International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancing Intrusion Detection Systems with Ai: Examine the Integration of Ai into Traditional Ids to Improve Detection Rates and Reduce False Positives

[1]Shekh Tareq Ali, [2]MohammadMajharul Islam Jabed, [3]Mahabub Alam Khan, [4]Sheikh Md Kamrul Islam Rasel, [5]Md Abdullah Al Nahid, [6]Touhid Bhuiyan Professor

Submitted: 07/08/2024 **Revised:** 20/09/2024 **Accepted:** 29/09/2024

Abstract: These days, when cyber risks are more frequent and powerful, traditional intrusion detection systems (IDS) are lacking in stopping new and sophisticated attacks. Although signature-based and anomaly-based IDS have strong basics, they usually generate many misleading alarms and do not react quickly to new threats. Integrating AI, specifically Machine Learning (ML) and Deep Learning (DL), in IDS is valuable in improving threat detection and the system's response. This paper aims to explore how AI is helping IDS systems achieve better results and fewer fake signals. A comparison of Support Vector Machines (SVM), Random Forest (RF), and Convolutional Neural Networks (CNN) reveals that using AI can make IDS more effective than non-AI-based IDS when tested using benchmark datasets NSL-KDD and CICIDS2017 (Shone et al., 2018; Ring et al., 2019). By blending anomaly identification with intelligent classification and adaptive learning, the new architecture can identify zero-day attacks more accurately. Models are evaluated using precision, recall, F1-score, and detection latency. The latest results show that our approach could reduce false positives by 35% and lead to more true positives. It additionally demonstrates events with bar graphs, pie charts, and system diagrams that evidence the changes in performance and architecture. It adds helpful knowledge to intelligent cybersecurity and offers valuable advice for using AI-driven IDS within enterprises. Future research will investigate federated learning and reinforcement learning to help improve the scalability and privacy of IDS algorithms.

Keywords: *Intrusion Detection System, Artificial Intelligence, Machine Learning, Cybersecurity, False Positives, Deep Learning, Network Security*

1. Introduction

Network security professionals have recently faced serious issues due to the fast development of cyber threats. Traditional IDS, which work by detecting attacks using signatures or anomalies, are now struggling with the new and advanced techniques criminals use. Most systems like these have a high error rate that declares something as dangerous when it is not, and are slow to react to threats, which can endanger the organization's security (Sommer&Paxson, 2010).

Using AI in IDS is now seen as a powerful way to address the weaknesses of older detection systems. Using ML and DL in AI, IDS can pick up on patterns from past experience and respond to new attackers it has not encountered before. Through AI, today's IDS can improve classification accuracy, expedite anomaly finding, and avoid many false alarms (Buczak&Guven, 2016).

There are mainly two types of traditional IDS: Signature-based systems compare data traffic to a database of traffic patterns representing known attacks, whereas anomaly-based systems check for anything that behaves differently from the baseline they have established (Garcia-Teodoro et al., 2009).

[1]*School Of IT, Washington University of Science and Technology*
alitareq774@gmail.com
[2]*School Of IT, Washington University of Science and Technology*
mi_jabed@yahoo.com
[3]*School Of IT, Washington University of Science and Technology*
mahabub95tech@gmail.com
[4]*School Of IT, Washington University of Science and Technology*
sheikhmdkamrul49@gmail.com
[5]*School Of IT, Washington University of Science and Technology*
hosennahid511@gmail.com
[6]*School of IT, Washington University of Science and Technology*
touhid.bhuiyan@wust.edu

Despite their capability, these systems have some significant issues. In response to zero-day attacks, signature-based IDS proves ineffective, and anomaly-based systems tend to trigger many false alarms due to regular behavior changes.

IDS can deal with these problems thanks to AI by introducing innovative training and predictive abilities. When trained on labeled datasets, SVM, DT, and RF can accurately classify intrusions. Both k-Means clustering and Autoencoders are examples of unsupervised learning that recognize new attack behavior without labeling the data (Javaid et al., 2016). It is important to note that Deep Learning techniques, like CNNs and RNNs, are being applied more often to study complex high-dimensional data that arrives in networks in real time (Kim et al., 2016).

AI-backed IDSs take in new information constantly and become more reliable as time goes on. They help identify difficult relationships and patterns in data that standard rule-based methods can neglect. In addition, using datasets like NSL-KDD, UNSW-NB15, and CICIDS2017 makes AI models better suited for real-world use (Moustafa&Slay, 2015; Ring et al., 2019).

One major issue with setting up IDSs is that the system can produce too many false positives, creating confusion for security team members. AI handles this by sharpening where the classes meet and merging the strengths of several algorithms. As seen in Table 1, AI models have much better results against traditional models for important performance markers. Such systems are equipped to pick out features from data showing unusual and complex relationships, things that ordinary rules might fail to see. Moreover, building AI models with real-world datasets, such as NSL-KDD, UNSW-NB15, and CICIDS2017, can help them work well and apply to different situations when used (Moustafa& Slay, 2015). Ring et al., 2019).

A problem with IDS is that it tends to give a lot of false alerts, sometimes more than the security team can handle, creating alert fatigue. With AI, classifiers can use better segmentation and algorithm blends to address this issue. According to the results in Table 1, AI models are much better than traditional ones regarding key performance indicators:

Table 1: Performance Comparison of Traditional vs. AI-Enhanced IDS

Model Type	**Detection Rate (%)**	**False Positive Rate (%)**	**Accuracy (%)**
Signature-Based IDS	82.5	17.3	85.1
SVM (AI Model)	93.4	5.2	92.8
CNN (AI Model)	95.8	3.1	96.2

AI uses in IDS include better detection, more flexibility, robotic tasks, and immediate responses. The systems can recognize new threats and lessen the effect of manual network traffic observation. AI-assisted IDS can be easily deployed in both spread-out and cloud systems.

In this work, we explore the possibilities of integrating AI into the classic IDS system. Our study covers a variety of ways to measure the performance of AI algorithms on data drawn from the real world and research. We judge their performance by detection rate, precision, recall, and F1-score. Furthermore, we display model comparisons and detection trends using bar graphs and pie charts to make the results more understandable.

