ISSN: 2182-2069 / E-ISSN: 2182-2077

Intelligent Information Security System for Language and History Education Using Machine Learning-based Intrusion Detection Algorithm

Dilrabo Bakhronova[1*], Ma'mura Narziyeva[2], Nilufar Yuldosheva[3], Umeda Zayniyeva[4], Jambul Yusupov[5], Bakhtiyor Uralov[6], Ibrokhim Sapaev[7], and Nodir Khikmatov[8]

[1*]Professor, Head of Department of Applied Sciences of Spanish Language, Doctor of Philology Uzbekistan State World Languages University, Uzbekistan. dbaxronova@uzswlu.uz, https://orcid.org/0000-0002-2012-7426

[2]Doctor of Philosophy in Philology (PhD), Associate Professor of Samarkand State University named after Sharof Rashidov, Uzbekistan. narziyevamamura164@gmail.com, https://orcid.org/0000-0003-0542-8826

[3]DSc, Professor, Department of Philology, Karshi State University, Uzbekistan. nilu.75@mail.ru, https://orcid.org/0000-0002-4920-7534

[4]Teacher, Samarkand State of Foreign Languages Institute, Uzbekistan. zayniyevaumeda@5gmail.com, https://orcid.org/0009-0009-5212-6231

[5]Kimyo International University in Tashkent, Tashkent, Uzbekistan. j.yusupov@kiut.uz, https://orcid.org/0000-0003-1758-6805

[6]Tashkent Institute of Irrigation and Agricultural Mechanization Engineers National Research University, Tashkent, Uzbekistan. bakhtior@mail.ru, https://orcid.org/0000-0001-9371-5563

[7]Head of the Department Physics and Chemistry, "Tashkent Institute of Irrigation and Agricultural Mechanization Engineers" National Research University, Tashkent, Uzbekistan; Scientific Researcher, University of Tashkent for Applied Sciences, Str. Gavhar 1, Tashkent, Uzbekistan; Western Caspian University, Scientific researcher, Baku, Azerbaijan. sapaevibrokhim@gmail.com, https://orcid.org/0000-0003-2365-1554

[8]Senior Lecturer, Department of Foreign Economic Activities, Tashkent State University of Oriental Studies, Uzbekistan. xnodir32@gmail.com, https://orcid.org/0000-0001-9987-0011

Received: December 29, 2024; Revised: January 30, 2025; Accepted: February 13, 2025; Published: February 28, 2025

Abstract

The recent spread of the internet has led to a significant increase in the use of personal online banking and e-commerce. Operations and marketing inside organizations, government, and banking sectors are experiencing tremendous growth, primarily through online shopping malls and web pages. There is a growing prevalence of cyber-attacks, including sophisticated strikes, unauthorized cyber intrusions, and the compromise of electronic data. Defensive measures, operational exercises, and Information Security (IS) incidents are not implemented effectively. This study uses Machine Learning to establish an Intrusion Detection System (IDS) for language and history education. IS

Journal of Internet Services and Information Security (JISIS), volume: 15, number: 1 (February), pp. 520-529.
DOI: *10.58346/JISIS.2025.I1.034*

*Corresponding author: Professor, Head of Department of Applied Sciences of Spanish Language, Doctor of Philology Uzbekistan State World Languages University, Uzbekistan.

training is being implemented for affiliated staff and educational subjects utilizing the Information Technology (IT) infrastructure to address IS incidents resulting from fraudulent emails and file attachments that commonly arise in government organizations and private enterprises. This approach could preempt IS incidents resulting from phishing emails and mitigate economic losses by averting data loss or system stoppage.

Keywords: Information Security, Language Education, Machine Learning, Intrusion Detection.

1 Introduction

The proliferation of the Internet has led to a significant increase in personal online banking and e-commerce utilization (Bădîrcea et al., 2021). Services and advertising within organizations, administration, and institutions are experiencing tremendous growth, primarily through online shopping malls and websites (Faisal & Dharmaraj, 2024). In these conditions, there is a rise in numerous illicit activities, such as the unlawful acquisition of financial data, such as private data and credit card details, corporate marketing strategies, new product data, significant internet service disruptions, and service incapacitation. Illegal activities refer to hacking or disseminating worm viruses aimed at unidentified individuals for language and history education (Vasani et al., 2023).

