International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

A comparative analysis of Network Intrusion Detection (NID) using Artificial Intelligence techniques for increase network security

MD Shadman Soumik *

Department of Information Technology (MSIT), Washington University of Science and Technology (WUST), 2900 Eisenhower Ave, Alexandria, VA 22314.

International Journal of Science and Research Archive, 2024, 13(02), 4014-4025

Publication history: Received on 19 November 2024; revised on 26 December 2024; accepted on 28 December 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.13.2.2664

Abstract

Network Intrusion Detection Systems (NIDS) are critical components of modern cybersecurity frameworks, designed to detect and mitigate malicious activities within networks. This study explores the application of Artificial Intelligence (AI) techniques, including Machine Learning (ML) and DL, for improving network security through accurate intrusion detection. Using the CIS-CICIDS2017 dataset, a comprehensive preprocessing pipeline involving data cleaning, SMOTE-based balancing, Min-Max normalization, and feature selection was employed. The Random Forest (RF) model demonstrated superior performance with an accuracy99.90%, precision97.78%, recall97.08%, and an F1-score97.41%. Comparative analysis with Decision Tree (DT), Stacked LSTM, and AdaBoost models highlighted RF's robustness in detecting and classifying network traffic. Future research aims to optimize feature engineering and explore hybrid AI models for improved real-time intrusion detection in dynamic network environments.

Keywords: NIDS; Artificial Intelligence; Machine Learning Algorithm; Random Forest Algorithm; Cyber Security; CICIDS2017

1. Introduction

In today's interconnected world, uninterrupted network access is vital for both business operations and personal life. The rapid proliferation of electronic devices connected to the Internet has created an expansive attack surface, providing malicious actors with unprecedented opportunities to exploit vulnerabilities[1][2]. The challenge lies in effectively defending networks from both known and emerging threats, especially as the number of attacks continues to rise each year. NIDS have become an indispensable component of modern cybersecurity strategies, serving as the first line of defense against hostile activities on Information and Communication Technology (ICT) networks[3][4].

NIDS are designed to continuously monitor network traffic and identify suspicious behavior that may compromise the confidentiality, integrity, or availability of network resources[5]. These systems utilize sophisticated algorithms to analyze traffic patterns and detect anomalies, allowing them to respond promptly to potential threats[6]. Intrusion detection extends beyond external attacks to include internal misuse, ensuring comprehensive protection of network assets[7][8]. While traditional IDS relied on predefined rules and signatures, advancements in Artificial Intelligence (AI) have revolutionized their capabilities, enabling them to adapt to evolving threats and detect previously unknown attack vectors[9].

An integration of AI, particularly ML and DL, has transformed the landscape of intrusion detection[10]. The categorisation of regular and problematic activity is made possible by ML approaches that use feature engineering to identify meaningful patterns by network data. On the other hand, DL models, with their deep architectures,

* Corresponding author: MD Shadman Soumik.

automatically learn complex features from raw data, eliminating the need for manual feature engineering[11][12][13]. Despite their effectiveness[14], DL models are often criticized for being "black box" systems due to their opaque nature and vast number of parameters.[15][16] Nevertheless, the growing computational power of GPUs has made it feasible to deploy ML and DL-based IDS, resulting in remarkable accuracy and performance in detecting intrusions[17][4].

As network security becomes a top priority for organizations, the role of AI-driven NIDS continues to expand. Enterprises now face sophisticated threats ranging from brute force and DDoS attacks to infiltration and insider threats[18][19]. Traditional security measures like firewalls, while effective at blocking external threats, fail to address internal attacks and advanced persistent threats (APTs)[20][21]. This is where AI-enhanced NIDS excel, providing dynamic, adaptive, and robust protection against an ever-evolving threat landscape[22]. By leveraging the power of AI, NIDS are not only capable of mitigating security breaches but also ensuring the resilience of networks in an increasingly digital world[23].

This study's objective is to create an AI-based Network Intrusion Detection (NID) system that enhances network security by accurately detecting and classifying network traffic. The study aims to address common challenges in network intrusion detection, such as class imbalance and complex attack patterns, by applying advanced preprocessing techniques and ML models. The goal is to build a robust solution capable of effectively mitigating cyber threats in diverse network environments. Main contribution of the study is listed below:

- Develops an AI-based NID system that accurately detects and classifies benign and malicious network activities.
- Implements a comprehensive preprocessing pipeline to handle class imbalance, data normalization, and feature selection.
- Applies SMOTE to balance the dataset, ensuring that minority classes are well-represented, thus improving model reliability.
- Uses Random Forest (RF) for effective classification, aggregating predictions to enhance accuracy and reduce model variance.
- Provides an in-depth assessment of network intrusion detection using the CIS-CICIDS2017 dataset, offering insights into anomaly identification in realistic network environments.

