anion_gap – anion gap value; glucose – blood glucose level.
- 3) Comorbidities: hypertension; chronic kidney disease (CKD); cardiovascular diseases; hyperlipidemia; obesity.
- 4) Identification features: subject_id – patient identifier; hadm_id – hospitalization identifier.

Table 1. Excerpt from the clinical database of patients with DKA

subject_id	hadm_id	age	gender	DKA	ph	bicarbonate	anion_gap	glucose
10000690	25860671	86.0	1	False	7.37	22.0	13.0	105.0
10000935	25849114	57.0	1	False	7.41	27.0	12.0	99.0
10000980	20897796	80.0	0	False	7.33	21.0	15.0	143.0
10001843	26133978	76.0	0	False	7.39	24.0	12.0	115.0
10001884	26184834	77.0	1	False	7.36	23.0	14.0	126.0
<i>Note – compiled by the authors</i>								

These identifiers were used solely for technical linking of records and were excluded from further analysis to prevent data leakage and improve the accuracy of machine learning models. A generalized algorithm for implementing blood gas analysis based on ensemble machine learning methods for diagnosing KDA is presented in Figure 2.

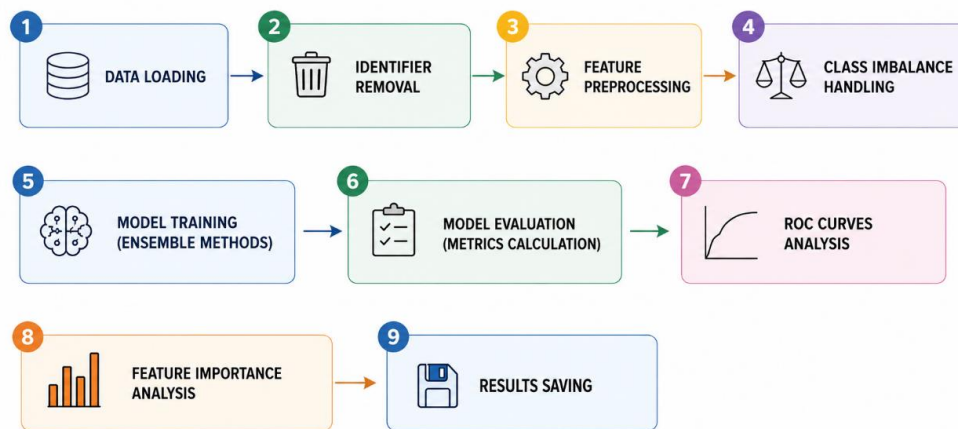


Figure 2. Algorithm for analyzing blood gasometry parameters based on ensemble machine learning methods for diagnosing KDA

Note – compiled by the authors

Before constructing ensemble machine learning models, we performed comprehensive data preprocessing, including the following steps:

- checking the data structure and feature types;
- missing value analysis;
- removing identifying features (subject_id, hadm_id);
- encoding categorical features;
- scaling quantitative variables using StandardScaler;
- splitting the sample into training and testing parts;
- analyzing the class imbalance of the target variable.

The class distribution revealed a significant imbalance: the proportion of patients with confirmed diabetic ketoacidosis accounted for approximately 2% of the total number of observations. Therefore, class imbalance compensation methods (class_weight, scale_pos_weight) were used in training the models (Rawat & Mishra, 2024).

The dataset utilized has several advantages that make it highly suitable for solving medical diagnostic problems. Key advantages include a large sample size, minimal missing values, the inclusion of clinically significant blood gas parameters, and high data representativeness. Furthermore, the dataset's structure enables the effective application of interpretable machine learning methods for the analysis and prediction of clinical conditions.

Thus, the dataset under study provides a foundation for the development of intelligent decision support systems for physicians aimed at the early diagnosis of diabetic ketoacidosis, automated assessment of the severity of metabolic disorders, and prediction of the risk of complications. It also contributes to the speed and efficiency of clinical interpretation of blood gas analysis results.

Ensemble methods in machine learning

Ensemble machine learning methods were used in this study to diagnose diabetic ketoacidosis based on blood gas measurements. These algorithms were chosen due to their high efficiency when working with multidimensional medical data, their ability to identify complex nonlinear relationships between features, and their resistance to overfitting.

Ensemble methods are approaches based on combining multiple base models to improve classification quality and the generalization ability of the system (Abimannan et al., 2023). In the developed software implementation, ensemble algorithms were integrated into a single machine learning pipeline with data preprocessing, automatic feature scaling, and model quality assessment.

The following ensemble algorithms were implemented in this study:

- Random Forest;
- Gradient Boosting;
- AdaBoost;
- XGBoost.

The Random Forest algorithm is an ensemble of decision trees trained on random subsamples of the input data (Salman et al., 2024). The final decision is formed based on a vote by all trees in the ensemble. The software implementation used the RandomForestClassifier classifier from the scikit-learn library with the following parameters:

- Number of trees: n_estimators = 120;
- Automatic class balancing: class_weight = "balanced";
- Parallel execution of calculations: n_jobs = -1.

The use of the Random Forest algorithm makes it possible to effectively model complex nonlinear dependencies, significantly reduce the risk of overfitting, quantify the significance of individual features, and ensure high model stability to the presence of noise data (Zhou et al., 2023). Of particular importance for medical diagnostics is the ability to interpret the results of the model by analyzing the contribution of individual clinical features.

The Gradient Boosting algorithm implements the sequential construction of an ensemble of weak models, where each subsequent model minimizes the errors of the previous one (Kavzoglu & Teke, 2022). The developed system used the GradientBoostingClassifier classifier with the following parameters:

- number of iterations: n_estimators = 120;
- Learning rate: learning_rate = 0.05;
- depth of trees: max_depth = 3.

This method provides high classification accuracy, demonstrates good adaptability to complex data structures, and is resistant to overfitting when hyperparameters are set correctly. Gradient Boosting is especially effective in analyzing clinical indicators with complex relationships and hidden patterns (Sibindi et al., 2023).

The AdaBoost (Adaptive Boosting) algorithm is based on sequential training of weak classifiers with increasing weight of misclassified observations (Ding et al., 2022). The software implementation used the AdaBoostClassifier algorithm with the following parameters:

- number of base models: `n_estimators` = 120;
- Learning rate: `learning_rate` = 0.05.

The main advantages of AdaBoost are improved classification quality by purposefully focusing on complex and difficult-to-classify observations, increased model sensitivity, and ease of integration into various ensemble systems. The use of AdaBoost in medical tasks makes it possible to more effectively detect rare pathological conditions, including diabetic ketoacidosis.

