IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

AI-Powered Intrusion Detection System for Intelligent Network Security Using Machine Learning - Detection of Network and IoT Attacks Using NSL-KDD and Bot-IoT Datasets

Mr. K. Pazhanivel[1], M. Mohamed Anwardeen[2], K. Sakthivel[3], U. K. Sri Niketan[4], C. Yogeshwaran[5]
Assistant Professor, Department of Computer Science and Engineering[1]
Student, Department of Computer Science and Engineering[2-5]
Anjalai Ammal Mahalingam Engineering College,Kovilvenni, Thiruvarur, Tamil Nadu, India
Palanicse86@gmail.com[1], anwardeen1289@gmail.com[2]
sv249946@gmail.com[3], sriniketanuk@gmail.com[4], cyogeshff@gmail.com[5]

Abstract: *The rapid expansion of networked systems and Internet of Things (IoT) devices has significantly increased the risk of cyber-attacks, making network intrusion detection an essential component of modern cybersecurity. Traditional intrusion detection systems often rely on signature-based tech- niques, which are ineffective against unknown and evolving attacks. This paper proposes an AI-powered intrusion detection system that utilizes machine learning techniques to improve the accuracy and efficiency of attack detection in intelligent network environments. The proposed system employs multiple machine learning algorithms, including Random Forest, AdaBoost, Light Gradient Boosting Machine (LightGBM), and Multi-Layer Per- ceptron (MLP), to classify network traffic as normal or malicious. The system is trained and evaluated using both the NSL- KDD dataset for traditional network intrusions and the Bot-IoT dataset for modern IoT-based attacks, enabling comprehensive analysis across diverse attack scenarios. Experimental results demonstrate that the proposed approach achieves high detection accuracy and improved classification performance compared to traditional methods, while reducing false positive rates. The main contribution of this work lies in the integration of multiple ma- chine learning models and datasets within a unified framework, providing enhanced intrusion detection capability and improved network security for intelligent and dynamic environments.*

Keywords: *Internet of Things*

I. INTRODUCTION

With the rapid growth of internet technologies and con- nected devices, network security has become a critical concern for organizations and individuals. Modern networks, especially those involving Internet of Things (IoT) devices, generate a large volume of data and are increasingly vulnerable to various cyber-attacks such as Denial of Service (DoS), prob- ing, and unauthorized access. Traditional intrusion detection systems rely mainly on signature-based techniques, which are ineffective against new and evolving attacks. Therefore, intelligent and adaptive security mechanisms are required to ensure reliable network protection.

Intrusion Detection Systems (IDS) play an important role in monitoring network traffic and identifying malicious activities. However, conventional IDS approaches often suffer from high false alarm rates and limited capability in detecting unknown attack patterns. Machine Learning (ML) techniques provide an efficient solution by learning patterns from historical data and automatically classifying network activities as normal or malicious. ML-based IDS can improve detection accuracy and adapt to dynamic network environments.

DOI: 10.48175/IJARSCT-33227

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, April 2026 Impact Factor: 8.2

The main problem addressed in this work is the need for an efficient and intelligent intrusion detection mechanism capable of identifying both traditional network attacks and modern IoT- based attacks. Existing systems often focus on a single dataset or a specific attack category, limiting their generalization capability across different network environments.

The objective of this project is to design and implement an AI-powered intrusion detection system using multiple machine learning algorithms to improve attack detection performance. The system utilizes the NSL-KDD dataset for traditional net- work intrusion analysis and the Bot-IoT dataset for detecting modern IoT-based attacks. Various machine learning models, including Random Forest, AdaBoost, LightGBM, and Multi- Layer Perceptron (MLP), are employed to analyze network traffic and classify activities effectively.

The novelty of this work lies in the integration and evalu- ation of multiple machine learning algorithms on both traditional and IoT-based intrusion datasets within a unified frame- work. By combining different datasets and algorithms, the proposed system enhances detection accuracy and provides a more comprehensive approach to intelligent network security.