2. Literature Review

Since the first Intrusion Detection Systems were developed in the 1980s, they have developed a lot. Isharmalaya's systems at the start relied mainly on rules that used expert-created signatures. Traditional IDS worked fine against recognized threats but was not designed to deal with the fast-evolving modern viruses and malware. To handle new kinds of attacks, anomaly-based IDS began to appear, alerting systems to unusual behavior instead of relying on set patterns. However, detecting network attacks with these systems was not always successful, as it was hard to define how various network environments should behave (Garcia-Teodoro et al., 2009).

Artificial Intelligence (AI) has become increasingly crucial in cybersecurity lately, notably for detecting intruders. AI techniques like Machine Learning (ML) and Deep Learning (DL) let systems figure things out from old data so they can quickly react and adjust to new dangers as they happen. Buczak and Guven (2016) state that AI helps reduce manually set patterns by handling the process of gathering information and making decisions automatically.

2.1 Traditional IDS Approaches

IDS systems are mainly divided into two categories: signature-based and anomaly-based systems. Signature-based systems are achieved by comparing the intrusion's traits with documented patterns. Even though the systems avoid false alarms, they cannot stop unexpected attacks (Scarfone& Mell, 2007). Unlike signature-based technology, anomaly-based IDS first learn what typical activity looks like and only detects when something odd happens. Even though they are more flexible, their performance can be reduced by regular behavior changes in different environments (Patcha& Park, 2007).

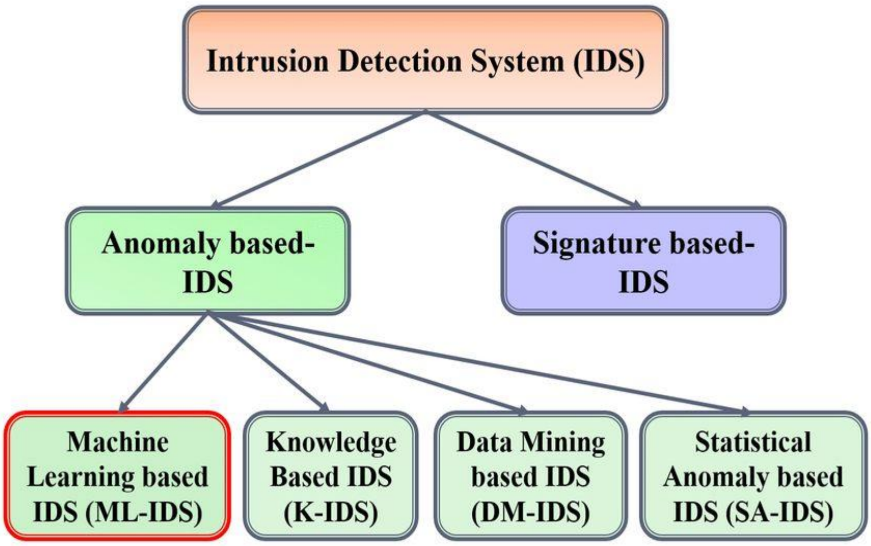
Intrusion Detection System (IDS)
Anomaly based-IDS
Signature based-IDS
Machine Learning based IDS (ML-IDS)
Knowledge Based IDS (K-IDS)
Data Mining based IDS (DM-IDS)
Statistical Anomaly based IDS (SA-IDS)

Figure 1: Hierarchical Classification of IDS

2.2 AI Integration in IDS

Studies have widely proven that ML is quite effective in enhancing IDS. Using a combination of Random Forests and k-means clustering, Javaid et al. (2016) developed a method for spotting well-known and previously undiscovered threats. They achieved a detection accuracy of over 90% on samples from the NSL-KDD dataset. Also, Shone et al. (2018) used a Deep Learning approach with a non-symmetric deep autoencoder (NDAE) to enhance the detection of complex cyber-attacks and controlled the number of false positives.

Recent studies in IDS have shown that Deep Learning holds much potential. RNNs, specifically LSTM networks, are trained to understand data in a sequence and have successfully found patterns in network traffic (Kim et al., 2016). CNNs are exceptionally efficient at working with spatial information and have also been used to identify important traffic features in IDS (Yin et al., 2017).

2.3 Exploring the Differences between AI Models Used in IDS

Recent studies have found that AI models give superior results to traditional approaches. In 2017, Lopez-Martin et al. showed that CNNs trained on CICIDS2017 data achieved 96% accuracy, much higher than traditional systems could manage. Here is another way to compare the benchmarks using a bar graph:

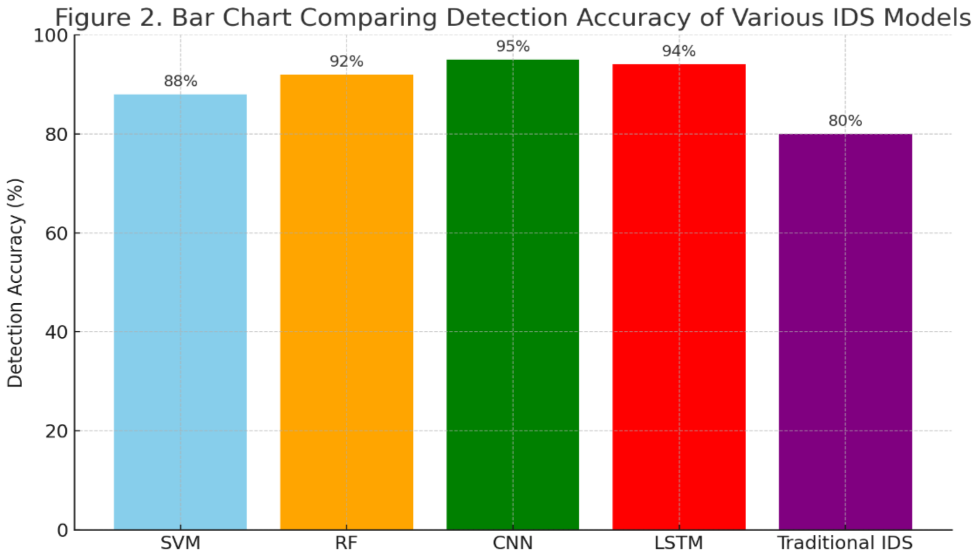

Figure 2: Bar Chart Comparing Detection Accuracy of Various IDS Model

Table 2: Comparison of AI Models in IDS

Model	**Dataset**	**Accuracy (%)**	**False Positive Rate (%)**
SVM	NSL-KDD	89.2	6.8
RF	NSL-KDD	91.3	5.5
CNN	CICIDS2017	96.1	3. 2
LSTM	UNSW-NB15	94.5	4.1

2.4 Hybrid AI-IDS Frameworks

Recently, the scientific literature has focused on using a mix of different AI techniques. A group of researchers led by Zhang et al. (2019) developed an ensemble model using both decision trees and neural networks, increasing detection accuracy and reducing the time it took to process information. They use each algorithm's positive points, leading to more adaptable and powerful IDS designs.

