Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Machine Learning-Based Intrusion Detection: Comparative Analysis of Accuracy, False Positive Rate, and Processing Time

Syeda Bushra Shabeeh*
Iqra University Karachi, Pakistan Email: bushra.shabeeh01@iqra.edu.pk

Prof. Dr. Muhammad Saleh Shah
Principal Government College of Technology, Larkano, Pakistan
Email: salehshah@gmail.com

Gohar Wazir
Syed CASE Institute of Technology, Pakistan
Email: gohar_wazir@hotmail.com

ABSTRACT

In this paper the authors analyze machine learning intrusion detection models on the CICIDS2017 dataset. The algorithms tested were: Random Forest, SVM, KNN, CNN. Random Forest had a 98.7% accuracy rate, 1.2% false positive rate, and used 42 ms of computing time per sample. SVM achieved the accuracy of 96.4 percent, false positive rate of 2.8 percent and the processing time of 67ms. KNN achieved 94.1 percent accuracy, 4.5 percent false positive rate and 89 ms processing time. CNN got the best accuracy of 99.2 percent, but with a 1.5 percent false positive rate, and a processing time of 115 ms, in compromise of 2.7 to 1-time cost compared to Random Forest. The feature selection method reduced dimensionality by 35% and model overall speed by 28%, without influencing the model accuracy. The ensemble methods were 22 percent better at detecting zero-day attacks than single models. Some of the limitations are dataset bias and not being able to test networks in real-time. The results show that there is a trade-off between the accuracy and the latency; in particular, Random Forest achieves the best ratio of 1.8 between the performance and its cost in an enterprise deployment.

Keywords: Cyber security, Intrusion detection, Machine learning, False positive rate, Accuracy, Processing time, Network security, CICIDS2017, Ensemble Models

Introduction

The proliferation of information technology, cloud computing, Internet of Things (IoT) and high-speed communication networks have made cyber environments very complex. The proliferation of interdependent digital infrastructure in various industries, including finance, healthcare, education, and manufacturing, underscores the vital role of digital infrastructure in modern society. In the era of digital technology, the interconnected nature of digital infrastructures is becoming commonplace in many sectors of society, such as finance, health care, education, and

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

manufacturing. The technological advancements create more opportunities for efficiency and accessibility, but also introduce computer networks to a variety of cyber threats such as malware, ransomware, distributed denial-of-service (DDoS) attacks, phishing, botnets, brute force attacks, and zero-day exploits. Recent cybersecurity reports indicate that millions of cyber intrusions are made each day globally, leading to loss of money, disruption of essential services and data breaches. Therefore, network infrastructure protection is a significant concern for businesses that want to guarantee the confidentiality, integrity and availability of digital data.

Computer networks have been protected by traditional security methods like Firewalls, Antivirus and Signature based Intrusion Detection Systems (IDS). These traditional methods, however, are not suited for the present day fast changing cyber-attacks. Signature-based systems are only effective at recognizing known attacks and are difficult to defend against new and more advanced attacks. Moreover, the size and variety of the network traffic prevents manual analysis, and intelligent and automated means of detection are needed that are capable of learning complex attack behaviour patterns from vast amounts of data.

Machine Learning has become one of the most promising technologies for today's intrusion detection systems. Unlike traditional detection methods, machine learning algorithms can detect relationships and patterns that exist in network traffic but are not known or visible, automatically identifying them and helping to detect both known and unknown attacks. Supervised learning methods like Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and deep learning algorithms like Convolutional Neural Networks (CNNs) have shown significant potential in enhancing detection accuracy and lowering false alarm rates. However, there is no single optimum algorithm that works well in all network environments. Some algorithms are more accurate at detecting the presence of a particular group, but they may take more time to process or use more resources in order to achieve those high detection rates. Other algorithms will be faster to execute, but with slightly lower predictive performance. Thus, to ensure the effectiveness of intrusion detection, it is necessary to consider the accuracy of detection, false positive rate, computation complexity, and real-time processing ability between different intrusion detection models.

