Volume 1 (2026)
Issue 1

TSARKA SCIENCE

REVIEW OF METHODS OF APPLYING MACHINE LEARNING TECHNOLOGY TO ENSURING INFORMATION SECURITY OF PERSONAL DATA IN MEDICAL SYSTEMS

I. Uvaliyeva[1]*, A. Toktarbayeva[2]

[1]D. Serikbayev East Kazakhstan Technical University, Oskemen, Kazakhstan.

[2]D. Serikbayev East Kazakhstan Technical University, Oskemen, Kazakhstan.

*Corresponding author: indirauvalieva@gmail.com

Abstract

This article explores the application of machine learning (ML) methods in the field of cybersecurity, with a focus on protecting sensitive data in medical information systems. As digitalization in healthcare accelerates, the volume and complexity of cyber threats also increase, demanding more intelligent and adaptive protection mechanisms. Machine learning offers advanced capabilities for detecting and neutralizing threats at an early stage through the automated analysis of large volumes of data. The proposed approach is based on the development of a framework that integrates various ML algorithms for dynamic threat identification. This system allows for real-time diagnostics and ensures high precision in detecting both known and emerging types of cyberattacks. The study highlights the importance of systematically collecting and analyzing statistical data to improve the prediction of potential vulnerabilities. Special attention is given to the application of machine learning in the medical sector, where data confidentiality is of critical importance. The integration of ML into healthcare information systems not only increases their security but also improves their adaptability and resilience. The findings of this study confirm that the use of modern algorithms significantly enhances the effectiveness of cybersecurity measures and helps maintain the privacy of medical records. Additionally, the study identifies several challenges and limitations, such as the need for algorithm adaptation to diverse operational conditions and the requirement for qualified personnel. Despite these challenges, machine learning remains a promising direction for improving the quality and reliability of data protection systems.The implementation

of machine learning in medical cybersecurity opens up new opportunities for creating secure, intelligent, and adaptive systems that effectively respond to the growing number of threats in modern healthcare.

Keywords: *Personal data security, machine learning in medical systems, information security, threats, cyberattack, machine learning methods, cybersecurity.*

1. Introduction

Computerization of almost all spheres of public life is increasingly penetrating the field of medicine every year. Modern conceptual approaches to the computerization of medicine, expressed in a person-centered approach to medical records and the creation of an electronic medical record of patients, have determined the directions of healthcare modernization, expressed in the development of electronic analogues of medical documents, the possibility of effective organized access to any set of medical records and primary research results of the patient by separating medical data from their source, the transition to electronic document management, integration of data on the health status of each person in specialized information processing centers of different levels. New opportunities have opened up for the development of digital medical technologies that allow remote consultations, examinations, and primary information processing in specialized centers, while reducing examination time and increasing diagnostic accuracy [1]. However, against the background of these positive changes, modern effective means of integration and rapid processing of personal medical data are successfully used by intruders who pose a threat to human rights and legitimate interests.

The current legislation of the Republic of Kazakhstan on personal data pays insufficient attention to data protection issues. There are no specific guarantees in Russian legislation against unauthorized collection of personal data [2]. Research on the problem of weak regulation of personal data accumulated on the Internet and information systems of government agencies has shown that currently the Law of the Republic of Kazakhstan “On Personal Data and their Protection” does not satisfy full-fledged legal regulation and protection [3].

The number of cyber attacks on medical organizations in Kazakhstan is increasing every year, as evidenced by the results of an information security study conducted by the State Technical Service [4]. The survey results showed that more than 60% of the threat of information leakage is primarily aimed at personal data that is being sent for sale. As a rule, the list of data that is offered for sale includes the most important information about users such as first name, last name, date of birth, phone number, e-mail, etc.

Over the past 5 years, there have been several attacks on medical organizations in our country, and the loudest during this period was in 2019, when the database with

the full information of 11 million patients of Kazakhstani medical organizations was leaked (DamuMed). The reason for the leak of personal data was an elementary mistake – unauthorized access to the organization's medical documents.