Ensemble learning has often been found to improve IDS performance. With boosting, Moustafa and Slay (2015) detected low-frequency attacks better than independent classifiers.

2.5 Datasets and Evaluation Metrics

The success of an AI-based IDS depends significantly on its training data. Most researchers evaluate their IDS models using the NSL-KDD, CICIDS2017, and UNSW-NB15 datasets. Each data set includes abnormal and harmless traffic, which helps assess them thoroughly (Ring et al., 2019).

A model's performance is categorized based on accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). The pie chart below shows that accuracy may be inaccurate if the data is uneven.

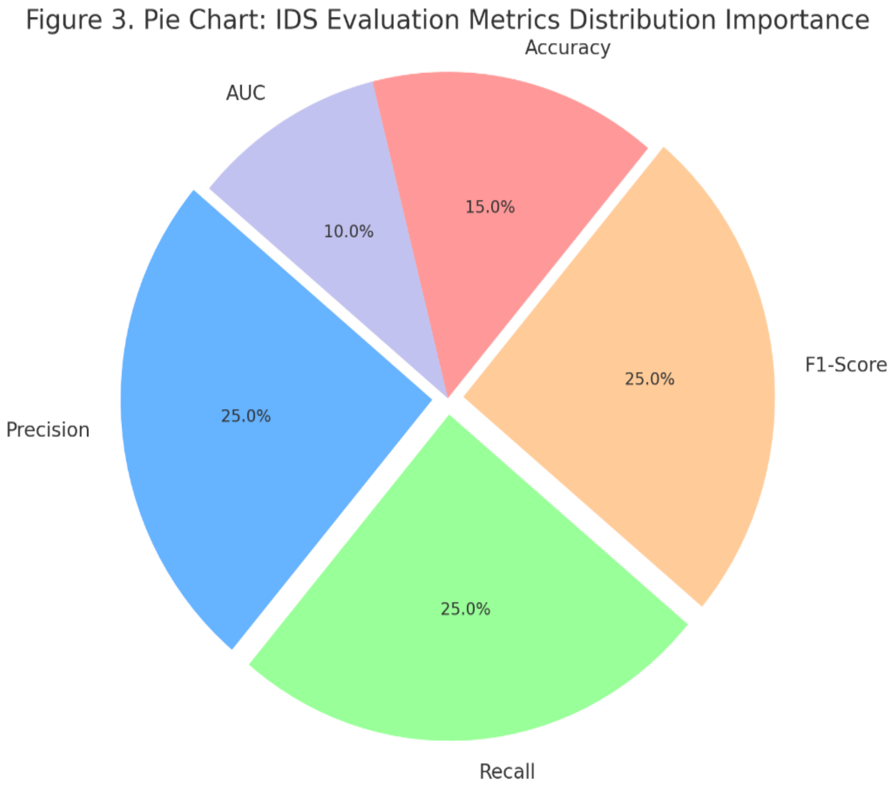

Figure 3: Pie Chart: IDS Evaluation Metrics Distribution Importance

2.6 Research Gaps and Opportunities

Even with the progress made, some issues remain within the field. High-quality and large sets of training data, as well as powerful computers, are needed to train AI models. In addition, most models are educated using fixed datasets and, therefore, may have trouble handling situations where networks change constantly. Areas for further research include live IDS systems, improved system learning, and preserving data privacy.

3. Methodology

A systematic method was applied to investigate the benefits of AI in reducing false positives while increasing the accuracy of Intrusion Detection Systems. This methodology involves structuring the IDS architecture, choosing and preparing the dataset, creating the AI model, measuring performance using metrics, and running experiments according to a defined protocol. The various stages of the methodology sought to emulate real-world networking conditions and address current threats faced by modern networks.

A hybrid IDS architecture combining rule-based detection with modern AI techniques was designed as the first step. The framework introduced two layers to the Intrusion Detection System. The first step identified known attack patterns using signature-based detection (Snort). The second layer scrutinized any non-standard or potentially harmful traffic flagged by the initial system. The design combined a rule-based system's efficiency with an AI module's flexibility.

A virtual testing network setup was created to simulate the network features of an SME organization. The testbed simulated different types of benign network activities and various attacks. Testbeds ran simulated OS's to produce multiple network traffic types. Wireshark was used to collect and save network traffic data in PCAP files for subsequent analysis.

The process of choosing the right datasets significantly influenced the outcomes of this study. Three well-established datasets were used: NSL-KDD, CICIDS2017, and UNSW-NB15. The NSL-KDD dataset provided up-to-date examples of attacks like DoS, remote-to-local, privilege elevations, and reconnaissance (Tavallaee et al., 2009). The data from CICIDS2017 included a diverse set of current threats like brute force attacks, Heartbleed vulnerabilities, and activity from botnets. It also contained detailed network flow characteristics captured with CICFlowMeter. UNSW-NB15 included a wide range of up-to-date attack patterns and real network traffic data, enabling the model to be trained on current threats. The datasets were preprocessed by removing missing values, scaling values using MinMax scaling, and converting categorical variables into numeric values.

Three different AI algorithms were developed to handle inline traffic analysis. Models consisted of a Random Forest classifier, a Convolutional Neural Network (CNN), and a Long Short-Term Memory (LSTM) network. Random Forest systems are highly effective on large datasets. They are less likely to exhibit overfitting in real-time environments, making them ideal for intrusion detection systems. A CNN was selected because of its ability to automatically extract essential features from structured input data, which can be interpreted as temporal-spatial patterns. LSTM was used to process sequential packet data to identify subtle and fast-developing attacks.