There is a proliferation of cyber hacking incidents, private data breaches, and security violations, with attack methodologies becoming increasingly sophisticated due to advancements in various technology fields for language and history education (Kim et al., 2019). The monthly report from 2020 indicates 19.5k instances of harmful code damage and 18.7k cases of hacking incidents reported and processed (Johnson et al., 2020).

Concerning several local and international security incidents (Reshmi, 2021), most persist as harmful code infections and phishing attacks via website accessibility (Yang, 2024). As numerous reports by media outlets and national entities regarding private data breaches are rising, and the severity and scope of violations in mobile internet and cloud environments are progressively escalating, the national economy, safety, and individual privacy are broadly jeopardized for language and history education (Aravindhan, 2023).

The generation and dissemination of malicious code and supplementary attacks have recently become exceedingly facile, and the variants are proliferating rapidly (Chen et al., 2025). Owing to the perilous access to websites and associated dangerous code infections, there is a rise in computer-related crimes, including information breaches resulting from fraudulent hacking or industrial espionage, as well as the leakage of corporate secrets by workers and consumer data (Armstrong & Tanaka, 2025).

When infringement incidents occur, the social milieu tends to place more tremendous guilt on the victims than the perpetrators for language and history education (Vasquez & Mendoza, 2024). As infringement incidents are attributed to the victims, the endeavor to achieve flawless prevention intensifies by applying and advancing diverse safety measures. Several Information Security (IS) (Rohan et al., 2023) measures are implemented to safeguard against these illicit activities, including intrusion-blocking systems to counter criminals, detection systems for intrusions, and antivirus solutions (Uvarajan, 2024).

Defenses and patches against illicit activities are not collaboratively shared but autonomously managed by particular organizations and enterprises (Steen & Hansen, 2024). Infringement incidents cannot be entirely averted with IS measures (Blaber & Rafiq, 2023). IS measures serve solely as a defense against specific threats. In this regard, safety measures must facilitate access routes permitted

within companies. Without an adequate online infrastructure for educating individuals about data security, significant damages arise due to ambiguous accountability for IS incidents.

There is an increasing need to create and implement a combined, complete educational system for responding to infringement incidents that can effectively address illegal activities with minimal staff. This study proposes the Intrusion Detection System (IDS) (Muhammad et al., 2023) of an IS training and education infrastructure utilizing the Internet for language and history education. Educational materials will be implemented for connected personnel and training participants associated with this network to address IS incidents resulting from fraudulent emails and document attachments that commonly occur in government organizations and private enterprises (Sapna & Singh, 2022).

2 Background

Intrusion detection equipment is classified into three primary groups: pattern-matching techniques, conventional Machine Learning (ML) (Ahmad et al., 2021) approaches, and Deep Learning (DL) (Bakhsh et al., 2023) methodologies. Initially, individuals mainly employed pattern-matching techniques for identifying breaches. The pattern-matching mechanism is the fundamental algorithm of an IDS that relies on matching features. The majority of methods have been evaluated for potential application in the past. Experiments demonstrate that the enhanced method expedites the matching speed and exhibits commendable temporal efficiency. A comparison is conducted among the naive strategies to see which is the most effective for patterns and detection of intrusions.

Pcap records have been utilized as datasets to evaluate the algorithm's effectiveness by analyzing their running durations. These conventional pattern identification algorithms have significant deficiencies that hinder their IDS. Pursuing a successful algorithm that achieves high efficacy and minimal false favorable rates remains the focus of ongoing research. The advancement of Artificial Intelligence (AI) (Mahbooba et al., 2021) has rendered the utilization of intelligent algorithms for IDS a prominent area of study.