In 2020, a massive leak of personal data and medical laboratory data was prevented. The vulnerability is related to an incorrect configuration of the company's server. Silantyev I. O. and Anikin I. V. (2023) believe that the most widely used method for detecting leaks is the use of machine learning in data content analysis, social network analysis, and device monitoring. Because it is this approach that makes it possible to detect abnormal patterns and behaviors that may indicate a possible leak of personal data [5].

Kazakh scientists emphasize that frequent attacks on medical organizations are associated with a lack of information confidentiality and data protection. According to R. R. Khamitov (2024), in Kazakhstan, one of the key threats is the possibility of unauthorized access to medical data [6].

The purpose of the study is to analyze the methods of applying machine learning technology to ensure the information security of personal data in medical systems.

2. Material and research methods.

The machine learning system is capable of detecting intrusions, as it uses support vector techniques, neural networks, and all these capabilities are based on decision trees that have effective meaningful patterns in anomaly detection systems [23].

Decision trees can analyze information and recognize critical system properties that illustrate malicious actions. What distinguishes the use of decision trees from other methods is that the decision tree provides a rich set of simple rules that can be easily integrated with real-time technologies [24].

The goal of the decision tree machine learning approach is to build a decision tree to check incoming traffic based on the available data set, which allows the information technology department to accurately group new requests. Of the many approaches available for building a decision tree, CRT (classification and regression trees) is of the greatest interest. For example, we will use CART for our intrusion detection system. The following formula is required for this approach (1) [25]:

$$S = \{(a_1, b_1), (a_2, b_2), \dots (a_n, b_n)\}, \quad (1)$$

Where $a_i = (a_i^1, a_i^2, \dots, a_i^{(n)})^T$, $i = 1, 2, \dots, N, a_i$ is an instance of the input data and specifies an entry for the network packet. N is a function in a_i. S – represents the number of package entries included in the set.

$b_i \in \{0, 1, 2, \dots, M-1\}$ – the class that means the output of each record.

For evaluation purposes, we need to enter the Score_indicator and the calculation time. T (calculation time) – build time and time required for detection using the proposed approach [25].

In order to characterize the sample data set as normal and abnormal, we will use four classification options, which are listed in Table 1.

Table 1. Instances of data set classification

Sl.No	Instances	Description
I	The positive truth	The usual case stands out unmistakably
II	Positive lies	An unusual case of mistaken identity is classified as common
III	Negative lies	A common case is mistakenly classified as unusual
IV	The negative truth	The unusual case stands out unmistakably

The variable E indicates the number of applicable cases among the detected examples (2):

$$E = \frac{PT}{PT+PF} \quad (2)$$

The variable T indicates the number of significant cases that have been identified in the total sum of the relevant examples. The variable can be obtained using the formula (3):

$$T = \frac{PT}{PT+NF} \quad (3)$$

Indicators E and T conflict from time to time, and the Score_indicator is a common indicator.

The Score_indicator indicator is calculated based on the average values of E and T, which are calculated using the formula (4):

$$Score\_indicator = \frac{(\theta^2+1)E*T}{\theta^2(E+T)} \quad (4)$$

Where θ=1, the evaluation indicator will be calculated using the following formula (5):

$$Score\_indicator = \frac{2ET}{E+T} \quad (5)$$

The dataset is used as input for the working model. The working model consists of three stages [25]: Stage 1. Pre-processing; Stage 2. Normalization of the decision tree; Stage 3. Building a decision tree.

The input data set usually consists of strings and numbers. Since the string value cannot be compared directly, we need to convert the string to digital form using a string manipulation operation, and this is done at the preprocessing stage. The construction of the system is shown in the figure 1 [25].