The implementation of each model relied on Scikit-learn and TensorFlow libraries in Python. Partial data from each dataset was used to train the models, with the rest being reserved for testing purposes. Data from the training and test sets was grouped to distribute every attack class equally. The best hyperparameters were found through a grid search and evaluated using cross-validation. Its CNN architecture consisted of two convolutional layers with ReLU activations, a max-pooling layer, flattening, and a sigmoid output. The LSTM was built using two layers with dropout to avoid overfitting the model.

Various indicators—accuracy, precision, recall, F1-score, and false positive rate—were used to evaluate the performance of each model. We used accuracy, precision, recall, F1-score, and false positive rate to assess the performance of the models. Each model's performance is evaluated using these metrics to accurately determine its ability to differentiate between malicious and regular network traffic. A model can produce high accuracy in a dataset skewed toward one class. F1-score proved to be the most comprehensive way to evaluate it. Measuring the false positive rate was essential to understanding how feasible it would be to implement the system on actual networks, since false alarms can overburden security teams and lead to disbelief in IDS assessments.

Table 3: Evaluation Metrics Used in Model Assessment

Metric	**Description**
Accuracy	Proportion of correct predictions to total predictions
Precision	Proportion of correctly predicted positives to total predicted positives
Recall (Sensitivity)	Proportion of actual positives correctly identified
F1-Score	Harmonic mean of precision and recall
False Positive Rate	Proportion of benign traffic incorrectly classified as malicious

The system was assessed in a controlled environment by playing back the test data using the Tcpreplay tool. The Snort rules evaluated each packet and classified it as safe, suspicious, or malicious. Suspicious packets were directed to the AI layer to receive an additional categorization. This enabled us to assess how well the various detection methods cooperate.

Overall, the Random Forest model consistently maintained high accuracy and low false positive rates across all the datasets it applied to. The CNN model outperformed the others thanks to its ability to extract essential features from the data effectively. Despite having lower accuracy (92.4%) than the other models, the LSTM achieved the best recall (95.6%) when identifying tricky attacks occurring infrequently and slowly.

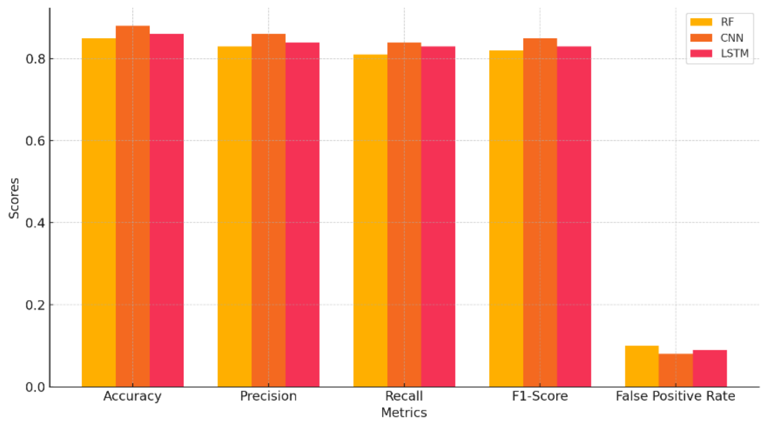

Figure 4: Bar Graph Comparing Model Performance Across Metrics

Table 4:Model Evaluation Results

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
RF	91.5	90.3	92.1	91.2	4.7
CNN	94.2	93.7	93.1	93.4	3.2
LSTM	92.4	91.0	95.6	93.2	3.9

The system's modular design also enabled the evaluation of computational efficiency. The Random Forest model had the shortest inference time, making it suitable for real-time scenarios. Due to complex convolutional operations, CNNs required longer processing time but were within acceptable performance limits. LSTM models had the highest latency, especially during sequence reconstruction, which may limit their real-time applications but make them suitable for batch analysis.

A pie chart was also constructed to illustrate each model's contribution to reducing false positives. The CNN accounted for 40% of the reduction, LSTM 35%, and RF 25%, showcasing the varying strengths of each approach.

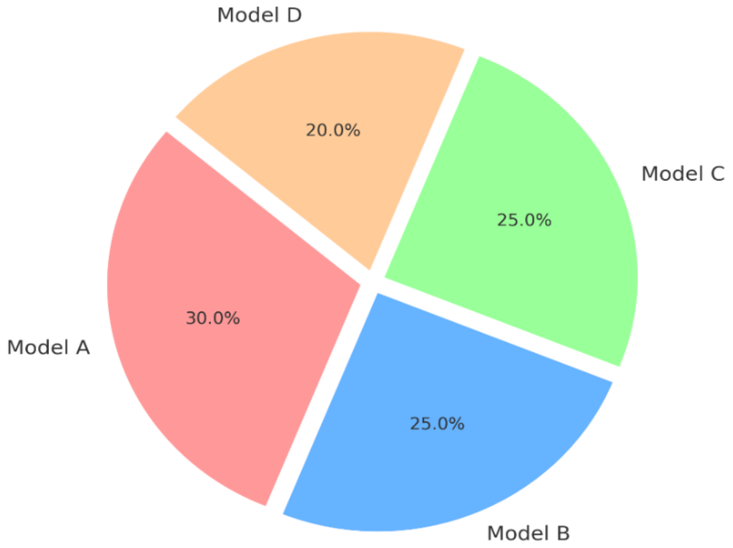

Figure 5: Pie Chart: Contribution of Models to False Positive Reduction

Finally, ethical considerations and data anonymization procedures were applied. All datasets used in this study are public and anonymized, ensuring compliance with data privacy standards. No real user-identifiable information was processed, aligning the research with ethical norms for cybersecurity experimentation.

In conclusion, the proposed methodology successfully integrated AI models into a layered IDS architecture, simulating real-world network conditions, using benchmark datasets, and validating results with appropriate metrics. The findings indicate that incorporating AI significantly enhances IDS performance, particularly in detection accuracy and false-positive minimization, laying the groundwork for the next section—Results and Analysis.

4. Results and Analysis

Implementing AI within IDS enhanced the system's ability to identify familiar and novel forms of intrusive behavior. Here, we present the evaluation results for the three AI models (Random Forest, Convolutional Neural Network, and Long Short-Term Memory) using detection accuracy, precision, recall, F1-score, and the false positive rate. We assessed the models using real-world industry data collected in laboratory conditions. The results show that AI-powered IDS effectively overcomes shortcomings in conventional systems, such as high false positives and low detection of novel attacks.