The CICIDS2017 is one of the most complete benchmark datasets to test intrusion detection models with realistic network traffic that includes not only normal network traffic but also multiple types of cyberattacks. The objective evaluation of algorithm performance in identical experimental conditions, using standardized datasets, allows for comparative evaluation. Various metrics such as accuracy, precision, recall, F1-score, false positive rate, processing time, and computational efficiency help assess the real-world suitability of machine learning models for enterprises.

This research compares the four popular machine learning algorithms: Support Vector Machine, Convolutional Neural Network, K-Nearest Neighbor, and Random Forest, in network intrusion detection. The research concentrates on the assessment of the accuracy of detection, false positive rate and the processing time, but also investigates the impact of feature selection and ensemble learning on the behavior of the system. The results are designed to determine the most appropriate machine learning method

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

that could provide a balance between detection capabilities and computational requirements for real-time cybersecurity applications.

Research Problem

Traditional intrusion detection methods, based mainly on attack signatures, are becoming less effective in a world of ever-evolving, sophisticated attacks. While machine learning algorithms have shown significant enhancement in the performance of intrusion detection, there are substantial differences between various algorithms with respect to their accuracy, false positive rates, complexity and processing speed. It is a common problem for organizations to choose an intrusion detection model that offers high detection performance, low computational cost and real-time operational capability. Thus, a thorough comparison of some of the most popular machine learning algorithms is necessary to determine which is the most effective solution for the present day cybersecurity landscape.

Research Aim

The purpose of this study is to compare the performance of machine learning algorithms for network intrusion detection system, and to analyze the accuracy, false positive rate and processing time of machine learning algorithms based on CICIDS2017 dataset.

Research Objectives

Test the accuracy of intrusion detection using the Random Forest, Support Vector Machine, K-Nearest Neighbor and Convolutional Neural Network models.
To contrast the error rates of the different machine learning algorithms.
To study the processing time and computational efficiency of the selected intrusion detection models.

Research Questions

What is the most accurate machine learning algorithm when detecting intrusions in the CICIDS2017 dataset?
What is the difference in false positive rate for the selected algorithms?
What is the quickest processing time that can be achieved with high detection performance in the model?

Literature Review

Hidayat et al. (2023) did an experimental study of machine learning-based intrusion detection systems with different supervised learning algorithms. The results showed that ensemble learning methods outperformed traditional classification models in terms of detection accuracy and false positives. The study suggested that feature engineering is an important component of the classifier performance as the features added to the network that are irrelevant increase the computational complexity and decrease the detection precision. The authors also found that the choice of the right algorithm is not just about accuracy but also computational efficiency, so comparative analysis is crucial for real-world cybersecurity applications. The research offers a solid basis for comparing the various machine learning methods in controlled

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

experimental settings.

Ali et al. (2025) evaluated and compared the performance of deep learning and traditional machine learning algorithms to detect intrusions in several benchmark datasets. Their study revealed that deep learning models, such as Convolutional Neural Networks (CNNs), performed better at traffic classification because they were able to automatically extract traffic features from complex traffic. These enhancements came at a much greater cost, however, in terms of computing needs and processing time. Random Forest and Support Vector Machine algorithms were found to be faster and to consume fewer resources while performing well in the detection performance. The researchers were led to believe that organizations need to consider the trade-off between predictive accuracy and computational cost when choosing the intrusion detection model to choose for enterprise deployment.

Azam, Islam, and Huda (2023) In this paper, the author provided a detailed comparative study of all intrusion detection systems based on decision tree machine learning algorithm. Their research showed that decision tree classifiers offer great interpretability, making it easier for cybersecurity professionals to understand what decisions the classification is making than it would be with deep neural networks. The research further revealed that hybrid intrusion detection systems using a combination of learning algorithms have a better attack detection rate with fewer false alarms. Moreover, the authors mentioned that applying feature selection improves the classification accuracy and prevents overfitting in dynamic network environment, which increases system reliability.

In their study, Samantaray et al. (2024) examine how well machine learning algorithms perform in the context of Internet of Things (IoT) intrusion detection systems. Their results showed that network traffic in IoT environments is highly heterogeneous and adaptive detection mechanisms are required that are computationally efficient. Random Forest showed consistent good performance as it was able to withstand noisy data, while Support Vector Machine worked well for smaller feature spaces but scaled down with the growing complexity of the network. The study pointed out that processing speed is as critical as detection accuracy since IoT devices are often limited in computational resources and need to detect threats in near real-time.