The pseudocode of the data set preprocessing process looks like this (6):

$$\begin{gathered} \textbf{\textit{Read}}: data\ set\ DS \\ \textbf{\textit{for}}\ x\ =\ 1\ to\ N\ do \\ \textbf{\textit{for}}\ y\ =\ 1\ to\ N \\ (St, n)\ =\ compute(processed\ DS) \\ \textbf{\textit{end for}} \\ \textbf{\textit{end for}} \\ \textbf{\textit{for}}\ z\ =\ 1\ to\ n\ do \\ processed\ DS(St)\ =\ D(j) \\ \textbf{\textit{end}}\ for \\ \textbf{\textit{return}} processedDS \end{gathered} \quad (6)$$

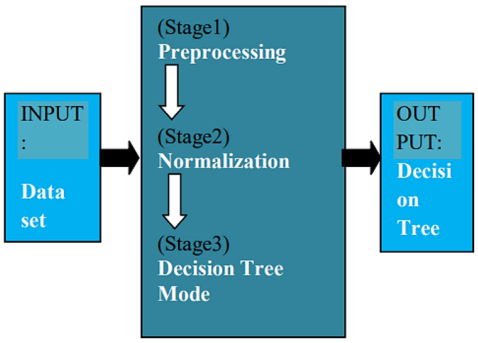

Figure 1. Creating an identifier using machine learning

First of all, we look through the DS input data set and check each St row in the DS dataset and get the n sections to compare using the compute() operation. Then, we call the supplant operation to replace St with a random number j. At the end, the processed data set is returned. And this dataset is used as input for the stage 2 [25].

After processing, the data set may not contain uniformity, as there may be fewer column sections in the sets, which may play a crucial role. Therefore, it is necessary to perform normalization, that is, to reduce the data of the processed data before it

is transferred to the detection algorithm. This process will look like this (7):

$$\begin{gathered} \boldsymbol{Read}: preprocessed\ data\ set \\ DS1, DS2\ =\ split\ dataset(processes\ DS, k\%) \\ A\_set\ =\ normalize(DS1) \\ B\_set\ =\ normalize(DS2) \\ \boldsymbol{Return}\ A\_set, B\_set \end{gathered} \tag{7}$$

The DS1 training data set is randomly selected by k% of the processed DS as the S1 training data set, and the remainder as the S2 training data set, which is (1-k%). The results of A_Set and B_set are obtained using the normalization operation. They will be used as source data in stage 3 [25].

In the process, we build a decision tree using the A_set training dataset as input to the CRT_fn() function, and then we get the output of the B_set control dataset. This process is described in the code (8):

$$\begin{gathered} \boldsymbol{Input}: A\_Set, B\_set \\ Dtmode\ =\ CRT\_fn(A\_set) \\ X\ =\ dt, ode(B\_set) \\ (score\_indicator, t)\ =\ generate() \\ \boldsymbol{Return}\ score\_indicator, computation\_time\ t \end{gathered} \tag{8}$$

The dtmode decision tree mode is achieved using the CRT_fn function by applying the methodology of classification and regression trees. This dtmode mode is applied to B_set to obtain X, and score_indicators and computation_time are generated using the generate() operation [25].

Researchers have shown that the best way to recognize malware in malware analysis is the K-means clustering method. In addition, the detection method is designed so that it can accurately identify an intrusion in a host-based system [24].

K-means clustering consists of three main stages: Stage 1. Data collection; Stage 2. Data preparation; Stage 3. Data analysis.

When data is collected, malware binaries are downloaded from a trusted website and sorted before they are used for an experiment. After the selection, the data was prepared using the selected malware in the control environment and the registry functions were extracted. Finally, the extracted features are combined into a database for data analysis [24].

The technological process of data clustering consists of 7 processes as shown in Figure 2. First, there is a process that extracts a regular registry file from a virtual machine. Then the process continues, which downloads the binary file and injects it into the same virtual machine. After that, the downloaded registry file is extracted from

the VM. In the next step, all files are saved to the database, and during the registry data extraction stage, the data is extracted and prepared. After that, it starts by clustering all the data files using K-means. Finally, at the end of the process, the output will be updated in another table in the same database [24].