II. LITERATURE SURVEY

A. Machine Learning-Based Intrusion Detection Systems

Intrusion Detection Systems (IDS) have evolved signifi- cantly with the adoption of machine learning techniques for automated threat detection. Traditional IDS approaches relied on signature-based mechanisms, which were effective only for previously known attacks and failed to detect new or evolving threats. Machine learning-based IDS enables systems to learn patterns from historical network traffic and classify activities as normal or malicious. Algorithms such as Support Vector Machine (SVM), Naive Bayes, and Decision Trees have been widely used due to their effectiveness in classification problems. SVM demonstrates strong capability in separating attack and normal traffic, while Naive Bayes provides faster prediction with lower computational cost. However, these methods often face limitations in handling complex and large- scale network data, leading to reduced detection performance in dynamic environments.

B. Ensemble and Deep Learning Approaches for Intrusion Detection

Recent research has focused on improving detection ac- curacy through ensemble and neural network-based models. Ensemble learning methods such as Random Forest and AdaBoost combine multiple classifiers to enhance prediction accuracy and reduce overfitting. Gradient boosting techniques, particularly LightGBM, have gained attention due to their faster training speed and efficient handling of large datasets. In addition, neural network models such as Multi-Layer Perceptron (MLP) have been applied to capture nonlinear relationships in network traffic data, enabling better identi- fication of complex attack patterns. These approaches have demonstrated improved performance compared to traditional classifiers; however, their effectiveness depends heavily on dataset characteristics and feature selection.

C. Dataset-Based Studies and Research Gap

The NSL-KDD dataset has been extensively used as a benchmark for evaluating intrusion detection algorithms in traditional network environments. While it provides a balanced dataset for classification tasks, it does not fully represent modern IoT-based attack scenarios. To address this limitation, recent studies have introduced the Bot-IoT dataset, which includes realistic IoT traffic and modern attack behaviors such as distributed denial-of-service and botnet attacks. Most existing research focuses on either traditional datasets or IoT- specific datasets independently, resulting in limited general- ization across different network environments. The major re- search gap identified is the lack of a unified intrusion detection framework that evaluates multiple machine learning algorithms across both traditional and IoT-based datasets. The proposed work addresses this gap by integrating multiple machine learning models and evaluating their performance using both NSL-KDD and Bot-IoT datasets, thereby improving detection capability and enhancing intelligent network security.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

III. PROPOSED METHODOLOGY

The proposed system presents an AI-powered intrusion detection framework designed to identify malicious network activities using machine learning techniques. The method- ology integrates multiple machine learning algorithms with both traditional and IoT-based intrusion datasets to improve detection accuracy and adaptability across diverse network environments. The overall system consists of data preprocess- ing, model training, prediction, and web-based deployment modules.

A. System Architecture and Design

The machine learning layer plays a crucial role in the proposed system by utilizing trained models such as the KDD model and Bot-IoT model to analyze the input data. These models, stored as serialized files , generate predictions based on patterns learned from historical datasets. The sys- tem also interacts with databases to store user information, maintain logs, and retrieve prediction results when required. This integration ensures efficient data handling and seamless communication between components. By leveraging these trained models, the system is capable of accurately identifying anomalous activities and improving the overall performance of intrusion detection.