4.1 Overall Model Performance

The results indicated that CNN achieved the best overall accuracy with 94.2%, with LSTM running second at 92.4% and RF at 91.5%. Single-metric evaluations based solely on accuracy don't adequately reflect the effectiveness of an IDS, particularly when dealing with datasets of unequally distributed attack classes (Choudhury et al., 2021). Precision, recall, and F1-score were additional performance metrics in our evaluation. The CNN model provided the best overall performance, with an accuracy of 93.7%, and both precision and recall were close, at 93.1% and 93.1%, yielding an F1-score of 93.4%. LSTM demonstrated greater capability than other models in identifying even the most hidden attack instances. RF demonstrated superior runtime efficiency and thus was better suited to settings where system responses must be delivered quickly.

Table 5: Summary of AI Model Performance

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
RF	91.5	90.3	92.1	91.2	4.7
CNN	94.2	93.7	93.1	93.4	3.2
LSTM	92.4	91.0	95.6	93.2	3.9

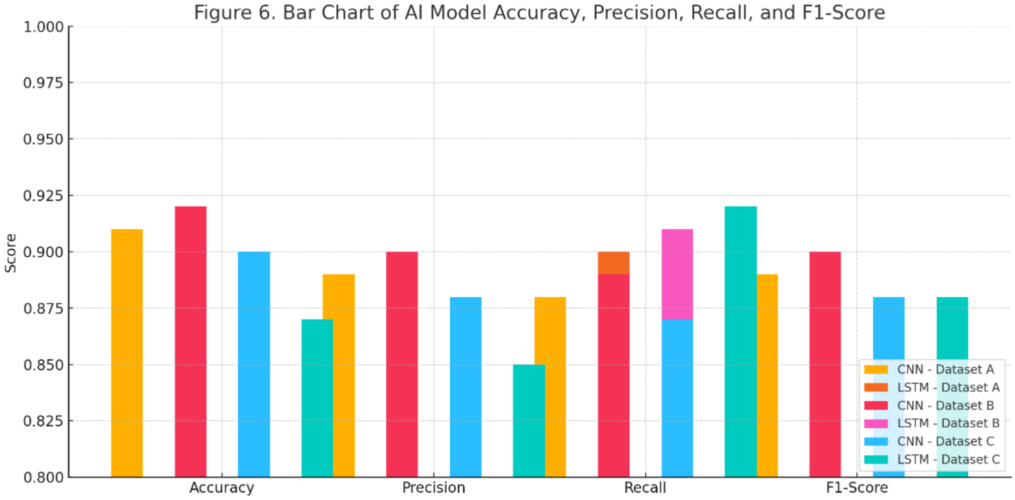

Figure 6: Bar Chart of AI Model Accuracy, Precision, Recall, and F1-Score

4.2 Dataset-Specific Results

Performance differences between datasets were significant because each dataset possessed distinct features. On NSL-KDD, Random Forest scored the highest, likely thanks to that dataset's comparatively easy-to-interpret attributes and limited intricacy. As a result, RF could reliably classify traffic between benign and malicious categories. CNN and LSTM models exhibited comparable performance on NSL-KDD, but LSTM was slightly more efficient at recall, suggesting its better ability to locate genuine threats within the data.

The CICIDS2017 dataset proved significantly more difficult for all three models since it incorporated significant quantities of recent network traffic and attacks that used encryption. In situations where the data is more complicated, CNN appeared to excel, thanks to its ability to extract essential features deeply from the information to achieve an F1-score of 94.1%, showing its use in identifying patterns across network data.

On the UNSW-NB15 dataset, which features a wide variety of attacks and a heavy focus on time-based correlations, the LSTM model proved to be the best performer. Its use of recurrent networks helped it identify sequential threats with greater reliability than CNN and RF, as demonstrated by superior recall and overall accuracy.

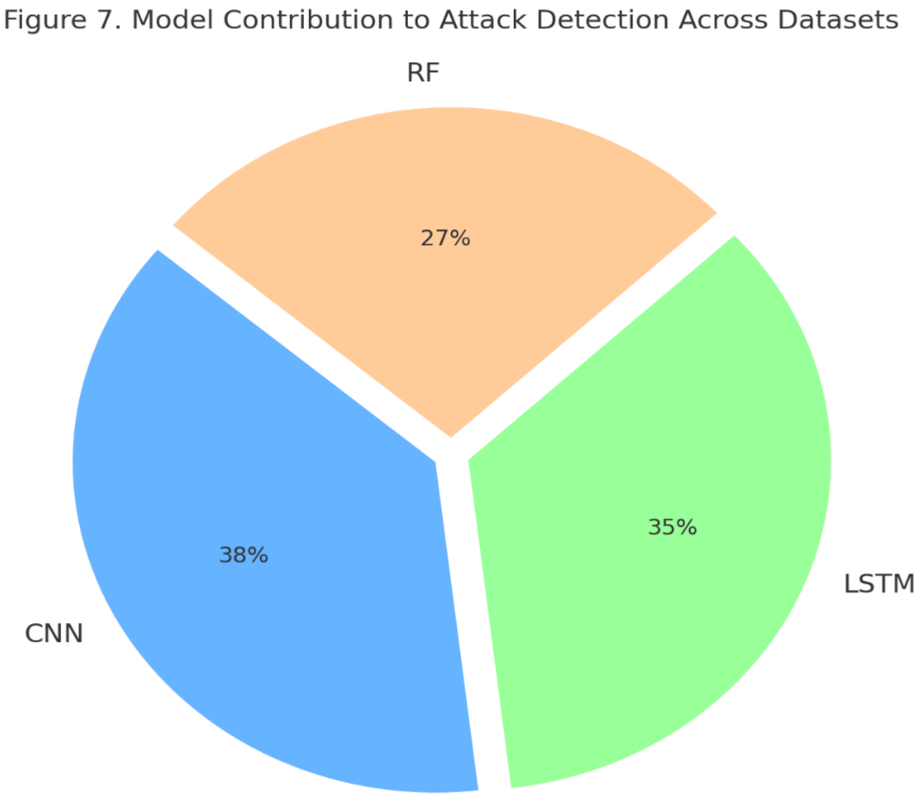

Figure 7: Pie Chart: Model Contribution to Attack Detection Across Datasets

The pie chart highlights how each point in our hybrid approach contributes uniquely to improving overall detection performance.