ML-based traffic detection of anomalies approaches have attained significant success in language and history education. The researchers introduce an innovative approach to choosing and categorizing features utilizing Support Vector Machine (SVM) technology. Experimental findings on the NSL-KDD Cup 99 malware detection database indicated that the categorizing accuracy of this technique, using all training features, attained 89%. The researchers integrate k-means clustering with the K-Nearest Neighbours (KNN) classification.

The experimental findings on the NSL-KDD dataset indicate that this strategy significantly enhances the accuracy of the KNN classification. The authors present a novel system that integrates abuse and detection of anomalies by applying the random forests technique. Experimental findings indicate that the hybrid system achieves an overall detection rate of 91.2% and an overall false positive rate of 3%. The NSL-KDD dataset's efficiency is assessed using an Artificial Neural Network (ANN). The detection rates achieved are 80.5% for IDS and 75.3% for attack-type categorization using the NSL-KDD database. A Decision Tree (DT)--based IDS is presented.

The results of feature selection experiments utilizing the Correlation-based Feature Selection (CFS) subset assessment approach indicate that the decision tree-based IDS exhibits superior accuracy. ML techniques have been presented and successfully implemented for an IDS. These techniques necessitate extensive preprocessing and intricate feature engineering of language and history education traffic information. Resolving the extensive invasion data classification issue with ML techniques is unfeasible.

3 System Architecture and Implementation

To establish a system for IS learning for individuals, this article outlines a configuration comprising an educational mail computer, portable computers, an agent structure, tracking and reporting Website Server (WS), and a WS for online threat learning. Concerning the education computer, the list has been compiled, and the mail transmission system infrastructure has been set up to evaluate the subjects utilizing the list for language and history education. A high-performance computer was established for continuous surveillance by the manager, capable of managing several concurrent traffic and scripts during the instruction on developing hosts. The database was established using MySQL (Figure 1).

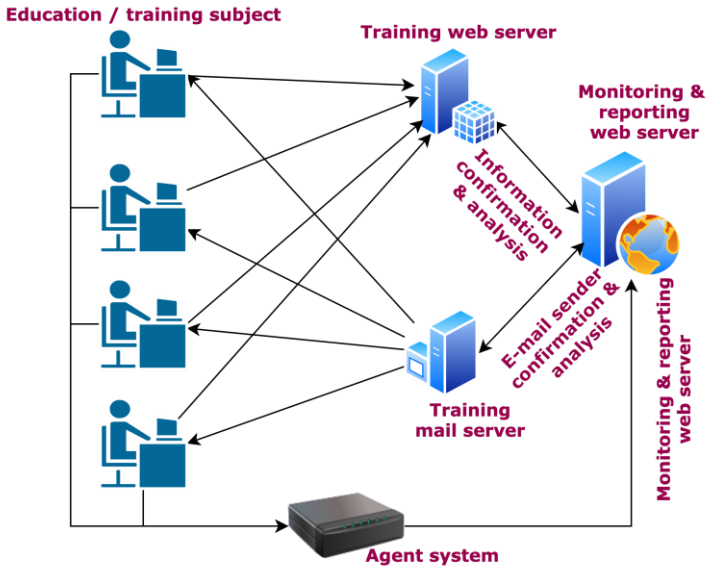

Figure 1: Designing of the System Elements

The educational mail WS was created for IS learning and designed to oversee the delivery of harmful and educational emails that could pose virtual IS threats to educational and training entities for language and history education. It was established that a mailing list of educational and training topics and the processes for sending and receiving educational correspondence could be determined. A desktop computer was utilized so that the educational and training individuals who got phishing messages from the instructional email WS could access the emails and get IS instructions.

An agent software was set up on a desktop computer and designed with functionalities for host monitoring by educational and training topics who discovered forged emails and for transferring their situational management capabilities regarding test and educational services. The tracking and reporting WS was designed to assess the dissemination outcomes of the agent system's situational handling capabilities and to disseminate data and information.