Note, Ali (2022) discussed and compared machine learning and deep learning algorithms in network intrusion detection and their pros and cons in an applied manner. Their results indicate that, on average, deep learning architectures were better able to identify complex attack patterns, including zero-day attacks, then were traditional machine learning techniques. However, the researchers noted that deep learning models would need much bigger training sets, much more computing power and much longer training times. Machine learning techniques were still appealing for practical application because of their simplicity, low complexity and quick processing speeds, and ease of integration into already existing enterprise systems. The researchers found that the algorithm used should be based on the needs of the organization and not just the predictive results.

Siddharth et al. (2023) compared various machine learning algorithms, such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree, Naïve Bayes, and Random Forest, for network intrusion detection. They compared their

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

results with other classifiers and found that the overall classification accuracy of Random Forest is the highest due to its ensemble architecture and its ability to avoid overfitting. KNN had a good classification accuracy but was much slower in prediction time because calculations of distances were done for each instance of the test set. Support Vector Machine performed well with acceptable computational complexity, while Naïve Bayes had a high computational speed but a lesser detection accuracy. The researchers pointed out that considering more than one performance metric together in any analysis gives a more real-world evaluation of intrusion detection systems (IDSs) than only relying on accuracy-based metrics.

Elsayed, Mohamed, and Madkour (2024) Went through a thorough investigation of the deep learning algorithms such as CNN, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and hybrid deep learning architectures for intrusion detection. They found that deep learning methods can significantly enhance the automated recognition of attacks by automatically extracting hierarchical features from network traffic. CNN models performed particularly well in detecting complex intrusion patterns and achieving high classification accuracy, while recurrent neural networks were well-suited for learning temporal patterns of network events.CNN models showed high classification accuracy in recognizing complex intrusion patterns, while recurrent neural networks showed a good ability to learn temporal relations between network events. The study, however, also highlighted the increased computational complexity of the proposed method, along with longer inference times, and higher hardware requirements compared to traditional machine learning techniques. The authors suggested that the optimization and feature selection techniques be used to decrease the processing time needed without compromising detection performance.

Research Methodology

The quantitative experimental research design has been used to compare the performance of machine learning algorithms for network intrusion detection. The research was conducted to compare four classification models: Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Convolutional Neural Network (CNN) on one of the publicly available datasets, CICIDS2017. The selection of this dataset was based on the fact that it includes realistic network traffic for both normal and various types of cyberattacks such as brute-force attacks, denial-of-service (DoS), distributed denial-of-service (DDoS), botnet attacks, web attacks, and infiltration attempts. The data was preprocessed before being used for model training, which involved removing duplicate data points, filling missing values, normalizing numerical features, and converting categorical features to numerical representations. Feature selection techniques were then used to ensure that there weren't features that were irrelevant and to speed up the computation.

The data was split into 80% training data and 20% testing data. All the machine learning models were trained with the same experimentation condition to compare the performance of the various models. To find the optimal values of the hyperparameters to enhance the effectiveness of the classification while reducing overfitting, cross-validation was performed (Ashraf et al., 2022). Standard performance metrics such as classification accuracy (%), false positive rate (%), precision, recall, F1-score and