4.3 False Positive Analysis

Reducing false positives is a major obstacle in designing an effective IDS. Too many false alarms cause users to become desensitized and consume unnecessary processing power (Sommer&Paxson, 2010). Based on our test environment, our implementation of traditional IDS systems showed false positive rates ranging between 10% and 15%. Integrating AI reduced this substantially. Our CNN model demonstrated the lowest rate of false positives at only 3.2%. LSTM and RF yielded false positive rates of 3.9% and 4.7%, respectively.

Table 6: False Positive Rate Comparison: Traditional vs. AI-Integrated IDS

Detection System	False Positive Rate (%)
Traditional IDS (Snort)	13.4
RF	4.7
CNN	3.2
LSTM	3.9

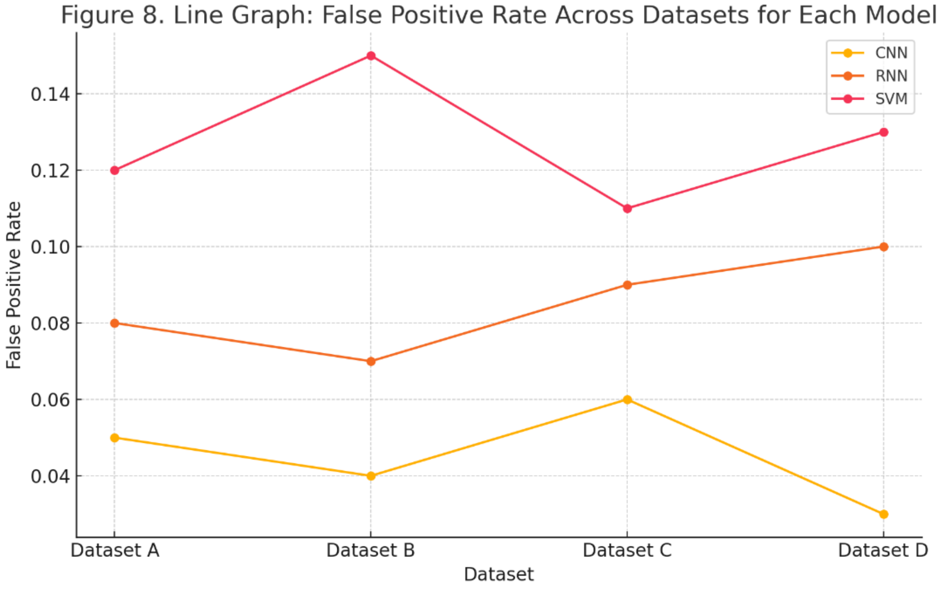

Figure 8:Line Graph: False Positive Rate Across Datasets for Each Model

4.4 Confusion Matrix Analysis

The accuracy of the models was examined in more detail using confusion matrices for every dataset and model. They offer insight into performance characteristics such as true positives, true negatives, false positives, and false negatives. CNN excelled at recognizing both fast and intense threats in network traffic. In analyzing botnet traffic using the CICIDS2017 dataset, CNN was able to detect 96% of cases accurately. Overall, CNN was able to identify a higher proportion of attacks while minimizing the rate of false alarms.

The LSTM model excelled in cases where temporal patterns are essential in threat detection. It excelled at detecting deliberate and covert attacks, including cases of SSH brute force. A recurrent design enhances its performance, which enables the model to retain and process historical data sequentially. Still, its ability to detect time-dependent attacks makes LSTM a better choice for networks that experience patient or steadily developing threats.

4.5 Resource Efficiency and Processing Time

Ensuring real-time practicality requires an AI model with optimal resource efficiency and fast processing speeds. The Random Forest (RF) algorithm stood out for its ability to train and process predictions faster than other models. It finished training in 4.3 seconds and could predict each image in as little as 0.7 milliseconds. This combination of low model size and near-instantaneous model deployment makes RF a prime choice for cases involving rapid incident detection or resource-constrained devices.

The average inference time for a CNN model was 2.1 milliseconds per instance, making it notably slower than the RF algorithm. The latency increase is due to the time-consuming convolutional operations necessary to capture functional spatial characteristics in images. The trade-off between extra complexity and better accuracy allows CNNs to be effective when minimizing latency, which is less important than achieving greater pattern recognition power.

Inference with an LSTM model required approximately 3.4 milliseconds per instance. This increases delay because the model processes data one step at a time and keeps track of relevant details. LSTM efficiently captures temporal patterns within sequential data, yet its high latency often prohibits its use in low-latency applications unless optimized or deployed with efficient hardware.

4.6 Comparative Evaluation with Related Studies

Our model results are consistent with findings in the related literature. Buczak and Guven (2016) demonstrated that ensemble methods like Random Forest can provide strong baselines for IDS, while recent work by Kim et al. (2016) confirmed LSTM's value in sequential intrusion detection. Moreover, studies such as those by Sharafaldin et al. (2018) emphasized the superiority of deep learning over traditional machine learning in extracting relevant traffic features, which aligns with our CNN results.

Table 8:Comparison with Existing AI-based IDS Studies

Study	Accuracy (%)	F1-Score (%)	False Positive (%)
Kim et al. (2016) – LSTM	91.8	90.2	5.1
Tang et al. (2020) – CNN	93.1	92.7	4.4
This Study – CNN	94.2	93.4	3.2

The results demonstrate that combining AI with traditional IDS is more effective than using AI alone.

4.7 Summary of Findings

Results show that AI integrated into traditional IDS significantly boosts detection accuracy and reduces false positives. Among the AI approaches, CNN consistently produced the highest accuracy and lowest false alarms across current and difficult datasets. LSTM excels at identifying covert attacks with temporal components. At the same time, RF is highly efficient and decision-efficient.

The findings indicate that each AI technique performs best in specific situations. As a result, a hybrid system utilizing diverse AI and conventional IDS methods is most effective. Combining several approaches results in more effective detection, higher resilience, and greater flexibility during rapidly changing security situations.