Structure of the paper

The format of this document is as follows: ML methods for NID are reviewed in Section II. Section III details the methodology, including the CICIDS2017 dataset and model implementation. Section IV presents experimental results and model comparisons. Section V concludes and future research directions.

2. Literature Review

An overview of earlier research on network intrusion detection (NID) using AI methods to improve network security is provided in this section.

In this study Sah and K, (2024), compared and tested two multiclass classifier algorithms—DR and RF—on the CICIDS-2017 dataset, which is part of the ISCX Consortium's ML CSV data set, for the purpose of anomaly-based intrusion detection in network traffic. Decision Tree classifiers had accuracy scores of 0.99867 respectively, with execution times of 44.3 seconds and 8 minutes and 35 seconds, respectively [24].

In this study Hu et al. (2024), the algorithm and a novel hybrid attention mechanism were both presented. Testing on the UNSW-NB15, CICIDS-2017, andCICIDS-2018 datasets confirmed an accuracy of 100%,99.79%, and98.10% for binary classification tasks, and 96.37%,98.12%, and99.06% for multiclassification problems [25].

In this study Arshad et al. (2023), examined a performance of classifiers on CICIDS-2017dataset using various feature selection/dimension reduction techniques. The best algorithm for a particular attack class was finally determined by comparing accuracy of the overall model for various attack classes, including DDoS, DOS) Web-based, Brute force, intrusions, scans, Bots, and Heartbleed. An outcome demonstrated that, with an accuracy of 99.91%, the XGBoost classifier with Boruta feature selection outperformed the other classifiers [26].

In this study Akoto and Salman, (2022), examined ML and DL intrusion detection and categorisation models. ML and DL models are trained and tested using the public CICIDS-2017 dataset. Three traditional ML models (LR, RF, the KNN) and three DL models (ConvlD, RNN, and a two-staged model) are used. An ANN is used for classification in the model, which is pre-trained using unsupervised DAE. At 99.5%, RF has the highest ML detection accuracy, while DAE-ANN comes in

at 98.7% for DL detection. Stepwise multi-classification is shown to be superior than the conventional single-stage multi-classification. In conclusion, RF outperforms DAE-ANN in classification, recording detection rates of 91.35% and 84.66%, respectively [27].

In this study Atefi, Hashim and Kassim, (2019), used the most recent dataset, CICIDS-2017, to concentrate on anomaly analysis for IDS classification. Abnormality study for classification was carried out by this research utilising KNN for ML and DNN for Deep Learning. One result is the performance of ML and DL in classification as measured by MCC. When comparing the two classifiers, DNN's accuracy was 0.9293% and KNN's was 0.8824% [28].

In this study He et al. (2019), studied an intrusion detection method that uses a hierarchical progressive network topology and is backed by technologies such as MDAE and LSTM. By applying their MDAE model to the characteristics of traffic data, the testing findings reveal that the detection model's performance may be enhanced by a margin of 2% to 5% [29].

Table 1 presents a comparative analysis of existing studies on network intrusion detection (NID) leveraging Artificial Intelligence techniques to enhance network security.

Table 1 Comparative Analysis of Network Intrusion Detection (NID) Using Artificial Intelligence Techniques