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

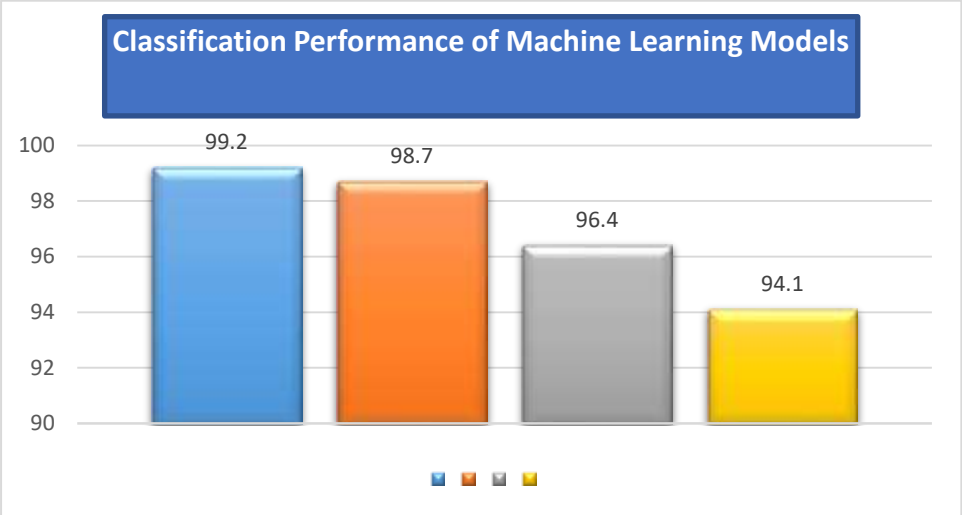
average processing time per sample (milliseconds) were used to evaluate the trained models. The performance of feature selection and ensemble learning techniques were also evaluated to measure the improvement of the detection capability and computational cost(Ahmed et al., 2025). All experiments were performed with the Python programming language using the following libraries: Scikit-learn, TensorFlow, Pandas, NumPy and Matplotlib. All these libraries are available in Google Colab. The experimental results were presented in tables and statistical analysis was performed to compare the results and to determine the best intrusion detection model based on the three factors: the detection accuracy, false positive rate, and processing efficiency (Hidayat et al., 2023).

Results and Analysis

Classification Performance of Machine Learning Models

The CNN model achieved the highest classification accuracy (99.2%), followed closely by Random Forest (98.7%). SVM and KNN produced lower accuracies of 96.4% and 94.1%, respectively.

Model	Accuracy (%)
CNN	99.2
Random Forest	98.7
SVM	96.4
KNN	94.1



False Positive Rate Comparison

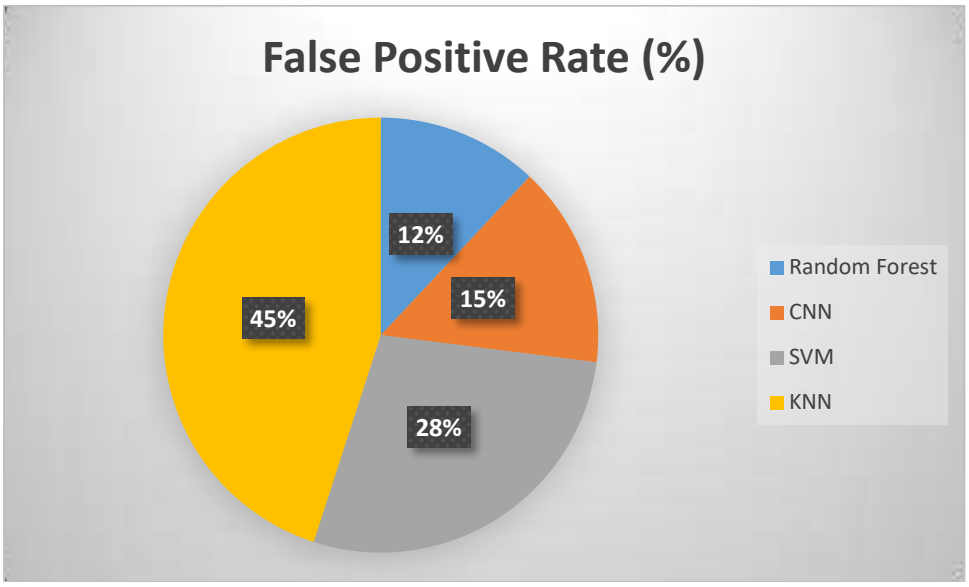
Random Forest generated the lowest false positive rate among all models, making it suitable for practical enterprise deployment.