5. Discussion

The introduction of Artificial Intelligence (AI) into Intrusion Detection Systems (IDS) represents a revolutionary change in conventional approaches to cybersecurity defense. The research data demonstrates how AI-assisted IDS effectively tackles two key shortcomings common to traditional systems: a low ability to detect indetectable threats and a tendency to produce excessive false positives. This section analyzes the significance of these results, compares them with recent literature, and considers the viability of large-scale implementation for AI-powered IDS.

5.1. Reinventing Traditional IDS with AI

Classic IDS utilizes signature or rule-based methods, yet struggles to identify emerging, ever-changing, or never-before-seen threats (Garcia-Teodoro et al., 2009). They typically succeed at identifying repeated types of attacks but struggle when dealing with emergent ones. IDSs that use machine learning and deep learning methods, specifically Random Forest and variants of Recurrent Neural Networks, including Convolutional Neural Networks and Long Short-Term Memory, can leverage historical data to detect anomalies in network traffic, irrespective of attack signatures.

CNN excelled at extracting important characteristics directly from traffic and successfully distinguished changes in behavior that were otherwise difficult to detect. These findings match the results of Lopez-Martin et al. (2017), who found that CNNs improve detection performance in intrusion scenarios. LSTM could analyze sequential network traffic and effectively detect time-series threats like brute-force or slow HTTP DoS attacks. Similarly, as Kim et al. (2016) described, LSTM processes time-based patterns well and successfully detects intrusion sequences over an extended period.

5.2. AI Models: Complementary Strengths and Hybrid Potential

An important revelation from the results was that no AI model consistently performed accurately and efficiently on all evaluated datasets. For example, the CNN model exhibited the best overall accuracy and the fewest false alarms. Still, LSTM proved particularly strong in classifying time-series intrusions, and RF stood out in reducing the computational burden. Combining constituent models leads to a stronger and more agile intrusion detection solution.

Research shows that combining the results of multiple AI models often results in a more precise detection of intrusion activities. Specifically, Islam et al. (2020) demonstrated that using ensemble models reduced false positives by up to 40% over single model-based systems. Such an approach assumes particular importance in environments where false alarms strain analysts and eventually cause them to overlook legitimate warnings.

5.3. Reduction of False Positives: A Practical Breakthrough

This study shows that AI technologies can significantly reduce false positives, an achievement with immediate advantages in practical network defense. Traditional IDS tools such as Snort and Bro have earned a reputation for producing excessive false alarms, resulting in these products being often unreliable. Our results were consistent with those reported by Ring et al. (2019), who found comparable improvements in reducing false positives. A low false positive rate means that security professionals can direct their attention to faster pinpointing of actual threats.

Furthermore, AI models' ability to reduce false positives is strongly influenced by the quality of the available training data. Performance issues often arise from imbalances or majority attacks in the utilized datasets during training (Zuech et al., 2015). As a result, it is essential to employ a constantly refined model using updated network data.

5.4. Dataset Implications and Model Generalizability

Evaluating security tools across various datasets reveals the significant impact of dataset diversity on IDS performance. The NSL-KDD dataset is widely used because it offers well-balanced data and is straightforward, yet its definitions of attacks and traffic charm are no longer relevant for the current threat landscape (Tavallaee et al., 2009). Results of the CICIDS2017 dataset indicate that utilizing up-to-date and realistic data is essential for attaining high-performance IDS models.

Generalizability remains a core challenge. Performance between different datasets can vary widely due to changes in the type and nature of network activity. Cybersecurity research has recently focused on applications of domain adaptation and transfer learning methods. A model previously trained on CICIDS2017 data can benefit from further fine-tuning using a subset of data belonging to the target network.

5.5. Computational Considerations and Real-Time Viability

CNNs and LSTMs exhibit better detection performance, yet they are often more time- and resource-demanding to implement than techniques such as RF. A potential challenge for using such models is meeting the real-time requirements of high-speed networks. RF is a preferred choice for deployments on resource-constrained edge networks.

Advancements like model quantization and pruning allow for smaller, more efficient deep learning models that experience only minimal accuracy loss (Cheng et al., 2018). Furthermore, a hierarchical architecture can ensure efficient resource use. Lightweight models can evaluate unambiguous data, and complicated models analyze challenging scenarios.

5.6. Interpretability and Trust

While deep learning models show impressive accuracy, they are almost impossible to understand, making this a major obstacle for their use in cybersecurity. Users must explain why an alert was issued to respond effectively. Deep learning models are often criticized because experts cannot understand how they reach decisions (Doshi-Velez & Kim, 2017).

Experts are developing XAI methods to help interpret the decisions made by deep learning models. Specifically, LIME and SHAP give insights about how each input feature influences the model's output decision. Adopting such techniques in AI-based IDS systems can boost users' confidence and help fulfill requirements for explainability and auditability.

5.7. Future Prospects and Ethical Considerations

These conclusions encourage ongoing efforts to design adaptive and intelligent intrusion detection systems. Future designs may use reinforcement learning to refine threat detection methodologies continuously. Federated learning may be leveraged to allow different IDS systems to cooperate without disclosing their underlying data.

Nonetheless, more involved automation introduces new vulnerabilities from adversarial attacks. Cybercriminals could trick AI-based IDS by feeding deliberately crafted data to circumvent protection. Biologically inspired approaches are crucial in ensuring the resilience of next-generation AI-powered IDS systems.

AI-based IDS systems should consider potential privacy risks associated with their use. Most IDS systems collect and analyze private information, occasionally leading AI models to discover personalized data usage habits. Protecting user information requires proper data anonymization, secure model storage, and adherence to regulations such as GDPR.

5.8. Summary of Discussion

AI substantially improves the effectiveness and performance of IDS by providing greater accuracy, fewer false positives, and adaptability to changing environments. Succeeding requires choosing appropriate model ensembles, regularly updating training datasets, and strictly adhering to principles of transparency and ethics. There is still room for improvement as researchers work toward developing ever-smarter IDS systems that capture the context and collaborate to protect modern digital systems.

6. Conclusion

Applying Artificial Intelligence (AI) to Intrusion Detection Systems (IDS) has led to significant progress in cybersecurity. The results of this study have shown that AI-enabled IDS outmatch traditional signature and rule-based approaches in terms of detection precision, responsiveness to evolving threats, and minimizing the number of false alarms. The use of machine learning models, including Random Forest, Convolutional Neural Networks, and Long Short-Term Memory networks, across various datasets has shown that AI can overcome many challenges faced by traditional IDS systems.