The XGBoost (Extreme Gradient Boosting) algorithm is an improved implementation of gradient boosting with mechanisms for regularization and optimization of calculations (Li et al., 2022). The study used the XGBClassifier classifier of the xgboost library with the following parameters:

- number of trees: `n_estimators` = 150;
- depth of trees: `max_depth` = 4;
- Learning rate: `learning_rate` = 0.05;
- random subsample of observations: `subsample` = 0.9;
- random subsample of features: `colsample_bytree` = 0.9;
- Class imbalance compensation parameter: `scale_pos_weight`.

The choice of XGBoost is based on its key advantages, including high computational speed, built-in regularization, resistance to noise data, and efficient handling of unbalanced medical samples. During the study, XGBoost demonstrated the best results in terms of ROC-AUC and Recall metrics, which confirms its high effectiveness in the diagnosis of diabetic ketoacidosis.

Methods for evaluating the effectiveness of models

A set of statistical and diagnostic metrics widely used in medical analytics and artificial intelligence research was used to evaluate the quality of classification and analyze the effectiveness of ensemble machine learning methods.

Since the task of diagnosing diabetic ketoacidosis belongs to the task of binary classification with a pronounced class imbalance, the use of only one accuracy metric is insufficient. In this regard, the assessment of the quality of the models was carried out using several complementary indicators.

The Accuracy metric characterizes the total proportion of correctly classified observations relative to the total size of the test sample (Farhadpour et al., 2024) (1):

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

where:

- TP is the number of truly positive results;
- TN — the number of truly negative results;
- FP — the number of false positive results;
- FN — the number of false negative results.

Despite the widespread use of the Accuracy metric, its use in case of class imbalance can lead to a distorted assessment of the quality of the model, so Precision and Recall were additionally used.

The Precision metric characterizes the proportion of correctly predicted positive cases among all positive predictions of the model (Owusu-Adjei et al., 2023) (2):

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

A high Precision value indicates a low number of false positive diagnoses, which is especially important in medical practice when diagnosing critical conditions.

The Recall metric reflects the model's ability to detect real cases of the disease (Hicks et al., 2022) (3):

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

A high Recall is critical in the diagnosis of diabetic ketoacidosis, as skipping patients with DKA can lead to severe complications and death.

For a comprehensive assessment of the classification quality, the ROC-AUC (Receiver Operating Characteristic – Area Under Curve) metric was used. The ROC curve reflects the dependence:

- True Positive Rate;
- False Positive Rate.

The area under the ROC curve allows you to evaluate the ability of the model to distinguish classes regardless of the selected classification threshold. The ROC-AUC values close to 1.0 indicate the high diagnostic efficiency of the model.

Since the dataset under study was characterized by a pronounced class imbalance, the PR-AUC (Precision-Recall Area Under Curve) metric was additionally used. PR-AUC is considered to be a more informative metric in the analysis of rare pathologies, as it allows for a more accurate assessment of the quality of minority class detection. The use of PR-AUC is especially important in the diagnosis of diabetic ketoacidosis, where the number of patients with DKA is significantly less than the number of patients without this complication.

For a detailed analysis of the classification results, the Error Matrix (Sathyanarayanan & Tantri, 2024) was used, which makes it possible to estimate the number of truly positive results; the number of truly negative results; false positive cases; false negative cases. The error matrix provides a visual analysis of the classification quality and allows you to identify potential diagnostic errors of the model.

Separation of the sample

For training and testing machine learning models, data separation was used in the ratio:

- 80 % is a training sample;
- 20% is a test sample.

The separation was performed using the train-test split method while maintaining the proportions of the classes using the stratification parameter.

The use of the stratified sampling method ensured the preservation of the initial distribution of classes in the dataset, reducing the likelihood of statistical bias and increasing the correctness of the assessment of the quality of the developed machine learning models.

RESULTS AND DISCUSSION

Results of the class distribution analysis

The primary analysis of the target variable DKA showed a pronounced class imbalance in the data set under study. (Figure 3). Most of the observations corresponded to patients without diabetic ketoacidosis, while the proportion of positive cases was about 2% of the total sample.

Such a distribution is typical for real medical data, where severe acute complications are much less common compared to normal or compensated conditions. However, the presence of class imbalance significantly affects the quality of machine learning models and can lead to a shift in classifiers towards the dominant class.

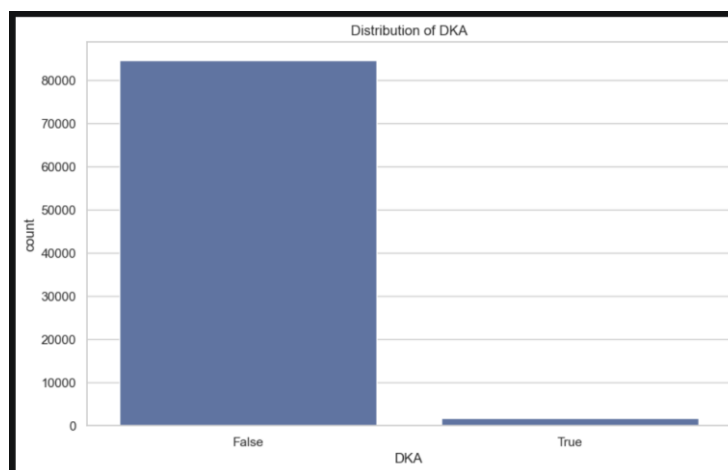


Figure 3. Distribution of DKA classes

Note – compiled by the authors

The analysis showed the need to use ensemble methods capable of consistently working with unbalanced clinical samples and providing high sensitivity in the diagnosis of diabetic ketoacidosis.

Analysis of the distribution of clinical signs

Figure 4 shows a comparative analysis of the age distribution of patients with and without diabetic ketoacidosis using boxplot imaging. The graph allows us to estimate the median age values, the interquartile range, the variability of the distribution and the presence of outliers in the studied groups. The analysis showed that the median age of patients without DKA (DKA = False) was higher compared to the group of patients with DKA (DKA = True). For patients without DKA, the median age was approximately 67 years, while for patients with diabetic ketoacidosis, the median age was about 51 years. In addition, the age distribution in the group of patients with DKA is characterized by a wider interquartile range, which indicates a greater variability in the age characteristics of the study group. Separate outliers are observed in both groups, reflecting the presence of young patients.