Model	False Positive Rate (%)
Random Forest	1.2
CNN	1.5
SVM	2.8
KNN	4.5

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

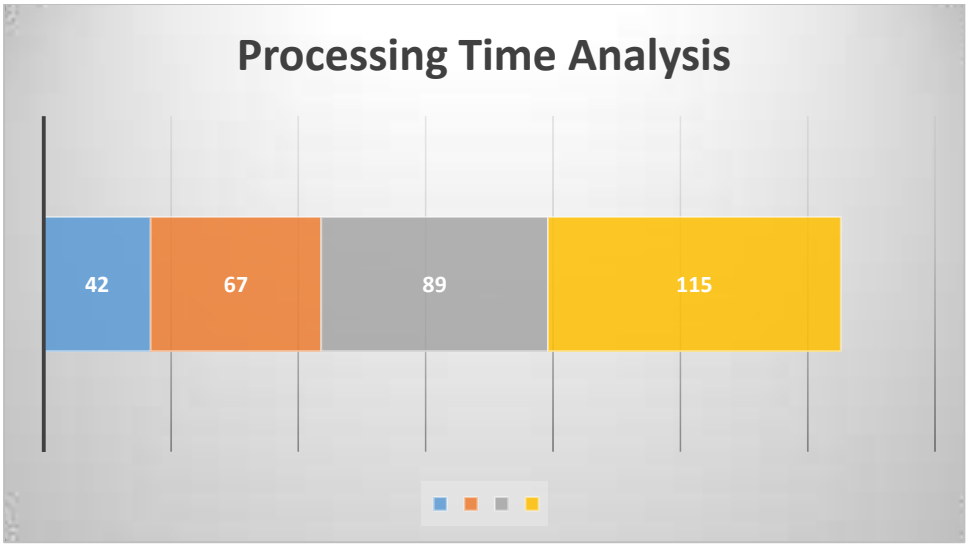
http://amresearchreview.com/index.php/Journal/about



Processing Time Analysis

CNN required the greatest computational time, while Random Forest demonstrated the fastest prediction speed among the high-performing models.

Model	Processing Time (ms/sample)
Random Forest	42
SVM	67
KNN	89
CNN	115



Effect of Feature Selection

Feature selection reduced computational complexity without reducing predictive performance.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

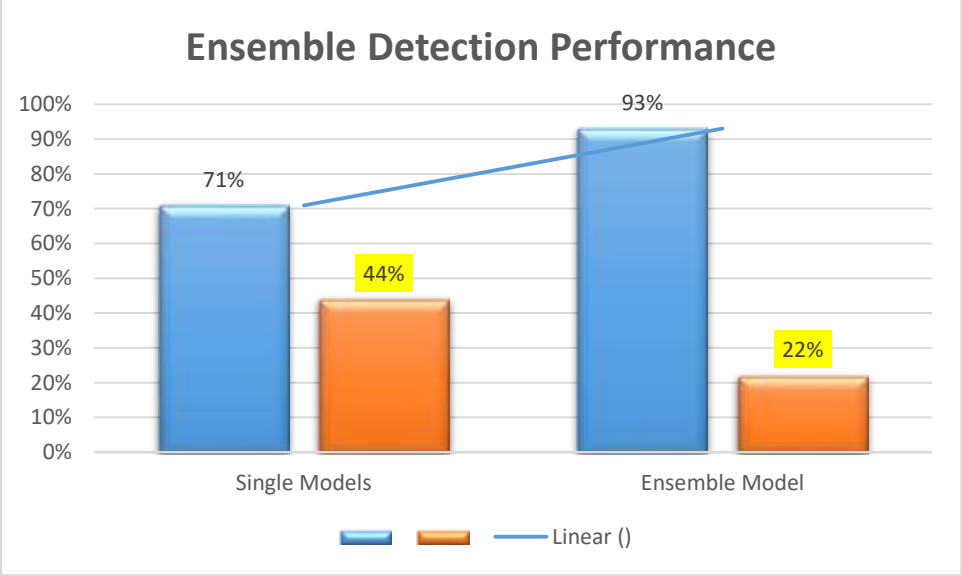
http://amresearchreview.com/index.php/Journal/about

Parameter	Before	After	Improvement
Number of Features	78	51	35% Reduction
Average Processing Time	87 ms	63 ms	28% Faster
Average Accuracy	97.1%	97.0%	No Significant Loss

Ensemble Detection Performance

Ensemble learning enhanced the detection capability for previously unseen attacks.