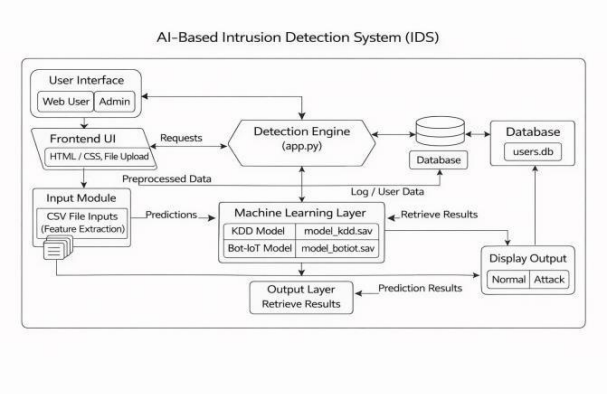

Fig. 1. Proposed AI-Based Intrusion Detection System Architecture

B. Machine Learning Algorithms

The proposed system employs multiple machine learning algorithms to enhance classification performance and ensure reliable intrusion detection. Each algorithm is selected based on its ability to address specific challenges in network intru- sion detection and to improve overall detection capability.

• Random Forest: Utilized for its ability to handle high- dimensional network traffic data effectively.Reduces over- fitting by combining multiple decision trees through ensemble learning.Provides stable and accurate classifi- cation for both normal and attack traffic.

• AdaBoost: Improves classification accuracy by assign- ing higher importance to misclassified samples during training. Iteratively enhances model performance by com- bining weak learners.Effective in improving detection of difficult or less frequent attack patterns.

• Light Gradient Boosting Machine (LightGBM): Se- lected for its faster training speed and computational efficiency. Handles large-scale datasets efficiently using gradient boosting techniques. Provides high prediction accuracy while reducing training time.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

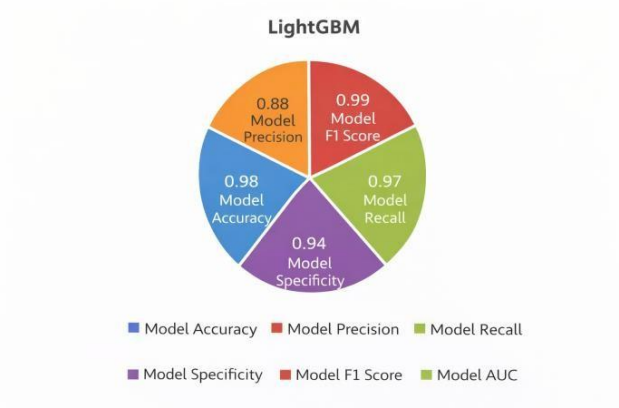

Fig. 2. LightGBM Model Performance metrics

• Multi-Layer Perceptron (MLP): Implemented as a neural network model to capture nonlinear relationships in network traffic data. Capable of learning complex attack patterns that are difficult to detect using traditional mod- els. Enhances detection performance for sophisticated

C. Dataset Description

Two benchmark datasets are used in this work to ensure comprehensive evaluation. The NSL-KDD dataset is employed to represent traditional network intrusion scenarios, including attacks such as Denial of Service, Probe, Remote-to-Local, and User-to-Root attacks. The Bot-IoT dataset is used to represent modern IoT-based attack environments, including botnet and distributed denial-of-service attacks. Using both datasets enables the system to learn diverse attack patterns and improves generalization across different network conditions.

D. Hyperparameter Configuration

Model performance is optimized through appropriate hy- perparameter selection. For the Random Forest model, parameters such as the number of decision trees, maximum tree depth, and minimum samples per split are tuned to balance accuracy and computational cost. AdaBoost utilizes the num- ber of estimators and learning rate to improve classification performance. LightGBM parameters including learning rate, number of leaves, and boosting iterations are adjusted for faster convergence. The MLP model configuration includes the number of hidden layers, activation functions, and learning rate to effectively learn complex traffic patterns.