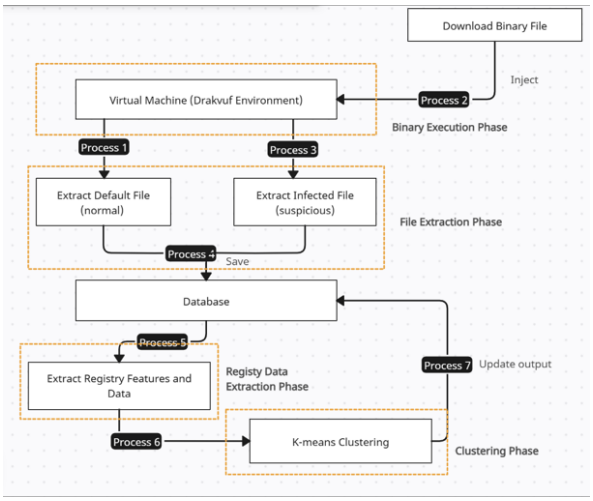

Figure 2. The technological process of data clusterization

All processes are divided into four main phases, and the stages of the detection model are described as follows:

- Stage 1. This is the execution of a binary file. At this stage, the binary file is run on a virtual machine, which is the Drakvuf environment. All actions are then recorded as a log.
- Stage 2. The file extraction stage. Then, at the extraction stage, all the data that represents malware actions. Two types of data are extracted: the default file (normal actions) and the infected file (suspicious actions).
- Stage 3. The stage of extracting registry data. After that, all the collected registry data is extracted and prepared at this stage, since the extracted data is imbalance data.
- Stage 4. Clustering stage. The last phase is the clustering phase in which balanced data is analyzed using a clustering algorithm using K-means to determine whether the data is malicious or not. The Euclidean distance formula is used to measure the distance between the centroid and the data points (9):

$$d(p,q) = d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \quad (9)$$

The effectiveness of K-means clustering detection in the field of intrusion detection

is usually assessed using the measurements shown in Table 2.

Table 2. Measurements of the effectiveness of K-means clustering

Sl.No	Instances	Description
I	The positive truth	Number of malware samples
II	Positive lies	The number of normal samples that were detected accurately
III	Negative lies	The number of malware samples that were falsely detected as common
IV	The negative truth	The number of normal samples falsely recognized as an attack

The detection coefficient (L) is calculated using the formula (10):

$$L = \frac{TP}{(TP+EP)} \quad (10)$$

In addition, malware functions are needed to detect attempts to break into uncontrolled data, since there is no numerical indicator that can be used to calculate the distance between points. The Elbow method is used to determine the K value, as well as the stop point after the result is plotted on the graph. To be precise, the Elbow method is a method of interpreting and verifying the consistency of cluster analysis results, designed to help find the best number of clusters in a dataset, as shown in Figure 3.

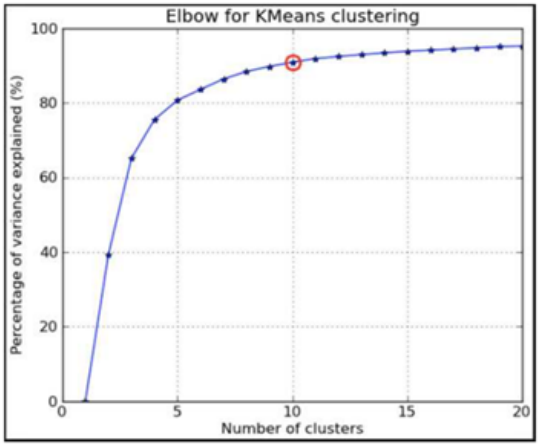

Figure 3. Graph of the Elbow method

Figure 3 shows the growth of the metric with an increase in the number of clusters.

With a small number of clusters (1-5), the increase in the explained variance is quite intense. Starting from about 8-9 clusters, the graph begins to “flatten out”, showing that a further increase in the number of clusters no longer gives a significant increase in the explained variance.

3. The results and discussion

The experimental analysis is performed by evaluating the intrusion detection system. The implementation is done using Python. First, you need to get from some devices that have data collection, and then a communication record is created as an information access point for the system. In this example, the dataset consists of two types: the previous one is a 20% dataset, and the second one is a complete dataset. Each data set consists of a fixed 41 characteristics and a class label. There are 4 types of attacks: 2 (unauthorized access to a local superuser), R2L (unauthorized access from a remote computer), DOS (denial of service), and probing (surveillance and other types of probing). During the experiment, all four types of attacks were detected. Table 3 provides detailed information about the attack subclasses related to these four main types of attacks. A total of 27 types of attack subclasses are listed.