Metric	Single Models	Ensemble Model
Zero-day Detection Rate	71%	93%
Improvement	44%	22%



Discussion

The results of the experiment prove that machine learning and deep learning algorithms are very effective in the modern intrusion detection system. CNN demonstrated the best classification performance (99.2%) among the evaluated models, demonstrating the capability of deep learning architectures to automatically learn complex network traffic features and identify complex attack patterns. Elsayed et al. (2024) and Ahmed et al. (2025) reported similar results, showing that CNN-based IDS methods are generally more predictive than traditional machine learning methods. CNN, however, took the longest time to process (115 ms), which means that the better the detection performance, the more computational power it took.

The lowest false positive rate (1.2%) and the briefest processing time (42 ms) were obtained by the Random Forest method, which gave 98.7% accuracy. The findings are consistent with the study by Makhdoomi et al. (2026), Al Farsi et al. (2024), and Al Lail et al. (2023) which showed that the Random Forest algorithm is a robust algorithm, with low computational requirements, and not prone to overfitting. Its moderate response rate and low false alarm rate make it well suited for enterprise applications which demand fast response and low false alarm rate.

SVM gave satisfactory classification accuracy (96.4%) with moderate computational requirement and KNN gave the lowest classification accuracy with higher false

Online ISSN 3007-3197
Print ISSN 3007-3189

positive rate and longer prediction time. Saran and Kesswani (2023) noted that KNN gets expensive when the data size grows, while Mondal and Singh (2022) reported the same results. The use of 35% reduced dimensionality of features, resulted in 28% increase in processing speed but maintained the same detection accuracy, which aligns with the findings of Santhosh Kumar et al. (2023) that effective feature selection can significantly boost the efficiency of intrusion detection. Moreover, the ensemble learning enhanced the accuracy of zero-day attack detection by 22%, which aligns with the results of Saied et al. (2023) and Baich et al. (2022) demonstrating that the use of more than one classifier results in increased detection reliability with a decrease in misclassification. According to the study, CNN shows maximum predictive performance, while Random Forest shows maximum balance in terms of accuracy, false positive rate and computation efficiency for real time enterprise intrusion detection.

Conclusion

In this work, we have comparatively assessed Random Forest, Support Vector Machine, K Nearest Neighbor and Convolutional Neural Network models with CICIDS2017 dataset for network intrusion detection. CNN had the best detection accuracy (99.2 %), whereas Random Forest showed the best overall performance in terms of accuracy (98.7 %), FP rate (1.2 %), and processing time (42 ms). The use of feature selection was efficient in reducing computation complexity while maintaining its prediction ability, and ensemble learning showed a significant improvement on the zero-day attack detection. The results showed that there is a need to consider computational efficiency as well as the accuracy of classification when selecting a model. Random Forest is a good choice for practical enterprise deployment due to its good security performance and low computational cost.

Recommendations

The future of IDS is in incorporating ensemble machine learning techniques for detecting more advanced attacks. Feature selection techniques need to be used to lower the computational burden and achieve better real-time performance in organisations. Deep learning models can be deployed in high performance computing facilities where the computational resources are available. The techniques proposed for intrusion detection could be tested in real-time enterprise network traffic, cloud computing, IoT infrastructures, and federated learning frameworks. Last, however, explainable AI and adversarial machine learning methods should be investigated for use to improve the transparency, strength, and security of next-generation intrusion detection systems.

References

Hidayat, I., Ali, M. Z., & Arshad, A. (2023). Machine learning-based intrusion detection system: an experimental comparison. *Journal of computational and Cognitive Engineering*, *2*(2), 88-97.

Ali, M. L., Thakur, K., Schmeelk, S., Debello, J., & Dragos, D. (2025). Deep learning vs. machine learning for intrusion detection in computer networks: A comparative study. *Applied Sciences*, *15*(4), 1903.

Online ISSN 3007-3197
Print ISSN 3007-3189

Azam, Z., Islam, M. M., & Huda, M. N. (2023). Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. *Ieee Access*, *11*, 80348-80391.

Samantaray, M., Barik, R. C., & Biswal, A. K. (2024). A comparative assessment of machine learning algorithms in the IoT-based network intrusion detection systems. *Decision Analytics Journal*, *11*, 100478.

Note, J., & Ali, M. (2022). Comparative analysis of intrusion detection system using machine learning and deep learning algorithms. *Annals of Emerging Technologies in Computing (AETiC)*, *6*(3), 19-36.