Ref.	Objective	Techniques	Dataset	Findings	Limitations & Future Work
[24]	Anomaly-based intrusion detection in network traffic using ML methods.	Decision Trees, Random Forests	CICIDS-2017	Achieved accuracy and F1-Score 99% for random forest model.	Higher computational cost for Random Forest. Future work could explore feature optimization and hybrid approaches for reducing execution time while retaining accuracy.
[25]	Introduced a hybrid attention mechanism with an enhanced algorithm for intrusion detection.	Hybrid attention mechanism, Effective Channel Layer, Curve Space Layer	UNSW-NB15, CICIDS-2017, CICIDS-2018	Accuracy for Binary Classification: 100%, 99.79%, 98.10%; Multiclass classification: 96.37%, 98.12%, 99.06%.	Needs validation on more datasets and real-world deployment scenarios. Future research could examine generalization capabilities and real-time performance.
[26]	Evaluated classifier performance on CICIDS-2017 using feature selection techniques.	XGBoost with Boruta feature selection	CICIDS-2017	XGBoost Accuracy: 99.91% For various attack classes (e.g., DDoS, DOS, Web-based, etc).	Needs exploration of real-time applicability and feature engineering for diverse datasets. Future work could focus on adaptive feature selection methods.
[27]	Compared ML and DL models for intrusion detection and categorization.	LR, RF, KNN, Convolutional 1D-CNN, RNN, Dense Autoencoder (DAE)+ANN	CICIDS-2017	Random Forest achieved 99.5% accuracy, outperforming DAE-ANN (98.7%). In categorization, RF scored 91.35%, while DAE-ANN achieved 84.66%.	DAE-ANN's performance in categorization was lower than RF. Future research could enhance DL models' generalizability and explore ensemble methods to combine ML and DL models for improved results.
[28]	Anomaly analysis for intrusion detection using updated dataset CICIDS-2017.	K-Nearest Neighbors (KNN), Deep Neural Network (DNN)	CICIDS-2017	DNN performed significantly better with MCC score of 0.9293 compared to	Future work could involve optimizing KNN and DNN models for real-time detection and exploring other DL methods like

				KNN with score 0.8824.	transformers for better MCC scores.
[29]	Proposed a multimodal-sequential intrusion detection approach with hierarchical progressive network.	Multimodal Deep Autoencoder (MDAE), LSTM	CICIDS-2017	Achieved 94% accuracy in binary classification and 88% in multi-class, with a 2%-5% improvement leveraging multimodal traffic data.	Requires validation across more datasets. Future research could explore scalability, multimodal data fusion, and improvements in LSTM models for real-time application and handling larger-scale traffic data.

3. Research Methodology

An aim of this study is to create an AI-based network intrusion detection (NID) system to enhance network security by accurately detecting and classifying network traffic. The results of the assessment will help to enhance the methodology's prediction for anomaly identification with the CIS-CICIDS2017 dataset, a comprehensive intrusion detection dataset created in a realistic network environment, simulating both benign and malicious traffic. Data pre-processing begins with cleaning, which involves removing null values, duplicates, and noisy data while converting categorical labels to numerical form. To address the dataset's class imbalance, SMOTE is applied, generating synthetic samples for minority classes to ensure balanced class distribution. Data normalization is performed using Min-Max Scaler, which scales features to a range of [0,1] for consistent input to machine learning models. Feature selection is conducted to reduce dimensionality and enhance model efficiency, with the most important features identified through importance scoring. The dataset is then split into training and testing sets in an 80:20 ratio. A Random Forest (RF) algorithm is used for classification, leveraging an ensemble of DT to predict intrusion. The RF model aggregates predictions from multiple trees to lower variance and enhance accuracy. The methodology incorporates both data-driven pre-processing steps and robust predictive modelling to optimize intrusion detection performance. The following is a full explanation of the data flow diagram phases, as illustrated in figure 1:

3.1. Data Collection

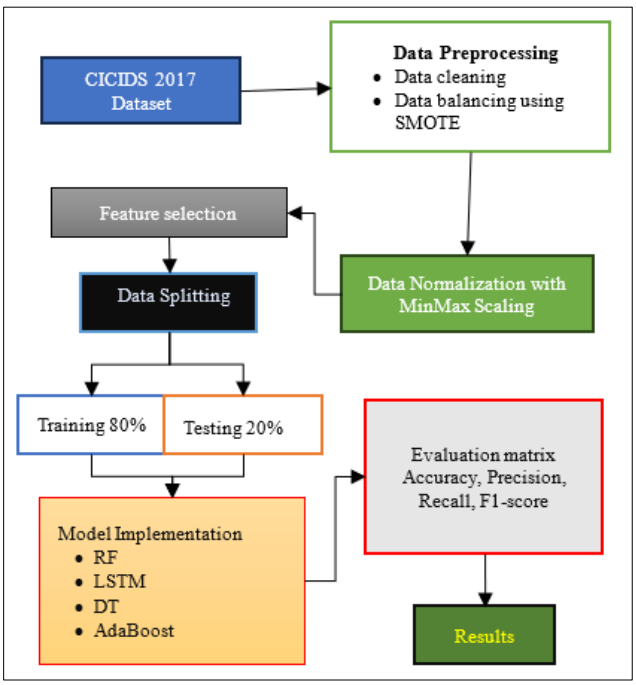

Figure 1 ML model flowchart for network intrusion detection.