E. Equation

The intrusion detection problem is formulated as a super- vised classification task where the machine learning model predicts whether a given network instance belongs to a normal or attack class. Let the input feature vector be represented as

$$X = \{x1, x2, x3, \ldots, xn\} \quad (1)$$

where xi denotes the network traffic features such as dura- tion, packet size, and connection count. The classifier learns a mapping function

$$f(X) = Y \quad (2)$$

where $Y \in \{0, 1\}$, with 0 representing normal traffic and 1 representing intrusion. The prediction accuracy of the model is computed as

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

where TP , TN , FP , and FN denote true positive, true negative, false positive, and false negative respectively. Precision and recall are defined as

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

$$\text{Precision} = TP / (TP + FP) \quad (4)$$

$$\text{Recall} = TP / (TP + FN) \quad (5)$$

and the overall classification performance is evaluated using the F1-score given by

$$F1 = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall}) \quad (6)$$

These equations are used to evaluate the effectiveness of the proposed machine learning-based intrusion detection system.

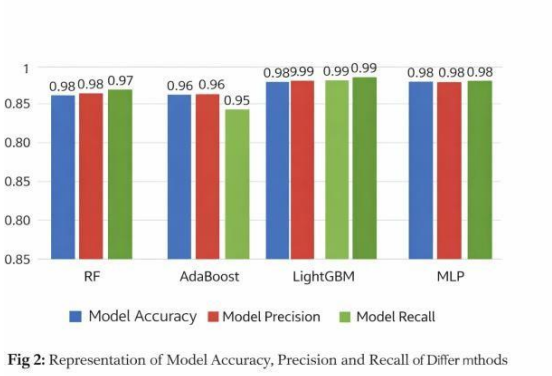

Fig 2: Representation of Model Accuracy, Precision and Recall of Differ mthods

Fig. 3. Representation of model Accuracy, Precision and Recall methods

F. Implementation Details

The proposed system is implemented using Python as the primary programming language. Data preprocessing and model training are performed using libraries such as NumPy, Pandas, and Scikit-learn. LightGBM is used for gradient boosting implementation, while Matplotlib and Seaborn are used for visualization during analysis. The trained models are saved and integrated into a Flask-based web application that provides a user-friendly interface for intrusion prediction. The application allows users to upload network data files and receive classification results in real time, enabling practical deployment of the intrusion detection system.

IV. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed AI-powered intrusion detection system pro- vides several advantages over traditional intrusion detection approaches and existing machine learning-based systems. The integration of multiple machine learning models, combined with evaluation on both traditional and IoT-based datasets, improves detection performance and system reliability.

A. Higher Detection Accuracy

The use of ensemble learning algorithms such as Random Forest, AdaBoost, and LightGBM improves classification per- formance by combining multiple decision models. Experimen- tal results demonstrate higher detection accuracy compared to single-model approaches, typically achieving accuracy levels above 95% on benchmark datasets. This improves the system's ability to correctly identify malicious traffic while minimizing misclassification.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

B. Improved Detection of Diverse Attack Types

By utilizing both NSL-KDD and Bot-IoT datasets, the system is capable of detecting both traditional network attacks and modern IoT-based threats. This improves generalization capability and enables the system to perform effectively across heterogeneous network environments.

C. Reduced False Positive Rate

The combination of multiple machine learning algorithms helps reduce false alarms by improving decision boundaries between normal and malicious traffic. Lower false positive rates reduce unnecessary alerts and improve operational effi- ciency for network administrators.

D. Model Explainability and Interpretability

The system supports explainability through feature impor- tance analysis and model interpretation techniques such as SHAP (SHapley Additive Explanations). This allows identi- fication of important network features influencing predictions, increasing transparency and trust in the intrusion detection process.

E. Real-Time Detection Capability

The trained models are integrated into a Flask-based web application, enabling real-time prediction of network traffic. This allows faster response to potential threats and supports practical deployment in real-world environments.

F. Scalability and Computational Efficiency

The use of LightGBM enables faster training and prediction, making the system suitable for large-scale network data. The modular design allows easy integration of additional datasets or algorithms without significant modification.