Table 3. Detailed information about attack subclasses

Classification of attacks	Subclasses
DOS(denial of service)	land, back, pod, neptune, teardrop, smurf, mailbomb, apache2
R2L(unauthorized access from a remote computer	imap, multihop, phf, spy, warezmaster, Xlock, warezclient, snmpgetattack, ftp_ write
U2R(unauthorized access to the local superuser)	buffer_overflow, loadmodule, perl, rootkit,guess_passwd
Sensing (observation and other types of sensing)	Ipsweep,nmap, portsweep, satan, saint

Figure 4a shows a comparison of the calculation time (in seconds) for three implementations of the naive Bayes classifier (GaussianNB, BernoulliNB, and MultinomialNB) applied to a 20% sample containing 27 types of attacks. The abscissa axis shows the models used, and the ordinate axis shows the time spent on data processing. As can be seen from Figure 4a, the GaussianNB model shows the longest calculation time (about 4.5–5 seconds), the BernoulliNB model works faster (about 3.5 seconds), and the MultinomialNB model shows the least time (about 2 seconds). Thus, when working with this sample (27 types of attacks), MultinomialNB turns out to be the most efficient in terms of computing speed, while GaussianNB requires significantly more time. The difference may be due to the peculiarities of the data representation

(type of features, their distribution) and algorithmic differences within the implementations of the naive Bayes classifier themselves.

Figure 4b shows a diagram comparing the calculation time (in seconds) for those 3 classifiers (GaussianNB, BernoulliNB, and MultinomialNB) trained on a complete dataset with eight types of attacks. The MultinomialNB model is the most efficient in terms of computing speed on this dataset, whereas the GaussianNB model requires significantly more resources. The difference in performance is explained by how each implementation handles the features and distribution of data (continuous, binary, discrete, etc.).

Figure 4c shows the results of the same three types of naive Bayes classifier applied to a complete dataset containing 27 types of attacks.

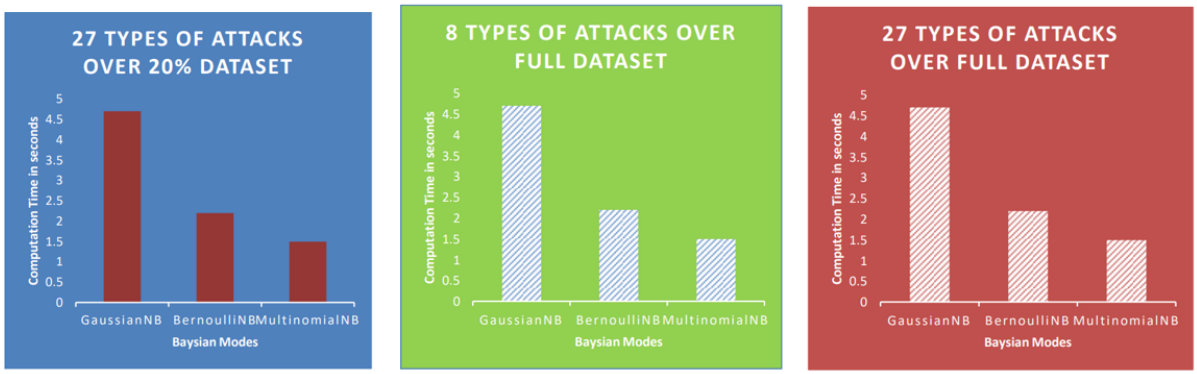

Figure 4. a) Diagram for 27 types of attacks on the entire dataset; b) Diagram for 27 types of attacks on 20% of the dataset; c) Diagram for 8 types of attacks on the entire dataset

As can be seen from Figure 4c, the MultinomialNB model turns out to be the most efficient in terms of computational speed for this complete data set with 27 types of attacks. The difference may be related to the specifics of the distribution of features and the features of each implementation of the naive Bayesian algorithm, where GaussianNB is focused primarily on continuous features, whereas BernoulliNB and MultinomialNB are optimized for binary or discrete data. Table 4 shows the total number of instances.