The subject's email delivery, reception, program execution, agent actions, and information gathering could be discerned. A WS for simulated IS threat learning was created, enabling training participants to open dangerous emails, view virtual safety threat Universal Resource Locators (URLs), or identify fraudulent messages on home computers following the detection of execution of code for language and history education. The online IS threat teaching WS and the surveillance and reporting WS can be interconnected to evaluate outcomes by recognizing the user accessing the virtual safety threat teaching WS.

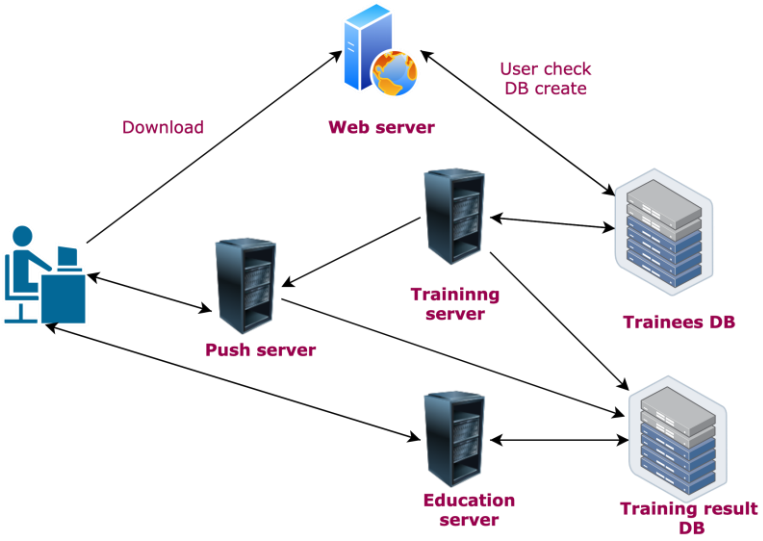

Figure 2: System flowchart

An education statistical evaluation solution was created by developing Test-Code utilizing JavaScript and OpenRelay. The browser-based exploit approach was examined, and the web-based Java Applet technique was implemented. This system generates emails containing user-activated links and triggers a pop-up warning window regarding IS policy within the WS upon link activation by the consumer. The execution outcomes and evaluation of information via simulated ML were assessed via a report upon completion, allowing for the classification of particular indexes and banks of questions at each stage for language and history education. Based on educational outcomes, supplementary IS training will be implemented through enhanced educational programs and the improvement of IS awareness courses (Figure 2).

This study presents the development of an Internet-based cyber IS training and instruction platform. The agent system facilitates the capability to manage learning for the mail WS's ability to dispatch a malicious email, the personal machine of the educational subject, the circumstances surrounding the harmful email, and the surveillance and reporting of the WS for virtual safety. The study sets up the machine as an internet WS for threat training. The instruction for the email WS executes processing activities and transmits harmful emails for language and history education. At this juncture, it is feasible to establish supplementary test centers to enhance the management capabilities of the agent.

Train on opening the malicious email that was sent to the machine of the education topic, the functioning of a website containing malware files, training, the web host for online threat instruction, the agent's execution of information-gathering capacities administration work, regarding the deletion of unconfirmed, unregistered illegal incoming mail. Working together with the WS for tracking and reporting, it is used to facilitate the analysis and validation of the permissions risk assessment for the WS concerning virtual protection.

The WS used for tracking and reporting is configured for training to address language and history education issues, including email deletion, email availability, execution of code, obtaining threat URLs in virtual safety, and incident reporting. It is designed to evaluate the outcomes of the execution. It creates a malicious email configuring the Test-code generator with links that induce user interaction according to the clicked content. Upon a user clicking a link, a warning box is programmed to appear for IS policy via the internet browser. To collect data to ensure that no agent-level behavior occurs during

training operations and to determine whether to disable or remove unconfirmed, undefined illegal emails from the agent structure, the web host will track and notify accordingly. Solutions are always required in the operating system that is being used, and the education personnel must be able to verify this. Educational and training topics aim to ascertain whether visiting a website linked to attachments in an offensive email occurs.