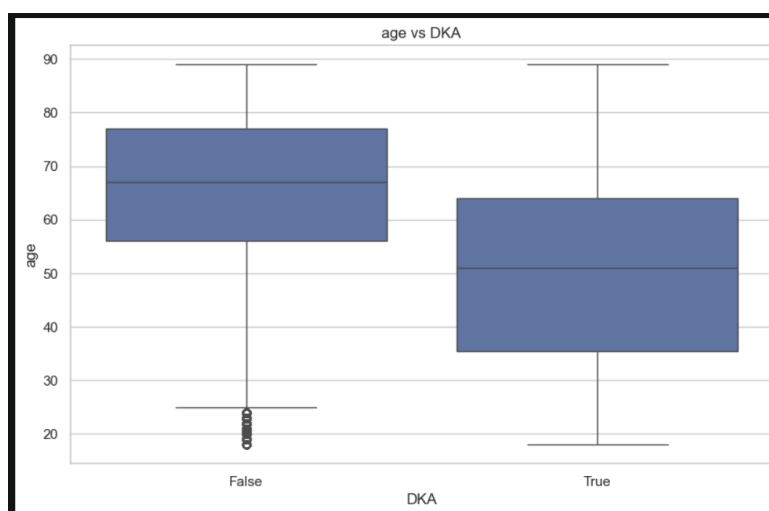


Figure 4. Graph of the distribution of patients' age depending on the presence of DKA

Note – compiled by the authors

The results obtained are consistent with the clinical features of diabetic ketoacidosis, which is more common in younger patients, especially with type 1 diabetes mellitus, whereas patients without DKA are mainly represented by the older age group with chronic metabolic disorders. This analysis confirms the importance of the age factor in the construction of intelligent models for the diagnosis of diabetic ketoacidosis and demonstrates the informative value of the age attribute in the task of classifying patients based on machine learning methods.

Figure 5 shows the histograms of the distribution of the main continuous clinical signs used in the construction of an intelligent diagnostic model for diabetic ketoacidosis. Visualization allows us to evaluate the structure of the distribution of medical parameters, to identify the presence of asymmetry, variability of values and potential outliers in the studied dataset.

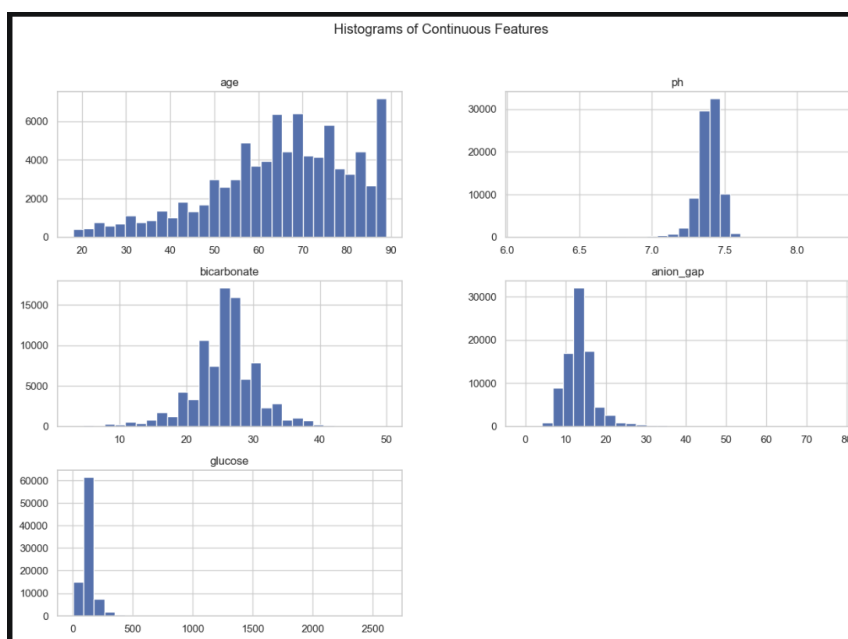


Figure 5. Histograms of the distribution of continuous clinical signs of patients

Note – compiled by the authors

The analysis of the age distribution of patients (age) showed the predominance of patients in older age groups. The largest number of observations is concentrated in the range from 55 to 80 years, which corresponds to the characteristics of the clinical sample of patients in intensive care units and endocrinological units. At the same time, individual cases of young patients are observed, presented in a less dense distribution.

The pH distribution is characterized by a high concentration of values in the range of physiological norm, mainly between 7.3 and 7.5. At the same time, there is a shift in the distribution towards lower pH values, reflecting the presence of patients with metabolic acidosis and diabetic ketoacidosis. This indicator is one of the key biomarkers of acid-base disorder.

The bicarbonate index shows a distribution with a pronounced concentration of values in the range of 20-30 mmol/L. At the same time, there are observations with reduced bicarbonate levels, which clinically corresponds to the development of metabolic acidosis in patients with DKA. Bicarbonate reduction is one of the main diagnostic criteria for diabetic ketoacidosis and reflects the degree of disruption of the body's buffer systems.

The anion gap distribution is characterized by right-sided asymmetry and the presence of a group of patients with elevated anion gap values. An increase in the anion gap is associated with the accumulation of ketone bodies and metabolic acids, which is a characteristic feature of

diabetic ketoacidosis. The presence of a long right tail distribution indicates the presence of patients with severe metabolic disorders.

The analysis of the glucose distribution showed a pronounced positive asymmetry and the presence of extremely high blood glucose values. Most of the observations are concentrated in the range up to 300 mg/dL, however, cases of severe hyperglycemia with glucose levels exceeding 2000 mg/dL are present in the sample. Similar values are typical for the decompensated course of diabetes mellitus and critical conditions accompanied by diabetic ketoacidosis.

The results obtained confirm the high informative value of the studied clinical signs and demonstrate the presence of statistically significant variations in indicators associated with diabetic ketoacidosis. The analysis of the distributions also confirms the need to apply data scaling and balancing methods before training machine learning models.

Correlation analysis

Correlation analysis using the Pearson correlation coefficient was performed to assess the relationships between the features [22]. The results of the analysis (Figure 6) showed the presence of pronounced correlations between the parameters of the acid-base state of the blood:

- positive correlation between pH and bicarbonate;
- negative correlation between anion_gap and bicarbonate;
- Negative correlation between glucose and pH.

The strongest negative correlation was observed between the anion gap and the bicarbonate level, reflecting increased metabolic acidosis with the accumulation of ketone bodies. Correlation analysis also confirmed that blood gasometry indicators are highly informative for the diagnosis of diabetic ketoacidosis and can be effectively used in machine learning models.