Table 4. Total number of copiess

Type of attack	Total number of training examples	Total number of test instances
Dos	1807	3168
U2R	70	105
U2R	216	2338
Probing	1391	2579
Normal	7500	10000
Total	10984	18200

The results of the experiment are compared using score_indicator and computation_time. To ensure that all types of attacks are accounted for, the data set is randomly divided into training and test datasets. Consequently, as a result of this, three models are formed, such as BernoulliNB, MultinomialNB, and GaussianNB in the naive Bayesian language.

The study demonstrates that the use of machine learning methods significantly increases the effectiveness of protecting confidential information systems. The results obtained confirm the working hypothesis that modern algorithms make it possible to automate threat diagnostics and detect them at an early stage with high accuracy. Such conclusions are consistent with previous studies, which emphasized the importance of systematic collection and analysis of statistical data for detecting cyber threats.

4. Conclusion

This study describes approaches to applying machine learning methods to ensure information security, which will make it possible to protect personal data in medical systems. The results obtained indicate a significant increase in the accuracy and timely response to cyber attacks and information leaks. In conclusion, the study highlights the significant potential of machine learning in protecting personal data in medicine. The results of the study open up promising opportunities for improving the quality of cybersecurity in the medical information system, which contributes to the further study and application of machine learning methods for data protection.

The study focuses not only on the potential of these methods to protect personal information, but also on their importance as a tool to counter threats and improve data security in medical information systems. The introduction of such technologies into the medical field promises to improve the quality of information systems and open up new opportunities for data protection, which will ultimately enhance the security of information in healthcare.

Financing. This research has been/was/is funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19679525).

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

References

1. Haleem, A., Javaid, M., Singh, R. P., & Suman, R. (2022). Medical 4.0 technologies for healthcare: Features, capabilities, and applications. Internet of Things and Cyber-Physical Systems, 2, 12-30.
2. Zhetpissov, S., Mussabekova, N., Talipova, Z., Dubovitskay, O., & Alibayeva, G. (2024). Vulnerability of personal data of Kazakhstani citizens and the need to implement the European experience. Rivista di studi sulla sostenibilità: 9, 2, 2024, 305-323.
3. Omurchiyeva, E., & Saudabayeva, D. (2024). Problemy primeneniya zakonodatelstva respubliki kazakhstan o personalnykh dannykh: riski dlya rabotodatelya. Scientific and legal journal «Bulletin of the Institute of Legislation and Legal Information of the Republic of Kazakhstan», 2(77).
4. Tarasov, D. S., Solodukhina, E. A., Marchenkova, V. E., & Skorikov, V. V. (2023). Protivodeystvie kiberterrorizmu v Belarusi, Kazakhstane i Azerbaydzhane. Postsovetskie issledovaniya, 6(5), 531-539.
5. Silantyev, I. O., & Anikin, I. V. (2023). Vyyavlenie utechek konfidentsialnoy informatsii v informatsionnykh sistemakh. Mezhdunarodnyy nauchno-issledovatelskiy zhurnal, (7 (133)).
6. Казанцев, А. А. (2023). Kharakteristika faktorov dlya sravneniya razlichnykh sposobov prokladki liniy elektroperedach v kontekste ikh sravneniya s uchetom elektrotekhnicheskoy bezopasnosti i ekonomicheskoy effektivnosti. Состав редакционной коллегии и организационного комитета.
7. Mazhar, T., Irfan, H. M., Khan, S., Haq, I., Ullah, I., Iqbal, M., & Hamam, H. (2023). Analysis of cyber security attacks and its solutions for the smart grid using machine learning and blockchain methods. Future Internet, 15(2), 83.
8. Ahsan, M., Nygard, K. E., Gomes, R., Chowdhury, M. M., Rifat, N., & Connolly, J. F. (2022). Cybersecurity threats and their mitigation approaches using Machine Learning—A Review. Journal of Cybersecurity and Privacy, 2(3), 527-555.
9. Mokhtari, S., Abbaspour, A., Yen, K. K., & Sargolzaei, A. (2021). A machine learning approach for anomaly detection in industrial control systems based on measurement data. Electronics, 10(4), 407.
10. Wen, S. F., Shukla, A., & Katt, B. (2025). Artificial intelligence for system security assurance: A systematic literature review. International Journal of Information Security, 24(1), 1-42.