4 Results

To objectively assess the accuracy and distinction of the proposed system, the research contrasts it with several relevant works that have been presented. The authors suggest an ML methodology for identifying breaches via Recurrent Neural Networks (RNNs). Compared to conventional classification techniques, like J48, Naive Bayesian (NB), and Random Forest (RF), Random Tree (RT), Multiple Layer Perception (MLP) the performance demonstrates superior accuracy and rates of identification while maintaining a low False Positive Rate (FPR), particularly in the context of multiple classifiers on the NSL-KDD dataset. The authors develop a DNN algorithm for an IDS and train it using the Database. Experimental findings validate that the ML methodology demonstrates significant potential for flow-based detection of anomalies in Software-Defined Networking (SDN) setups. In, the authors advocate for employing a conventional ML technique, Recurrent Neural Network (RNN), Convolutional Neural Networks (CNN), Deep Belief Network (DBN), Long Short Term Memory (LSTM), Bidirectional LSTM (BLSTM), Bayesian Attack Trees (BAT) to identify cyber breaches. The experimental findings demonstrate that the accuracy of this IDS system surpasses that of systems utilizing classic ML techniques and innovative ML approaches in multi-class categorization. These studies use the identical database NSL-KDD to categorize language and history education network traffic. These studies are not only fresh and very relevant to IS monitoring, but they demonstrate exceptional accuracy. Figure 3 and Figure 4 illustrate the comparative results of these studies on the NSL-KDD database.

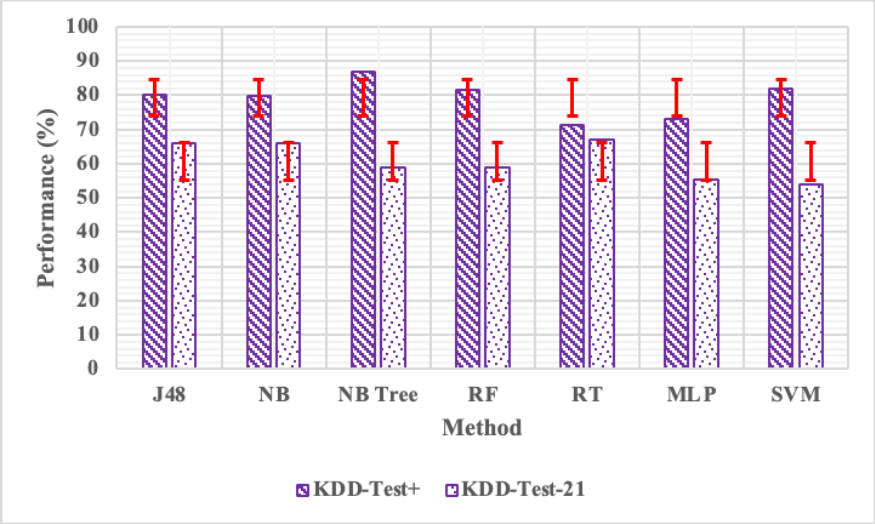

Figure 3: ML Model Analysis

Figure 3 and Figure 4 illustrate that the suggested model outperforms other approaches in accuracy, achieving 89.4% and 75.3% on the KDDTest+ and KDDTest-21 datasets, respectively. In contrast to the algorithm, the authors utilize conventional ML techniques to identify anomalous traffic. In other words, it is necessary to manually construct traffic characteristics and finalize the collection and choice of network activity before model retraining for language and history education. The proposed model

utilizes the gathered traffic as its primary input. The attention process extracts essential features from the results generated by the model. Results from experiments indicate that the proposed system can autonomously extract characteristics through end-to-end learning, yielding superior classification outcomes compared to manual design techniques.