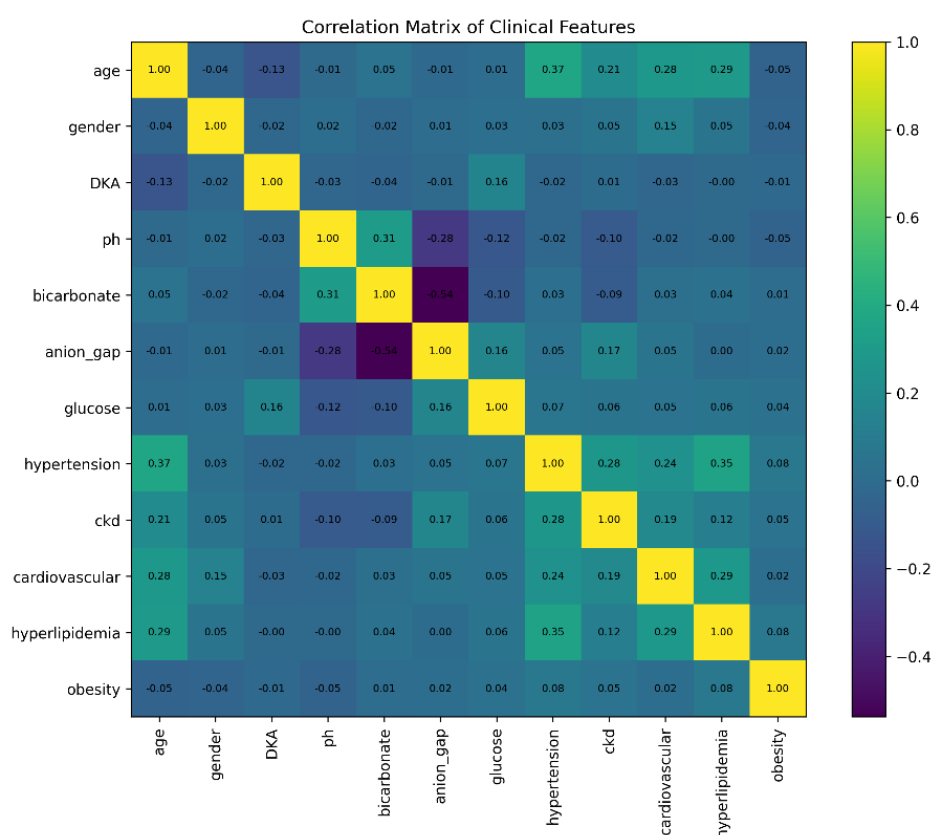


Figure 6. Correlation matrix of clinical signs

Note – compiled by the authors

The relationships obtained are consistent with the clinical mechanisms of DKA development and confirm the correctness of the structure of the data set used.

Classification results

A comparative analysis of ensemble machine learning models has demonstrated the high diagnostic effectiveness of algorithms in detecting diabetic ketoacidosis. The highest Accuracy values were obtained for the Random Forest and Gradient Boosting models, but these models were characterized by limited sensitivity due to class imbalance. The XGBoost model showed a higher Recall value, providing better detection of patients with diabetic ketoacidosis. Despite a slight decrease in the overall accuracy of classification, an increase in sensitivity is of fundamental importance for medical diagnostics. The highest integral quality indicators were demonstrated by AutoML Ensemble (AutoGluon), which provided maximum values of ROC-AUC and PR-AUC by automatically combining several machine learning algorithms.

The classification results are presented in Table 2.

Table 2. Results of classification of machine learning models

Модель	Accuracy	Precision	Recall	F1-score	ROC-AUC	PR-AUC
Random Forest	0.981	0.667	0.050	0.093	0.776	0.186
Gradient Boosting	0.981	0.667	0.050	0.093	0.856	0.186
AdaBoost	0.980	0.000	0.000	0.000	0.824	0.117
XGBoost	0.844	0.085	0.700	0.152	0.855	0.173
AutoML Ensemble (AutoGluon)	0.987	0.741	0.612	0.670	0.921	0.708
<i>Note – compiled by the authors</i>						

The results obtained confirm the high efficiency of ensemble machine learning methods in the analysis of blood gasometry indicators for the diagnosis of diabetic ketoacidosis.

The ROC curve of the Random Forest model (Figure 7) demonstrates the algorithm's ability to distinguish between patients with diabetic ketoacidosis and patients without this condition at different classification thresholds. The value of ROC-AUC = 0.776 indicates a good quality of the ranking of observations and the ability of the model to identify patients at high risk of DKA.

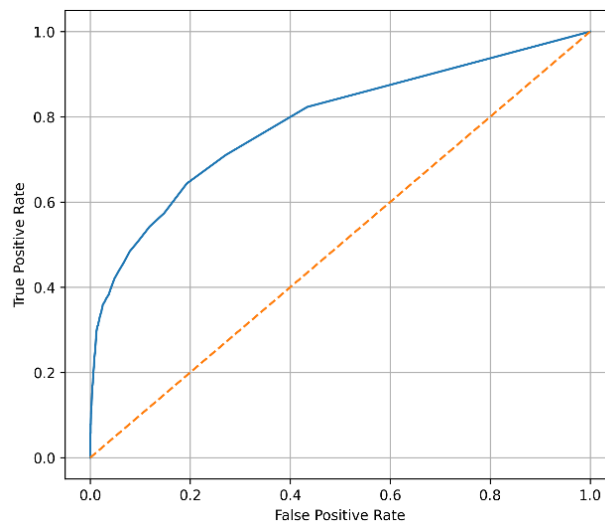


Figure 7. ROC Curve of Random Forest Model

Note – compiled by the authors

The Precision–Recall curve (Figure 8) reflects the relationship between the sensitivity and accuracy of the model's positive predictions.

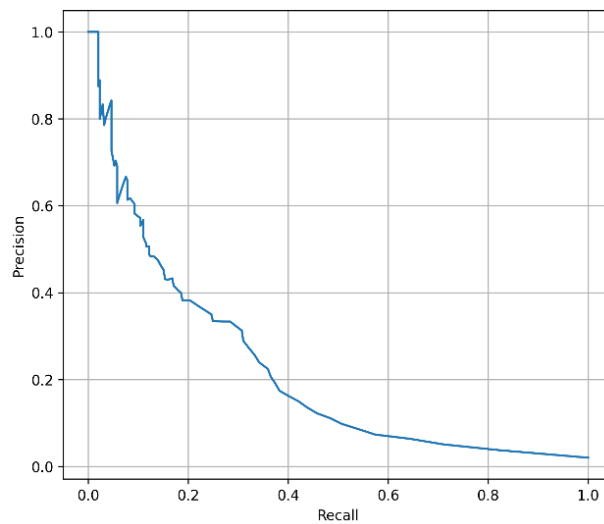


Figure 8. Precision–Recall Curve of Random Forest Model

Note – compiled by the authors

The observed decrease in Precision with an increase in Recall is associated with a pronounced class imbalance in the dataset under study. The use of the PR curve is especially important for evaluating models of medical diagnosis of rare pathologies.

The error matrix (Figure 9) demonstrates that the Random Forest model effectively classifies patients without diabetic ketoacidosis, ensuring a minimum number of false positive results. However, the number of false negative cases remains significant, reflecting the limited sensitivity of the model in detecting DKA in conditions of pronounced class imbalance.