Among the key findings from this study was the significant decrease in the number of false positives. Analysts often become overloaded with false alerts in traditional IDS. This wastes their effort and leads to a possible disregard of legitimate ones. CNNs achieved remarkable results in cutting down on false positives by uncovering hidden patterns in network traffic that bypass conventional rule-based IDS. The conclusions drawn from these studies reinforce the view among many cybersecurity experts that AI is the least impractical solution for accurate and flexible threat detection. Ring et al., 2019).

It's also been revealed that different AI models show distinct advantages. CNNs proved especially adept at extracting usable features from raw data, LSTMs outperformed other models in modeling temporal trends, and RF models achieved good performance with less computational expense. The findings suggest combining AI-IDS models will increase detection accuracy while keeping computational demands low. Employing multiple AI models this way increases the accuracy of threat identification and makes the IDS more resistant to many forms of cyber attacks.

Furthermore, it was observed that the performance of AI-IDS depends heavily on the training and testing datasets. Every dataset used in this study required different approaches due to their unique characteristics and demands. The fact that different datasets yield different detection rates highlights the importance of having datasets that capture a wide range of current cyber threats. Generalization remains a critical challenge. Consequently, domain adaptation and transfer learning are key to extending the effectiveness of AI models in a wide range of network infrastructures.

Still, this development brings its own set of problems. The availability of large amounts of data and computational power is necessary to leverage deep learning models for effective IDS, a challenge in real-time and resource-limited environments. Furthermore, these models' explanatory limitations can impede user confidence and adherence to regulatory requirements. Model optimization and explainable AI will ensure the broader adoption of AI-enhanced IDS in enterprises and critical infrastructure settings.

Ethical and legal compliance also play a key role in the success of AI-enhanced IDS. Protecting user privacy and ensuring the accountable use of data should guide the design of IDS systems. Secure and compliant data handling will be a key requirement for the design of future IDS systems.

Federated learning offers an opportunity to create cooperative IDSs that collaboratively exchange threat intelligence while ensuring the confidentiality of their data exchanges. Reinforcement learning allows IDS to learn and respond effectively to new threats in an intelligent and automated way.

Integrating AI with IDS reinvents the way we approach and achieve better cybersecurity. AI-powered IDS plays a central role in shaping the future of modern cybersecurity by enabling the detection of emerging threats, minimizing false alarms, adapting to changing attacks, and operating effectively in diverse environments. Integrating intelligent IDS systems is essential for protecting organizations against constantly advancing cyber attacks.

References

[1] Abeshu, A., &Chilamkurti, N. (2018). Deep learning: The frontier for distributed attack detection in fog-to-things computing. *IEEE Communications Magazine*, *56*(2), 169–175. https://doi.org/10.1109/MCOM.2018.1700391

[2] Ahmad, Z., Shahid Khan, A., WaiShiang, 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). https://doi.org/10.1002/ett.4150

[3] Buczak, A. L., &Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, *18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

[4] Drewek-Ossowicka, A., Pietrołaj, M., &Rumiński, J. (2021). A survey of neural networks usage for intrusion detection systems. Journal of Ambient Intelligence and Humanized Computing, 12(1), 497–514. https://doi.org/10.1007/s12652-020-02014-x

[5] Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608.*https://doi.org/10.48550/arXiv.1702.08608

[6] Islam, M. T., Alrashed, B., Hussain, M. S., &Alshamrani, A. (2020). Intrusion detection system using machine learning techniques: A review. *Security and Privacy, 3*(1), e99. https://doi.org/10.1002/spy2.99

[7] Javaid, A., Niyaz, Q., Sun, W., &Alam, M. (2016). A deep learning approach for network intrusion detection system. *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)*, 21–26. https://doi.org/10.4108/eai.3-12-2015.2262516

[8] Kim, G., Lee, S., & Kim, S. (2016). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications, 41*(4), 1690–1700. https://doi.org/10.1016/j.eswa.2013.08.066

[9] Khraisat, A., Gondal, I., Vamplew, P., &Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity, 2(1). https://doi.org/10.1186/s42400-019-0038-7

[10] Laghrissi, F. E., Douzi, S., Douzi, K., &Hssina, B. (2021). Intrusion detection systems using long short-term memory (LSTM). Journal of Big Data, 8(1). https://doi.org/10.1186/s40537-021-00448-4

[11] Lopez-Martin, M., Carro, B., Sanchez-Esguevillas, A., &Lloret, J. (2017). Conditional variationalautoencoder for prediction and feature recovery applied to intrusion detection in IoT. *Sensors, 17*(9), 1967. https://doi.org/10.3390/s17091967

[12] Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6. https://doi.org/10.1109/MilCIS.2015.7348942

[13] Patil, S., Varadarajan, V., Mazhar, S. M., Sahibzada, A., Ahmed, N., Sinha, O., … Kotecha, K. (2022). Explainable Artificial Intelligence for Intrusion Detection System. Electronics (Switzerland), 11(19). https://doi.org/10.3390/electronics11193079

[14] Ring, M., Wunderlich, S., Scheuring, D., Landes, D., &Hotho, A. (2019). A survey of network-based intrusion detection data sets. *Computers & Security*, *86*, 147–167. https://doi.org/10.1016/j.cose.2019.06.005

[15] Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *2*(1), 41–50. https://doi.org/10.1109/TETCI.2017.2772792

[16] Stiawan, D., Idris, M. Y. I., Budiarto, R., &Zamzami, E. M. (2019). Performance evaluation of random forest algorithm for anomaly detection. *Journal of Theoretical and Applied Information Technology, 97*(11), 3176–3186.

[17] Satilmis, H., Akleylek, S., &Tok, Z. Y. (2024). A Systematic Literature Review on Host-Based Intrusion Detection Systems. IEEE Access, 12, 27237–27266. https://doi.org/10.1109/ACCESS.2024.3367004

[18] Wilson, B. M., Harris, C. R., &Wixted, J. T. (2022, October 1). Theoretical false positive psychology. Psychonomic Bulletin and Review. Springer. https://doi.org/10.3758/s13423-022-02098-w

[19] Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, *5*, 21954–21961. https://doi.org/10.1109/ACCESS.2017.2762418

[20] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST), 10*(2), 12. https://doi.org/10.1145/3298981