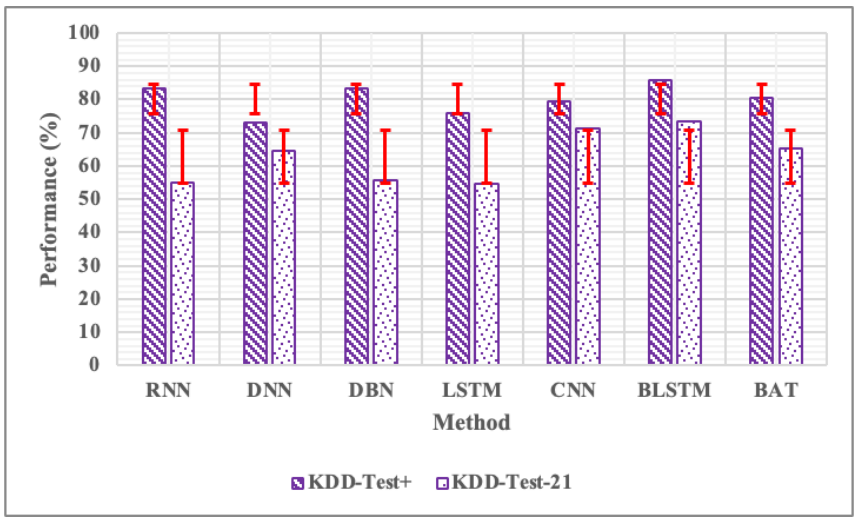

Figure 4: DL Model Analysis

The research evaluates current studies utilizing DL models for irregular traffic identification. Figure 4 illustrates that the suggested model got superior results on the KDDTest+ and KDDTest-21 assessment sets. In the KDDTest + set, the precision of the proposed system exceeds that of CNN and RNN by 3.42% and 1.85%, correspondingly. In the KDDTest-21 dataset, the precision of the suggested system surpasses that of CNN and RNN by 5.41% and 6.5%, respectively. The proposed system exhibits superior accuracy compared to CNN, as CNN is primarily optimized for image information interpretation. CNN has a static convolution core incapable of capturing extended contextual details, which hinders the feature collection of time series datasets. The suggested networks surpass RNN, LSTM, and BLSTM by integrating an attention strategy that effectively captures essential features and acquires additional contextual data for language and history education.

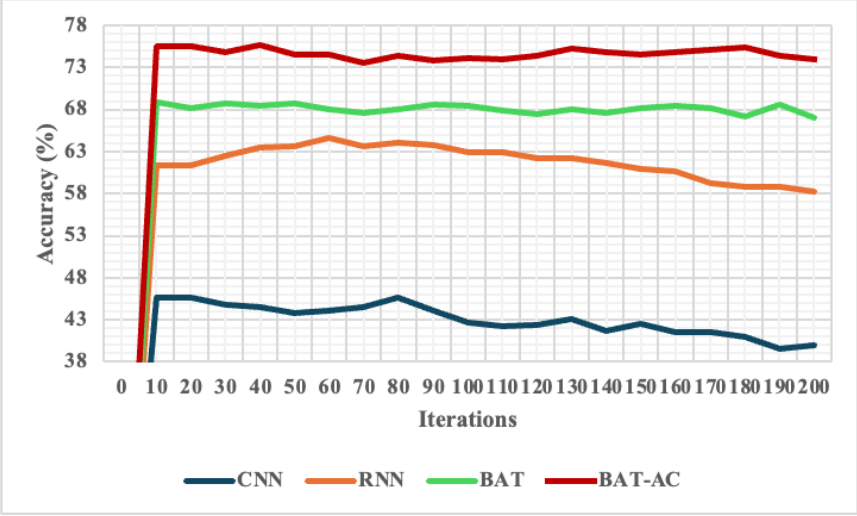

Figure 5: Accuracy Analysis

With the increase of iterations, the precision of each model exhibits a general rising trajectory. Figure 5 illustrates that the precision rate of the test set with the suggested algorithm is not only the most rapid but exhibits minimal fluctuation following 20 iterations. The precision of the suggested system stays unaltered primarily. With increased repetitions, the suggested model's precision progressively improves, ultimately attaining an optimal form. The suggested model has greater precision than the suggested model due to its ability to capture broad data, demonstrating the benefits of multiple convolutional layering. The RNN model exhibits minor variations in the precision of the iterative procedure for language and history education. The RNN approach exhibits a more rapid improvement, although it demonstrates inferior accuracy compared to the BAT and suggested models. The CNN method has a diminishing rate of improvement and shows the poorest efficiency among all algorithms. The suggested system achieves a precision of 89.7% in identifying time series information, proving its efficacy as an IDS.