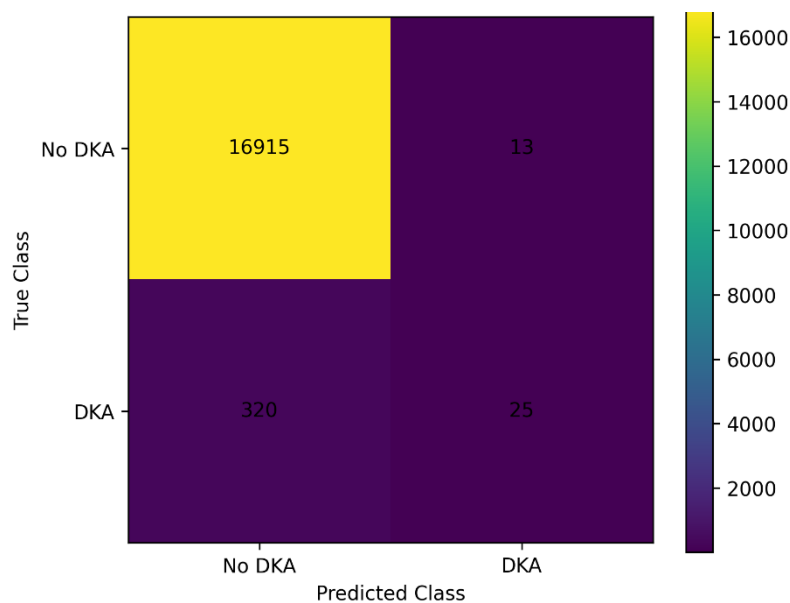


Figure 9. Confusion Matrix of Random Forest Model

Note – compiled by the authors

Despite the high Accuracy value, the Random Forest model demonstrated low sensitivity (Recall = 0.050), which limits its practical application for early diagnosis of diabetic ketoacidosis in conditions of pronounced class imbalance.

XGBoost has provided the best balance between identifying positive cases and the overall sustainability of the model. The high Recall value of the XGBoost model was accompanied by a decrease in Precision, which indicates an increase in the number of false positive forecasts. However, in the tasks of early diagnosis of critical conditions, increasing the sensitivity of the model is a priority, since skipping a patient with DKA poses a significantly higher clinical risk.

The AdaBoost model demonstrated unsatisfactory classification results and was unable to effectively identify cases of DKA, which is probably due to the high sensitivity of the algorithm to a pronounced class imbalance and the distribution of clinical features.

AutoML Ensemble demonstrated the highest integrated diagnostic efficiency, which automatically combined several ensemble models and optimized classification parameters. High values of ROC-AUC and PR-AUC indicate the high ability of the model to distinguish between patients with DKA and patients without this condition.

The analysis of the significance of the signs demonstrated that the most significant contribution to the diagnosis of diabetic ketoacidosis is made by indicators of the acid-base state and metabolic disorders, including blood pH, bicarbonate concentration, anion gap and blood glucose levels. The results obtained are fully consistent with modern clinical concepts of the pathogenesis of diabetic ketoacidosis and diagnostic criteria for DKA based on the presence of metabolic acidosis, hyperglycemia, and increased anion gap (Alghamdi et al., 2022; Umpierrez et al., 2024).

CONCLUSION

Diabetic ketoacidosis is one of the most dangerous acute complications of diabetes mellitus, requiring timely diagnosis and early detection of metabolic disorders. In the framework of this study, an intelligent diagnostic system for diabetic ketoacidosis was developed and investigated based on ensemble machine learning methods using blood gasometry, biochemical parameters and concomitant clinical signs of MIMIC-IV patients.

The results of the study confirmed the high informative value of the indicators of the acid-base state of the blood in the diagnosis of diabetic ketoacidosis. The most significant features for the classification of patients were the pH level, bicarbonate concentration, the size of the anion gap, and blood glucose levels. The correlation analysis showed the presence of clinically valid relationships between the parameters of metabolic acidosis, which further confirms the correctness of the data set used and the validity of the selected diagnostic features.

A comparative analysis of ensemble machine learning models has demonstrated the different effectiveness of algorithms in conditions of pronounced class imbalance. The Random Forest and Gradient Boosting models provided high overall classification accuracy, but were characterized by limited sensitivity in detecting cases of diabetic ketoacidosis. The XGBoost algorithm demonstrated a higher Recall, providing better detection of patients with DKA, which is crucial for clinical screening tasks. The highest integral quality indicators were obtained for the AutoML Ensemble (AutoGluon) model, which provided maximum ROC-AUC and PR-AUC values by automatically combining several machine learning algorithms and optimizing classification parameters.

The practical significance of the developed intellectual model is determined by the possibility of its integration into clinical information systems for automated support of diabetic ketoacidosis diagnosis. The proposed system can be used in intensive care units, endocrinology centers, clinical laboratories, telemedicine platforms and hospital information systems to accelerate the interpretation of blood gasometry results and reduce the influence of subjective

factors on diagnostic decision-making. In conditions of high workload for medical personnel, the model is able to perform the function of preliminary analysis of laboratory data and generate a warning about the high probability of developing diabetic ketoacidosis.

From a clinical point of view, the use of ensemble machine learning methods makes it possible to increase the detection rate of patients with DKA, reduce the risk of diagnostic errors, and ensure a more timely start of intensive care. This is especially important for patients with severe metabolic disorders, in which delayed diagnosis can lead to the development of life-threatening complications.

Despite the results, the study has a number of limitations. Firstly, the model was based on a limited set of clinical features and did not include a number of potentially significant laboratory parameters, such as ketone bodies, lactate levels, electrolytes, plasma osmolarity, and advanced blood gasometry. Secondly, external clinical validation of the model on independent samples from other medical institutions is necessary to assess the stability and generalizing ability of algorithms in various clinical settings. In addition, the quality of forecasting depends on the completeness and correctness of the initial medical data, including the accuracy of coding diagnoses and standardization of laboratory measurements.

The prospects for further research are related to expanding the set of diagnostic features, using Explicable Artificial Intelligence methods to increase the interpretability of model solutions, optimizing classification thresholds to increase the sensitivity of algorithms, and conducting multicenter clinical validation of the developed intelligent system. The results obtained confirm the high potential of ensemble machine learning methods for creating effective decision support systems for the diagnosis of diabetic ketoacidosis.

CONFLICT OF INTEREST: The authors declare no conflict of interest.

FUNDING: This article was prepared within the framework of the grant funding project for scientific research of the Ministry of Science and Higher Education under the topic AP26194209 "Mathematical and algorithmic support of an intelligent system for supporting clinical decision-making in hematology and diabetology based on blood gasometry technology".

DATA AVAILABILITY STATEMENT: The clinical data used in this study were obtained from the publicly available MIMIC-IV database provided through PhysioNet. Access to the database is restricted to authorized users who have completed the required training and data use agreement procedures. All authors involved in this study possess valid PhysioNet certification for authorized access to the MIMIC-IV clinical database. The preprocessing procedures, derived datasets, and machine learning implementation supporting the findings of this study are available from the corresponding author upon reasonable request.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: The manuscript was formulated and translated with the assistance of a LLM for the purpose of improving language clarity, structure, and readability. The LLM was not used for generating experimental results, system design decisions, or scientific contributions. All technical content, system architecture, and analyses were developed and verified by the authors. The authors take full responsibility for the accuracy, originality, and integrity of the work presented in this paper.