The Canadian Institute for Cybersecurity's 2017 CICIDS dataset is a popular tool for researchers interested in intrusion detection. It was built in a realistic network environment with a variety of hardware components and software operating systems, including Ubuntu, Mac OS, Windows, and firewalls. The dataset mimics the actions of 25 people who

inadvertently generate innocuous traffic using protocols including HTTP, HTTPS, FTP, SSH, andemail. Also included are typical attack scenarios from 2016, such as brute force, DoS, DDoS, infiltration, Heartbleed, botnet, and port scan techniques. The dataset, available as a CSV file on the University of New Brunswick's website, is comprehensive but exhibits a significant class imbalance, with over 70% of traffic being benign and some attack types contributing less than 1% of the total.

The detailed explanations of eacg step in a flowchart are provided below.

3.1.1. Data Preprocessing

Data preprocessing entails nothing more than converting raw data into a more comprehensible format. Incomplete, inconsistent, duplicated, or noisy real-world data does occur [30]. Data preprocessing is a series of operations that transforms raw data into a processed and understandable form. The following pre-processing steps are listed given below:

3.1.2. Data cleaning

Data cleaning involves identifying and addressing corrupt, inaccurate, or irrelevant records within a dataset or database[31]v. Data that is missing, inconsistent, noisy, or superfluous may be found and dealt with in this process by altering, deleting, or replacing it. When cleaning the CICIDS-2017 dataset, we remove any rows that include null values, duplicates, or empty cells (such as Inf or NaN). We then transform categorical information, such as labels, into numerical values so that machine learning algorithms may use them[32].

3.2. Data Balancing Using SMOTE

An approach to rectifying unbalanced datasets, SMOTE (Synthetic Minority Oversampling Technique) involves creating synthetic samples for the minority class. It ensures a more uniform distribution of classes by generating new samples by interpolation between existing data points instead of duplicating them. This helps in making ML models more effective at tackling unbalanced classification problems[33].

3.2.1. Data Normalization with MinMax Scaler

Min-Max Scaler is a technique used for normalizing data by scaling features to a specified range, typically [0, 1][34]. The transformation is performed using the formula in equation (1).

$$X` = \frac{x - x_{min}}{x_{max} - x_{min}} \quad \text{.....(1)}$$

Where, x_{max} and x_{min} are the maximum and the minimum values for feature x, respectively [35].

3.2.2. Feature Selection

Feature selection is a typical approach to reduce the issue of unnecessary and excessive features. Feature selection strategies often lower the data dimensionality for training[36]. Figure 2 displays the feature rating.

The figure 2 displays the feature importance scores in a machine learning model. The x-axis lists the various features involved in the model, while the y-axis shows their respective importance values. The height of each bar indicates how significant a feature is in contributing to the model's predictions. Features on the left side, such as "Bwd Packet Length Max" and "Average Packet Size," have the highest importance, while features towards the right have minimal impact on the model. This plot helps in identifying which features are most influential, aiding in feature selection and model optimization.