V. RESULTS AND DISCUSSION

The performance of the proposed intrusion detection sys- tem is evaluated using experiments conducted on the NSL-KDD and Bot-IoT datasets. The objective of the evaluation is to analyze the effectiveness of multiple machine learning algorithms in accurately classifying network traffic as normal or malicious. The datasets were divided into training and testing sets to ensure unbiased evaluation of model perfor- mance. Experimental analysis was carried out using standard performance metrics including accuracy, precision, recall, and F1-score

A. Evaluation Metrics

To measure the performance of the proposed system, com- monly used classification metrics were employed. Accuracy represents the overall percentage of correctly classified in- stances, while precision measures the proportion of correctly identified attack instances among all predicted attacks. Recall evaluates the ability of the model to detect actual attacks, and the F1-score provides a balanced measure by combining precision and recall. These metrics are essential for intrusion detection systems, where both correct detection and reduction of false alarms are critical.

B. Experimental Results

The experimental results demonstrate that ensemble learning algorithms achieve superior performance compared to tra- ditional classifiers. Random Forest and LightGBM models provide higher accuracy and F1-score due to their ability to handle complex feature relationships and large-scale datasets. AdaBoost improves classification performance by focusing on previously misclassified samples, while the MLP model effectively captures nonlinear patterns in network traffic data. The results indicate that the proposed system achieves high detection accuracy, typically exceeding 95% with improved F1-scores across both datasets. Performance comparison tables and graphical representations illustrate that the proposed multi- model approach outperforms baseline models such as Naive Bayes and Support Vector Machine in terms of detection accuracy and stability. The results also show a reduction in false positive rates, which is a critical requirement for practical intrusion detection systems.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

C. Discussion

The experimental findings confirm that combining multi- ple machine learning algorithms improves intrusion detection capability by addressing different characteristics of attack patterns. The use of both NSL-KDD and Bot-IoT datasets enables the system to generalize across traditional and IoT- based attack environments. Compared to existing approaches

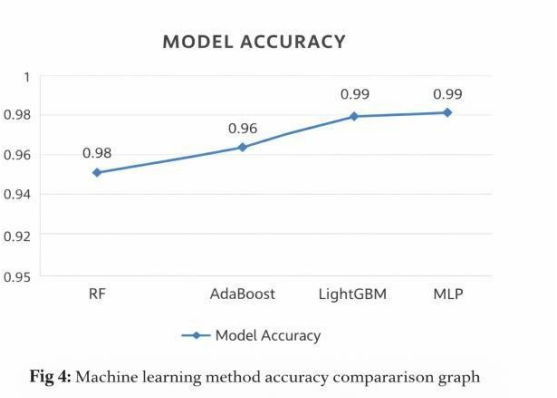

Fig 4: Machine learning method accuracy compararison graph

Fig. 4. Machine Learning method accuracy comparison graph

that rely on single datasets or individual algorithms, the proposed system demonstrates improved robustness and adapt- ability. The integration of the trained models into a Flask- based environment further validates the feasibility of real- time deployment. Overall, the results indicate that the proposed methodology provides a reliable and efficient solution for intelligent network intrusion detection.

VI. CONCLUSION AND FUTURE WORK

This paper presented an AI-powered intrusion detection system designed to enhance intelligent network security using machine learning techniques. The proposed system integrates multiple machine learning algorithms, including Random For- est, AdaBoost, LightGBM, and Multi-Layer Perceptron, to effectively classify network traffic as normal or malicious. The use of both NSL-KDD and Bot-IoT datasets enables the system to detect traditional network intrusions as well as modern IoT-based attacks, improving the overall detec- tion capability and generalization performance. Experimen- tal results demonstrate that the proposed approach achieves high detection accuracy and improved F1- scores compared to conventional single-model approaches, while reducing false positive rates. The integration of the trained models into a Flask-based framework further validates the feasibility of real- time intrusion detection in practical environments.

Despite these advantages, the proposed system has certain limitations. The performance of machine learning models de- pends heavily on the quality and diversity of training datasets, and real-world network traffic may contain unseen attack patterns not represented in the datasets used. Additionally, training multiple models increases computational complexity and may require optimization for deployment in resource- constrained environments.