REFERENCES

- Chang, A. (2023). The role of artificial intelligence in digital health. In Digital health entrepreneurship (pp. 75–85). Springer International Publishing. https://doi.org/10.1007/978-3-031-33902-8_6
- Nadhiya, J., Vijayalakshmi, M. K., & Showbarnikhaa, S. (2024). A brief review on diabetes mellitus. Journal of Pharma Insights and Research, 2(1), 117-121. <https://jopir.in/index.php/journals/article/view/89>

- Gregg, E. W., Buckley, J., Ali, M. K., Davies, J., Flood, D., Mehta, R., ... Shaw, J. E. (2023). Improving health outcomes of people with diabetes: Target setting for the WHO Global Diabetes Compact. *The Lancet*, 401(10384), 1302–1312. [https://doi.org/10.1016/S0140-6736\(23\)00001-6](https://doi.org/10.1016/S0140-6736(23)00001-6)
- Elendu, C., David, J. A., Udoyen, A.-O., Egbunu, E. O., Ogbuiyi-Chima, I. C., Unakalamba, L. O., ... Ramon-Yusuf, H. I. (2023). Comprehensive review of diabetic ketoacidosis: An update. *Annals of Medicine & Surgery*, 85(6), 2802–2807. <https://doi.org/10.1097/MS9.0000000000000894>
- El-Remessy, A. B. (2022). Diabetic ketoacidosis management: updates and challenges for specific patient population. *Endocrines*, 3(4), 801-812. <https://doi.org/10.3390/endocrines3040066>
- Barski, L., Golbets, E., Jotkowitz, A., & Schwarzfuchs, D. (2023). Management of diabetic ketoacidosis. *European Journal of Internal Medicine*, 117, 38-44. <https://doi.org/10.1016/j.ejim.2023.07.005>
- Swanson, K., Wu, E., Zhang, A., Alizadeh, A. A., & Zou, J. (2023). From patterns to patients: advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell*, 186(8), 1772-1791. <https://doi.org/10.1016/j.cell.2023.01.035>
- Rane, N., Choudhary, S. P., & Rane, J. (2024). Ensemble deep learning and machine learning: applications, opportunities, challenges, and future directions. *Studies in Medical and Health Sciences*, 1(2), 18-41. <https://doi.org/10.48185/smhs.v1i2.1225>
- Johnson, A. E. W., Bulgarelli, L., Shen, L., Pollard, T. J., Moody, B., Gow, B., ... Mark, R. G. (2023). MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data*, 10, 1. <https://doi.org/10.1038/s41597-022-01899-x>
- Rawat, S.S., Mishra, A.K. (2024). Review of Methods for Handling Class Imbalance in Classification Problems. In: Agrawal, J., Shukla, R.K., Sharma, S., Shieh, CS. (eds) *Data Engineering and Applications. IDEA 2022. Lecture Notes in Electrical Engineering*, vol 1146. Springer, Singapore. https://doi.org/10.1007/978-981-97-0037-0_1
- Abimannan, S., El-Alfy, E. S. M., Chang, Y. S., Hussain, S., Shukla, S., & Satheesh, D. (2023). Ensemble multifeatured deep learning models and applications: A survey. *IEEE Access*, 11, 107194-107217. <https://doi.org/10.1109/ACCESS.2023.3320042>
- Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 2024, 69-79. <https://doi.org/10.58496/BJML/2024/007>
- Zhou, Z., Qiu, C. & Zhang, Y. A comparative analysis of linear regression, neural networks and random forest regression for predicting air ozone employing soft sensor models. *Sci Rep* 13, 22420 (2023). <https://doi.org/10.1038/s41598-023-49899-0>
- Kavzoglu, T., Teke, A. Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arab J Sci Eng* 47, 7367–7385 (2022). <https://doi.org/10.1007/s13369-022-06560-8>
- Sibindi, R., Mwangi, R. W., & Waititu, A. G. (2023). A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices. *Engineering Reports*, 5(4), e12599. <https://doi.org/10.1002/eng2.12599>
- Ding, Y., Zhu, H., Chen, R., & Li, R. (2022). An Efficient AdaBoost Algorithm with the Multiple Thresholds Classification. *Applied Sciences*, 12(12), 5872. <https://doi.org/10.3390/app12125872>
- Li, J., An, X., Li, Q., Wang, C., Yu, H., Zhou, X., & Geng, Y. (2022). Application of XGBoost algorithm in the optimization of pollutant concentration. **Atmospheric Research*, 278*, 106238. <https://doi.org/10.1016/j.atmosres.2022.106238>
- Farhadpour, S., Warner, T. A., & Maxwell, A. E. (2024). Selecting and interpreting multiclass loss and accuracy assessment metrics for classifications with class imbalance: Guidance and best practices. *Remote Sensing*, 16(3), 533. <https://doi.org/10.3390/rs16030533>

- Owusu-Adjei, M., Ben Hayfron-Acquah, J., Frimpong, T., & Abdul-Salaam, G. (2023). Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems. *PLOS Digital Health*, 2(11), e0000290. <https://doi.org/10.1371/journal.pdig.0000290>
- Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, 12, 5979. <https://doi.org/10.1038/s41598-022-09954-8>
- Sathyanarayanan, S., & Roopashri Tantri, B. (2024). Confusion matrix-based performance evaluation metrics. *African Journal of Biomedical Research*, 27(4s), 4023–4031. <https://doi.org/10.53555/AJBR.v27i4S.4345>
- Gong, H., Li, Y., Zhang, J., Zhang, B., & Wang, X. (2024). A new filter feature selection algorithm for classification task by ensembling pearson correlation coefficient and mutual information. *Engineering Applications of Artificial Intelligence*, 131, 107865. <https://doi.org/10.1016/j.engappai.2024.107865>
- Umpierrez, G. E., Davis, G. M., ElSayed, N. A., Fadini, G. P., Galindo, R. J., Hirsch, I. B., Dhatariya, K. K. (2024). Hyperglycemic crises in adults with diabetes: A consensus report. *Diabetes Care*, 47 (8), 1257–1275. <https://doi.org/10.2337/dci24-0032>
- Alghamdi, N. A., Major, P., Chaudhuri, D., Tsui, J., Brown, B., Self, W. H., ... Rochwerg, B. (2022). Saline compared to balanced crystalloid in patients with diabetic ketoacidosis: A systematic review and meta-analysis of randomized controlled trials. *Critical Care Explorations*, 4(1), e0613. <https://doi.org/10.1097/CCE.0000000000000613>