5 Conclusion

The proliferation of the Internet has led to a significant rise in personal online banking and e-commerce utilization. There is an incessant occurrence of cyber hacking, sensitive data breaches, and IS violations, with attack methodologies becoming increasingly sophisticated due to advancements in numerous technology fields. In these situations, there is a rise in multiple illicit activities, including the unlawful acquisition of monetary data such as personal data and credit card numbers, corporate marketing strategies, new product data, significant internet service disruptions, and service incapacitation.

The generation and dissemination of malicious software and supplementary attacks have recently become exceedingly facile, and the variants are proliferating rapidly. Owing to the perilous access to websites and associated dangerous code infections, there is a rise in computer-related crimes, such as data thefts resulting from fraudulent hacking or industrial spying, involving the disclosure of corporate secrets by workers and the leakage of customer data. Without a successful online platform for educating about data safety, significant damages arise due to ambiguous accountability for IS incidents.

This study presents the development of a cyber data safety training and education platform utilizing the internet for language and history education. Data literacy was implemented using this system aimed at affiliated staff or training topics to address IS incidents resulting from malicious emails and attached documents that frequently affect government agencies and private businesses using ML. Data on the harmful attachment files associated with training and educational subjects was recognized and examined on the virtual IS threat training WS for language and history education. The present research assessed the impact of ML and instruction on mitigating the harm caused by hackers and viruses by analyzing implementation results and providing relevant data and information.

References

[1] Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, *32*(1), e4150. https://doi.org/10.1002/ett.4150

[2] Aravindhan, S. (2023). A Flexible Structure's Active Vibration Suppression Using Smart Materials. *Association Journal of Interdisciplinary Technics in Engineering Mechanics*, *1*(1), 55-61.

[3] Armstrong, D., & Tanaka, Y. (2025). Boosting Telemedicine Healthcare Assessment Using Internet of Things and Artificial Intelligence for Transforming Alzheimer's Detection. *Global Journal of Medical Terminology Research and Informatics*, *2*(1), 8-14.

[4] Bădîrcea, R. M., Manta, A. G., Florea, N. M., Popescu, J., Manta, F. L., & Puiu, S. (2021). E-commerce and the Factors Affecting its Development in the Age of Digital Technology: Empirical Evidence at EU–27 level. *Sustainability*, *14*(1), 101. https://doi.org/10.3390/su14010101

[5] Bakhsh, S. A., Khan, M. A., Ahmed, F., Alshehri, M. S., Ali, H., & Ahmad, J. (2023). Enhancing IoT network security through deep learning-powered Intrusion Detection System. *Internet of Things*, *24*, 100936. https://doi.org/10.1016/j.iot.2023.100936

[6] Blaber, M., & Rafiq, H. (2023). What Makes Agile Powerful to Boost Innovation for the Larger Organizations. *Global Perspectives in Management, 1*(1), 17-31.

[7] Chen, C., Liu, J., Tan, H., Li, X., Wang, K. I. K., Li, P., ... & Dou, D. (2025). Trustworthy federated learning: Privacy, security, and beyond. *Knowledge and Information Systems*, *67*(3), 2321-2356. https://doi.org/10.1007/s10115-024-02285-2

[8] Faisal, T., & Dharmaraj, A. (2024). Perception of E-learning among the Students Studying in Higher Education Institutions with Reference to Kerala. *Indian Journal of Information Sources and Services, 14*(4), 66–72. https://doi.org/10.51983/ijiss-2024.14.4.11

[9] Johnson, C., Khadka, B., Basnet, R.B., & Doleck, T. (2020). Towards Detecting and Classifying Malicious URLs Using Deep Learning. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 11*(4), 31-48.

[10] Kim, K., Ko, E., Kim, J., & Yi, J. H. (2019). Intelligent Malware Detection Based on Hybrid Learning of API and ACG on Android. *Journal of Internet Services and Information Security*, *9*(4), 39-48.