11. Bhosale, K. S., Nenova, M., & Iliev, G. (2021, September). A study of cyber attacks: In the healthcare sector. In 2021 sixth junior conference on lighting (lighting) (pp. 1-6). IEEE.
12. Manoharan, A., & Sarker, M. (2023). Revolutionizing Cybersecurity: Unleashing the Power of Artificial Intelligence and Machine Learning for Next-Generation Threat Detection. DOI: https://www. doi. org/10.56726/IRJMETS32644, 1.
13. Kotenko, I. V., Saenko, I. B., Lauta, O. S., Vasilyev, N. A., & Sadovnikov, V. E. (2024). Ataki i metody zashchity v sistemakh mashinnogo obucheniya: analiz sovremennykh issledovaniy. Voprosy kiberbezopasnosti, (1), 59.
14. Stamp, M. (2022). Introduction to machine learning with applications in information security. Chapman and Hall/CRC.
15. Zarour, Mohammad, et al. “Ensuring data integrity of healthcare information in the era of digital health.” Healthcare technology letters 8.3 (2021): 66-77.
16. Zhetpisov, S. K., Alibayeva, G. A., & Dubovitskaya, O. B. (2023). Zashchita personalnykh dannykh v epokhu tsifrovizatsii: Konstitutsionno-pravovoy aspekt. Scientific and legal journal «Bulletin of the Institute of Legislation and Legal Information of the Republic of Kazakhstan», 3(74).
17. Sarker, I. H. (2023). Machine learning for intelligent data analysis and automation in cybersecurity: current and future prospects. Annals of Data Science, 10(6), 1473-1498.
18. Mokhtari, S., Abbaspour, A., Yen, K. K., & Sargolzaei, A. (2021). A machine learning approach for anomaly detection in industrial control systems based on measurement data. Electronics, 10(4), 407.
19. Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). Machine learning and deep learning approaches for cybersecurity: A review. IEEE Access, 10, 19572-19585.
20. Zhang, C., Jia, D., Wang, L., Wang, W., Liu, F., & Yang, A. (2022). Comparative research on network intrusion detection methods based on machine learning. Computers & Security, 121, 102861.
21. Chinedu, P. U., Nwankwo, W., Masajuwa, F. U., & Imoisi, S. (2021). Cybercrime detection and prevention efforts in the last decade: an overview of the possibilities of machine learning models. Rigeo, 11(7).
22. Hesham, M., Essam, M., Bahaa, M., Mohamed, A., Gomaa, M., Hany, M., & Elsersy, W. (2024, July). Evaluating Predictive Models in Cybersecurity: A Comparative Analysis of Machine and Deep Learning Techniques for Threat Detection. In 2024 Intelligent Methods, Systems, and Applications (IMSA) (pp. 33-38). IEEE.
23. Jiang L., Zhang H., Cai Z. A novel bayes model: Hidden naive bayesю. IEEE Transactions on knowledge and data engineering. – 2008. – Т. 21. – №. 10. – С. 1361-

1371.

24. Martínez Torres J., Iglesias Comesaña C., García-Nieto P. J. Machine learning techniques applied to cybersecurity. International Journal of Machine Learning and Cybernetics. – 2019. – T. 10. – №. 10. – C. 2823-2836.
25. Shilpashree S. et al. Decision tree: A machine learning for intrusion detection. Int. J. Innov. Technol. Explor. Eng. – 2019. – T. 8. – №. 5.

Information about authors

Uvaliyeva Indira Makhmutovna - PhD, D. Serikbayev East Kazakhstan Technical University.
e-mail: indirauvalieva@gmail.com
ORCID: 0000-0002-2117-5390

Aiman Toktarbayeva Bolatkyzy - Master of Computer Science, D. Serikbayev East Kazakhstan Technical University,
e-mail: ayman_toktarbayeva@mail.ru
ORCID: 0009-0007-8404-2193