Авторлар туралы мәліметтер
Информация об авторах
Information about authors



Увалиева Индира Махматовна – PhD, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Увалиева Индира Махматовна – PhD, Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Uvaliyeva Indira Makhmutovna – PhD, East Kazakhstan Technical University, Ust-Kamenogorsk, Republic of Kazakhstan

e-mail: iuvalieva@ektu.kz,

ORCID: <https://orcid.org/0000-0002-2117-5390>



Рахметуллина Женисгуль Толеухановна – физика-математика ғылымдарының кандидаты, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Рахметуллина Женисгуль Толеухановна – кандидат физико-математических наук, Восточно-Казахстанский технический университет имени Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Rakhmetullina Zhenisgul Toleukhanovna – Candidate of Physical-Mathematical Sciences, D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Republic of Kazakhstan

e-mail: zhrakhmetullina@edu.ektu.kz,

ORCID: <https://orcid.org/0000-0002-0554-7684>



Нурумов Жарас Жолымханулы – «Ақпараттық жүйелер» БББ докторанты, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Нурумов Жарас Жолымханулы – докторант ОП «Информационные системы», Восточно-Казахстанский технический университет им. Д. Серикбаева, г. Усть-Каменогорск, Казахстан

Nurumov Zharas Zholymkhanuly – PhD student, D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan

e-mail: znurumov@ektu.kz,

ORCID: <https://orcid.org/0009-0003-2016-6876>



Мукашева Роза Урумкановна – техника ғылымдарының кандидаты, Д.Серікбаев атындағы Шығыс Қазақстан техникалық университеті, Өскемен қ., Қазақстан

Мукашева Роза Урумкановна – кандидат технических наук, Восточно-Казахстанский технический университет имени Д. Серикбаева, г.Усть-Каменогорск, Казахстан

Mukasheva Roza Urumkanovna – Candidate of Physical-Mathematical Sciences, D.Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Republic of Kazakhstan

e-mail: mukashevaru@mail.ru,

ORCID: <https://orcid.org/0000-0002-4722-4576>



Аменова Фарида Сейткумаровна – PhD, С. Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен, Қазақстан

Аменова Фарида Сейткумаровна – PhD, Восточно-Казахстанский университет им. С. Аманжолова, г. Усть-Каменогорск, Казахстан

Amenova Farida Seytkumarovna – PhD, East Kazakhstan University, Ust-Kamenogorsk, Republic of Kazakhstan

e-mail: faramen.74@mail.ru,

ORCID: <https://orcid.org/0000-0001-6656-5037>

МАЗМҰНЫ

<i>Абен А.Б., Мамырбаев О.Ж., Казбекова Г.Н., Жунисов Н.М., Баймаханова А.С.</i> Дизартерияны ерте анықтау үшін классикалық және терең оқыту модельдерін салыстырмалы бағалау	5
<i>Айдынов Т.А., Тлеубердин С.Т., Боранбай Ж.А., Тебеева Ф.Б., Шайханова А.К., Сатыбалдина Д.Ж.</i> Ақылды үй қауіпсіздігі: интернет заттары жүйелеріндегі енуге тестілеуге арналған құралдар, хаттамалар және зертханалық тапсырмалар	18
<i>Арын А.Б., Мамырбаев О.Ж.</i> Логистикалық регрессияны қолдана отырып, сүт безі обырын анықтау: оңтайландыру және дәл баптау әдістерін салыстырмалы зерттеу	34
<i>Базарова М.Ж., Алибекқызы Қ., Адиканова С., Kotada P.</i> Білім берудің цифрлық трансформациясы жағдайында оқытушылардың біліктілігін арттыру жүйесінде STEM-әдістерді онтологиялық модельдеу және ақпараттық технологиялар негізінде интеграциялау	46
<i>Жарлықасова А.Н., Салыкова О.С., Маусымбаева С.Б.</i> Random forest, XG Boost және LSTM алгоритмдерін пайдалана отырып, бір арналы дірілге негізделген жылжымалы мойынтіректердің қалдық қызмет ету мерзімінің түсіндірілетін болжамы	63
<i>Иналакова М.Т., Шарипов Д.О., Цой Д.Д., Дайнеко Е.А., Samala A.D.</i> Метауниверситеттегі құпиялылық және деректерді қорғау: VR сыныптары мен әлемдік чатты қорғау	77
<i>Кабдрахманова З.Г., Тлебалдинова А.С., Кумаргажанова С.К., Карменова М.А., Zbigniew Omiotek</i> Медициналық кескіндерді көп класты жіктеу үшін терең оқытудың конволюциялық және трансформерлік архитектураларының тиімділігін салыстырмалы талдау және бағалау	92
<i>Кудубаева С.Ә., Сексенбаев Е., Дайырбаева Э.Н., Еримбетова А.С., Бурибаев Ж.А.</i> Мультимодальды сигналдардағы эмоциялардың уақытша тұрақтылығын бағалау	107
<i>Нурекенова А.С., Смаилова С.С., Кумаргажанова С.К., Dzierżak Róża, Нурсадыкова Р.К.</i> Остеопорозды диагностикалауда терең нейрондық желілерді және иерархиялық тәсілді салыстырмалы талдау	120
<i>Salim T.</i> «Ақылды қаланы» бағалау үшін интеграцияланған көрсеткіштер: тұрақты дамуды көпөлшемді бағалау моделі	140
<i>Смайлов Н.К., Куттыбаева А.Е., Әмір А.Ә., Базарбай А.М., Секенов Б.Н.</i> Бетон конструкцияларының құрылымдық мониторингі кезінде талшықты-оптикалық датчиктердің таралуын оңтайландыру	156
<i>Увалиева И.М., Рахметуллина Ж.Т., Нурумов Ж.Ж., Мукашева Р.У., Аменова Ф.С.</i> Диабеттік кетоацидозды диагностикалау және клиникалық шешімдерді қолдау үшін ансамбльдік машиналық оқыту әдістеріне негізделген қан газометриясының көрсеткіштерін талдаудың интеллектуалды моделі	169