[11] Mahbooba, B., Timilsina, M., Sahal, R., & Serrano, M. (2021). Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. *Complexity*, *2021*(1), 6634811. https://doi.org/10.1155/2021/6634811

[12] Muhammad, A. R., Sukarno, P., & Wardana, A. A. (2023). Integrated security information and event management (siem) with intrusion detection system (ids) for live analysis based on machine learning. *Procedia Computer Science*, *217*, 1406-1415. https://doi.org/10.1016/j.procs.2022.12.339

[13] Reshmi, T. R. (2021). Information security breaches due to ransomware attacks-a systematic literature review. *International Journal of Information Management Data Insights, 1*(2), 100013. https://doi.org/10.1016/j.jjimei.2021.100013

[14] Rohan, R., Pal, D., Hautamäki, J., Funilkul, S., Chutimaskul, W., & Thapliyal, H. (2023). A systematic literature review of cybersecurity scales assessing information security awareness. *Heliyon*, *9*(3). https://doi.org/10.1016/j.heliyon.2023.e14234

[15] Sapna, P. F., & Singh, R. L. R. (2022). Electrical Load Forecasting Techniques & Methods: An Overview. *International Journal of Advances in Engineering and Emerging Technology*, *13*(2), 254-262.

[16] Steen, R., & Hansen, T. B. (2024). Collaborative defense in the Arctic: Strengthening Norway's oil sector resilience through knowledge sharing and vigilance against drone threats. *Risk, Hazards & Crisis in Public Policy*. https://doi.org/10.1002/rhc3.12302

[17] Uvarajan, K. P. (2024). Advances in quantum computing: Implications for engineering and science. *Innovative Reviews in Engineering and Science*, *1*(1), 21-24. https://doi.org/10.31838/INES/01.01.05

[18] Vasani, V., Bairwa, A. K., Joshi, S., Pljonkin, A., Kaur, M., & Amoon, M. (2023). Comprehensive analysis of advanced techniques and vital tools for detecting malware intrusion. *Electronics*, *12*(20), 4299. https://doi.org/10.3390/electronics12204299

[19] Vasquez, E., & Mendoza, R. (2024). Membrane-Based Separation Methods for Effective Contaminant Removal in Wastewater and Water Systems. *Engineering Perspectives in Filtration and Separation, 1*(1), 21-27.
[20] Yang, Z. (2024). The Impact of Environmental Assessment of Green Innovation on Corporate Performance and an Empirical Study. *Natural and Engineering Sciences, 9*(2), 94-109. https://doi.org/10.28978/nesciences.1569137

Authors Biography

Dilrabo Bakhronova is a Doctor of Philology and a Professor who leads the Department of Applied Sciences of the Spanish Language at Uzbekistan State World Languages University. Her research primarily focuses on philology and applied linguistics.

Ma'mura Narziyeva is an Associate Professor with a PhD in Philology. She is affiliated with Samarkand State University and is known for her work in linguistic research.

Nilufar Yuldosheva is a Doctor of Science and Professor in the Department of Philology at Karshi State University. She has made significant contributions to philological research and academia.

Umeda Zayniyeva is an educator at the Samarkand State Institute of Foreign Languages. Her work revolves around language studies and education.

Jambul Yusupov is a faculty member at Kimyo International University in Tashkent. He is actively involved in research and has contributed to various academic projects.

Bakhtiyor Uralov is a researcher and faculty member at the Tashkent Institute of Irrigation and Agricultural Mechanization Engineers. His expertise lies in agricultural engineering and related fields.

Ibrokhim Sapaev is a distinguished researcher with a leadership role in the "Physics and Chemistry" department at Tashkent Institute of Irrigation and Agricultural Mechanization Engineers. He also holds research positions at the University of Tashkent for Applied Sciences and Western Caspian University.

Nodir Khikmatov is a Senior Lecturer specializing in foreign economic activities at Tashkent State University of Oriental Studies. His work focuses on international economic relations and academic research.