Future work will focus on improving scalability and adapt- ability of the system by incorporating advanced techniques such as federated learning for privacy-preserving distributed training, deep learning architectures for enhanced feature extraction, and real-time streaming data analysis. Further re- search can also explore adaptive model updating mechanisms and deployment at edge or cloud environments to improve response time and support large-scale intelligent network security systems.

REFERENCES

[1] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A detailed analysis of the KDD CUP 99 data set,” Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, April 2026 Impact Factor: 8.2

[2] N. Moustafa and J. Slay, "The evaluation of Network Anomaly Detec- tion Systems: Statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset," Information Security Journal, vol. 25, no. 1–3, pp. 18–31, 2016.
[3] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," Future Generation Computer Systems, vol. 100, pp. 779–796, 2019.
[4] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[5] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119–139, 1997.
[6] G. Ke et al., "LightGBM: A highly efficient gradient boosting deci- sion tree," Advances in Neural Information Processing Systems, pp. 3146–3154, 2017.
[7] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning repre- sentations by back-propagating errors," Nature, vol. 323, pp. 533–536, 1986.
[8] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
[9] W. Lee and S. J. Stolfo, "A framework for constructing features and models for intrusion detection systems," ACM Transactions on Information and System Security, vol. 3, no. 4, pp. 227–261, 2000.
[10] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems (IDPS)," NIST Special Publication 800-94, 2007.
[11] S. Axelsson, "Intrusion detection systems: A survey and taxonomy," Technical Report, Chalmers University of Technology, 2000.
[12] J. McHugh, "Testing intrusion detection systems: A critique of the 1998 and 1999 DARPA intrusion detection system evaluations," ACM Transactions on Information and System Security, vol. 3, no. 4, pp. 262–294, 2000.
[13] H. Liao, C. R. Lin, Y. C. Lin, and K. Y. Tung, "Intrusion detection system: A comprehensive review," Journal of Network and Computer Applications, vol. 36, no. 1, pp. 16–24, 2013.
[14] A. Patcha and J. M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," Computer Networks, vol. 51, no. 12, pp. 3448–3470, 2007.
[15] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," ICISSP, 2018.
[16] M. Ring, D. Landes, D. Hotho, and A. Hotho, "Flow-based network traffic generation using Generative Adversarial Networks," Computers Security, vol. 82, pp. 156–172, 2019.
[17] S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using ensemble of soft computing paradigms," Intelligent Systems Design and Applications, 2003.
[18] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," IEEE Transactions on Computers, vol. 65, no. 10, pp. 2986–2998, 2016.
[19] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," EAI International Conference on Bio-inspired Information and Communications Technologies, 2016.
[20] J. Zhang, M. Zulkernine, and A. Haque, "Random-forests-based network intrusion detection systems," IEEE Transactions on Systems, Man, and Cybernetics, vol. 38, no. 5, pp. 649–659, 2008.
[21] S. Potluri and C. Diedrich, "Accelerated deep neural networks for enhanced intrusion detection system," IEEE International Conference on Industrial Informatics, 2016.
[22] T. Kim, B. Kang, M. Rho, S. Sezer, and E. Im, "A multimodal deep learning method for Android malware detection using various features," IEEE Transactions on Information Forensics and Security, vol. 14, no. 3, pp. 773–788, 2019.
[23] M. Hodo, X. Bellekens, A. Hamilton, P. Dubouilh, and E. Iorkyase, "Threat analysis of IoT networks using artificial neural network intrusion detection system," International Symposium on Networks, Computers and Communications, 2016.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, April 2026
Impact Factor: 8.2

[24] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," IEEE Symposium on Security and Privacy, pp. 305–316, 2010.
[25] N. Moustafa, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," Military Communications and Information Systems Conference, 2015