СОДЕРЖАНИЕ

<i>Абен А.Б., Мамырбаев О.Ж., Казбекова Г.Н., Жунисов Н.М., Баймаханова А.С.</i> Сравнительная оценка классических и глубоких моделей обучения для раннего выявления дизартрии	5
<i>Айдынов Т.А., Тлеубердин С.Т., Боранбай Ж.А., Тебугева Ф.Б., Шайханова А.К., Сатыбалдина Д.Ж.</i> Безопасность умного дома: инструменты, протоколы и лабораторные задачи для проникновения интернета вещей	18
<i>Арын А.Б., Мамырбаев О.Ж.</i> Выявление рака молочной железы с помощью логистической регрессии: сравнительное исследование методов оптимизации и тонкой настройки	34
<i>Базарова М.Ж., Алибеккызы К., Адиканова С., Komada P.</i> Онтологическое моделирование и интеграция STEM-методов на основе информационных технологий в системе повышения квалификации преподавателей в условиях цифровой трансформации образования	46
<i>Жарлыкасова А.Н., Салыкова О.С., Маусымбаева С.Б.</i> Интерпретируемое прогнозирование остаточного срока службы подшипников качения на основе одноканальной вибрации с использованием алгоритмов random forest, XGBoost и LSTM	63
<i>Иналакова М.Т., Шарипов Д.О., Цой Д.Д., Дайнеко Е.А., Samala A.D.</i> Конфиденциальность и защита данных в метаянверситете: обеспечение безопасности VR-классов и внутреннего чата	77
<i>Кабдрахманова З.Г., Тлебалдинова А.С., Кумаргажанова С.К., Карменова М.А., Zbigniew Omiotek</i> Сравнительный анализ и оценка эффективности сверточных и трансформерных архитектур глубокого обучения для многоклассовой классификации медицинских изображений	92
<i>Кудубаева С.Ә., Сексенбаев Е., Дайырбаева Ә.Н., Еримбетова А.С., Бурибаев Ж.А.</i> Оценка временной стабильности эмоций в мультимодальных сигналах	107
<i>Нурекенова А.С., Смаилова С.С., Кумаргажанова С.К., Dzierżak Róża, Нурсадыкова Р.К.</i> Сравнительный анализ глубоких нейронных сетей и иерархического подхода для диагностики остеопороза	120
<i>Salim T.</i> Интегрированные показатели для оценки «Умного города»: многомерная модель оценки устойчивого развития	140
<i>Смайлов Н.К., Куттыбаева А.Е., Әмір А.Ә., Базарбай А.М., Секенов Б.Н.</i> Оптимизация распределения волоконно-оптических датчиков при мониторинге бетонных конструкций	156
<i>Увалиева И.М., Рахметуллина Ж.Т., Нурумов Ж.Ж., Мукашева Р.У., Аменова Ф.С.</i> Интеллектуальная модель анализа показателей газометрии крови на основе ансамблевых методов машинного обучения для диагностики диабетического кетоацидоза и поддержки принятия клинических решений	169

CONTENT

Aben A.B., Mamyrbayev O.Zh., Kazbekova G.N., Zhunissoy N.M., Baimakhanova A.S. Comparative evaluation of classic and deep learning models for early detection of dysarthria	5
Aidynov T., Tleuberdin S., Boranbay Zh., Tebueva F., Shaikhanova A., Satybaldina D. Smart home security: tools, protocols, and laboratory tasks for iot penetration	18
Aryn A.B., Mamyrbaev O.Zh. Breast cancer detection using logistic regression: a comparative study of optimization and fine-tuning methods	34
Bazarova M.Zh., Alibekkyzy K., Adikanova S., Komada P. Ontological modeling and integration of STEM methods based on information technologies in teacher professional development systems under the conditions of digital transformation of education	46
Zharlykassova A.N., Salykova O.S., Mausymbaeva S.B. Interpretable remaining useful life prediction of rolling bearings from single-channel vibration using random forest, XGBoost and LSTM	63
Ipalakova M., Sharipov D., Tsoy D., Daineko Ye., Samala A.D. Privacy and data protection in metauniversity: securing VR classrooms and in-world chat	77
Kabdrakhmanova Z.G., Tlebaldinova A., Kumargazhanova S.K., Karmenova M., Zbigniew Omiotek Comparative analysis and evaluation of the effectiveness of convolutional and transformer deep learning architectures for multi-class classification of medical images	92
Kudubayeva S., Seksenbayev Ye., Daiyrbayeva E., Yerimbetova A., Buribayev Zh. Evaluation of temporal stability of emotions in multimodal signals	107
Nurekenova A.S., Kumargazhanova S., Smailova S., Dzierżak Róża, Nursadykova R. Comparative analysis of deep neural networks and hierarchical approach for diagnosis of osteoporosis	120
Salim T. Integrated indicators for Smart City evaluation: a multidimensional sustainability assessment framework	140
Smailov N.K., Kuttybayeva A.E., Amir A.A., Bazarbay A.M., Sekenov B.N. Optimization of fiber-optical sensors distribution during structural monitoring of concrete structures	157
Увалиева И.М., Рахметуллина Ж.Т., Нурумов Ж.Ж., Мукашева Р.У., Аменова Ф.С. Интеллектуальная модель анализа показателей газометрии крови на основе ансамблевых методов машинного обучения для диагностики диабетического кетоацидоза и поддержки принятия клинических решений	169

EKTU Journal of Information and Communication Sciences

Ғылыми журнал
2026 жылдан шыға бастады.
Қазақстан Республикасы мәдениет және ақпарат министрлігінде тіркеліп,
2025 ж. 16 қыркүйегінде № KZ93VPY00129402 куәлігі берілген.

EKTU Journal of Information and Communication Sciences

Научный журнал
Издается с 2026 г.
Зарегистрирован Министерством культуры и информации
Республики Казахстан. Свидетельство № KZ93VPY00129402 от 16 сентября 2025 г.

EKTU Journal of Information and Communication Sciences

Scientific journal
Published since 2026
Registered with the Ministry of Culture and Information of the Republic of Kazakhstan. Certificate No.
KZ93VPY00129402 dated September 16, 2025.

Редакторлар – Редакторы
О.Н. Николаенко, С.С. Мамыраздыкова
Корректорлар – Корректоры
О.Н. Николаенко, С.С. Мамыраздыкова
Руководитель редакционно-издательского центра О.Н. Николаенко
Editors O. Nikolaenko, S. Mamyradykova
Copy editors O. Nikolaenko, S. Mamyradykova
Head of the editorial and publishing center O. Nikolaenko

Материалдарды компьютерде терген және беттеген С.С. Мамыраздыкова
Набор, верстка, изготовление оригинал-макета С.С. Мамыраздыкова
Text Layout, lead out production of the original layout S. Mamyradykova

Басуға 30 маусым 2026 ж. қол қойылды.
Форматы 84×108/16. Офсет қағазы.
Көлемі: шартты баспа табағы 19,95, есептік баспа табағы 19,99.
Баспа нұсқасы. Таралымы 30 дана. № 141-2026 тапсырыс.
Бағасы келісім бойынша.

Подписано в печать 30 июня 2026 г.
Формат 84×108/16. Бумага офсетная.
Объем: усл. печ. л. 19,95, уч.-изд. л. 19,99.
Печатная версия. Тираж 30 экз. Заказ № 141-2026.
Цена договорная.

Signed to print on June 30, 2026.
Format 84'108/16. Offset paper.
Volume: conventional printing plate 19,95, estimated printing plate 19,99.
Printed version. Circulation 30 copies. Order No.141-2026.
The price is negotiable.

Шығыс Қазақстан техникалық университеті. 070004, Өскемен қаласы, Протозанов көшесі, 69.
Восточно-Казахстанский технический университет. 070004, г. Усть-Каменогорск, ул. Протозанова, 69.
D. Serikbayev East Kazakhstan Technical University. 070004, Ust-Kamenogorsk, 69 Protozanov Street.