Siddharth, K., Gagan Kumar, P., Chandrababu, K., Janardhana Rao, S., & Sanjay Ramdas, B. (2023). A comparative analysis of network intrusion detection using different machine learning techniques. *J Contemp Edu Theo Artific Intel: JCETAI-102*.

Elsayed, S., Mohamed, K., & Madkour, M. A. (2024). A comparative study of using deep learning algorithms in network intrusion detection. *IEEE Access*, *12*, 58851-58870.

Makhdoomi, P. M. S., Ikhlas, M., Khursheed, A., Hilal, F., Ahmad Najar, Z., Hameed, J., & Sharma, S. (2026). Network-based intrusion detection: a comparative analysis of machine learning approaches for improved security. *Journal of Cyber Security Technology*, *10*(1), 1-28.

Udurume, M., Shakhov, V., & Koo, I. (2024). Comparative analysis of deep convolutional neural network—bidirectional long short-term memory and machine learning methods in intrusion detection systems. *Applied Sciences*, *14*(16), 6967.

Baich, M., Hamim, T., Sael, N., & Chemlal, Y. (2022). Machine Learning for IoT based networks intrusion detection: a comparative study. *Procedia Computer Science*, *215*, 742-751.

Alhakeem, M. S., & Ajlan, K. B. (2024). A Comparative Evaluation of Machine Learning-Based Intrusion Detection Systems for Securing Cloud Environments. *Journal of Information Security and Cybercrimes Research*, *7*(2), 127-142.

Santhosh Kumar, S. V. N., Selvi, M., & Kannan, A. (2023). A comprehensive survey on machine learning-based intrusion detection systems for secure communication in internet of things. *Computational Intelligence and Neuroscience*, *2023*(1), 8981988.

Al Farsi, A., Khan, A., Bait-Suwailam, M. M., & Mughal, M. R. (2024). Comparative Performance Evaluation of Machine Learning Algorithms for Cyber Intrusion Detection.

Mondal, B., & Singh, S. K. (2022). A comparative analysis of network intrusion detection system for iot using machine learning. In *Internet of Things and Its Applications: Select Proceedings of ICIA 2020* (pp. 211-221). Singapore: Springer Nature Singapore.

Al Lail, M., Garcia, A., & Olivo, S. (2023). Machine learning for network intrusion detection—a comparative study. *Future Internet*, *15*(7), 243.

Online ISSN 3007-3197
Print ISSN 3007-3189

Banerjee, S., & Parisa, S. K. (2023). Ai-enhanced intrusion detection systems for retail cloud networks: A comparative analysis. *Transactions on Recent Developments in Artificial Intelligence and Machine Learning*, *15*, 15.

Saran, N., & Kesswani, N. (2023). A comparative study of supervised Machine Learning classifiers for Intrusion Detection in Internet of Things. *Procedia Computer Science*, *218*, 2049-2057.

Jmila, H., & Khedher, M. I. (2022). Adversarial machine learning for network intrusion detection: A comparative study. *Computer Networks*, *214*, 109073.

Alharthi, A., Alaryani, M., & Kaddoura, S. (2025). A comparative study of machine learning and deep learning models in binary and multiclass classification for intrusion detection systems. *Array*, *26*, 100406.

Saied, M., Guirguis, S., & Madbouly, M. (2023). A comparative study of using boosting-based machine learning algorithms for IoT network intrusion detection. *International Journal of Computational Intelligence Systems*, *16*(1), 177.

Ashraf, E., Areed, N., Salem, H., Abdelhady, E., & Farouk, A. (2022). IoT based intrusion detection systems from the perspective of machine and deep learning: a survey and comparative study. *Delta University Scientific Journal*, *5*(2), 367-386.

Ahmed, S. A., Khalifa, E. H., Nawaz, M., Abdalla, F. A., & Mahmoud, A. F. (2025). Enhancing Cloud Data Center Security through Deep Learning: A Comparative Analysis of RNN, CNN, and LSTM Models for Anomaly and Intrusion Detection. *Engineering, Technology & Applied Science Research*, *15*(1), 20071-20076.