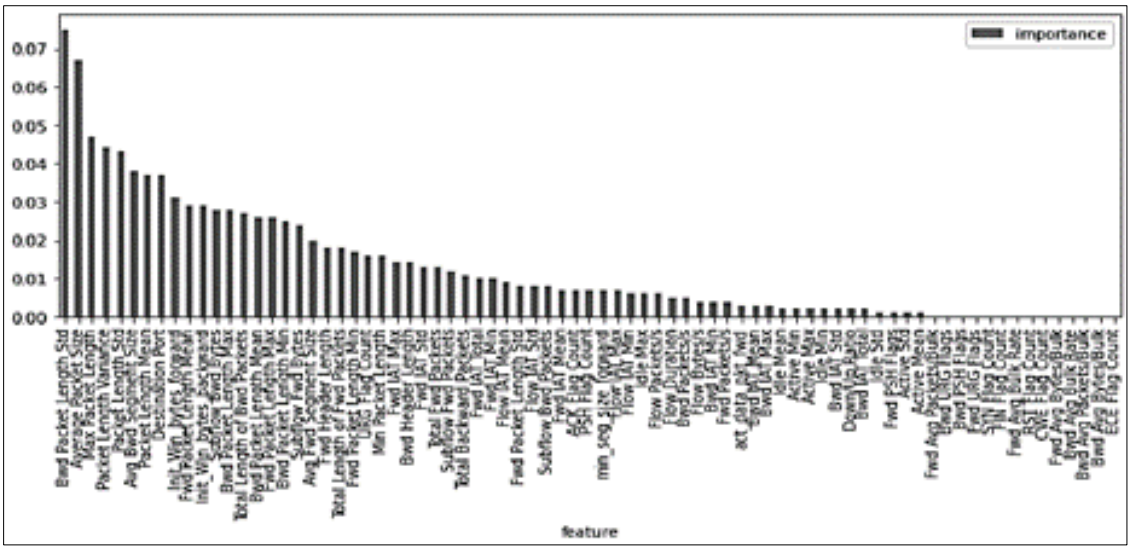

Figure 2 Importance score of each feature in CIC-IDS2017 dataset.

3.3. Data Splitting

In this study, Training and testing sets make up the CIC-IDS2017 dataset. Eighty percent of the dataset is used for training purposes, while the remaining twenty percent is kept aside for testing purposes.

3.4. Prediction ML model for Network Intrusion Network (Random Forest):

The RF algorithm is made up of many decision trees. A forecast is produced by each of these trees. Next, the algorithm takes into account the majority of these predictions to arrive at a judgement [37][38]. The capacity to handle outliers, the fact that it does not need scaling, and the fact that it can be used to both regression and classification issues are some of RF's benefits. One drawback is that this approach uses a lot of processing power since it works with a lot of decision trees, which makes model training take longer. Figure 3 provides a graphical illustration [32].

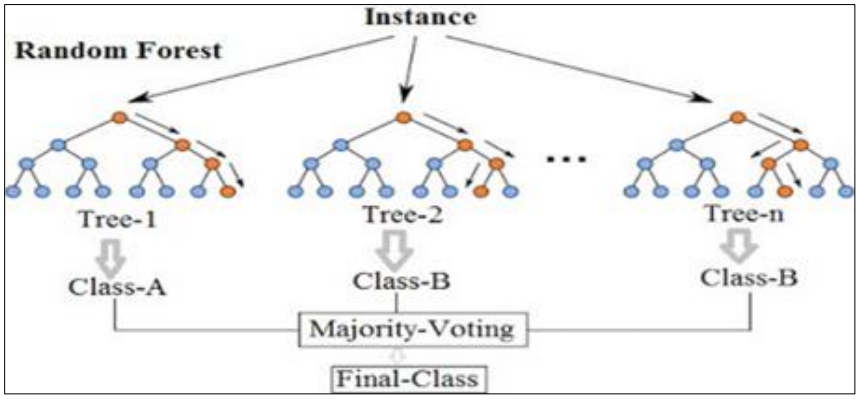

Figure 3 Graphical representations of Random Forest

Mathematically, the RF algorithm can be represented as equation (2):

$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x)$ .....(2)

Where:

- $\hat{y}$ is the final prediction.
- T is the number of trees in the forest.
- $f_t(x)$ is the prediction of the t-th tree for input x.

Each tree $f_t(x)$ is built based on a subset of the data and a subset of features, leading to a diverse collection of models. The aggregation of these predictions reduces variance and improves model accuracy.

3.4.1. Evaluation parameters

This section details the evaluation metrics utilized to assess network intrusion detection models, including accuracy, precision, recall, and F1-score. With accuracy standing for the model's general correctness, precision for the percentage of true positives detected properly, recall for the model's capacity to catch all relevant outliers, and F1-score for a balanced statistic that harmonises recall and precision.

3.4.2. Confusion Matrix.

According to, a confusion matrix is a technique that may be used to analyse how well a classifier recognises tuples from distinct classes [39][40][41]. Figure 4 shows the confusion matrix graphically. The True-Positive and True-Negative values tell us when the classifier got the data classification right, while the False-Positive and False-Negative values tell us when the classifier got the data classification incorrect.

- **TP (True Positive):** The quantity of information that has both positive predicted and positive actual value.
- **FP (False Positive):** The quantity of information including both positive and negative predictive value.
- **FN (False Negative):** For both the actual and forecast values, there is a large quantity of data.
- **TN (True Negative):** The quantity of information having both a negative predicted value and a negative actual value.

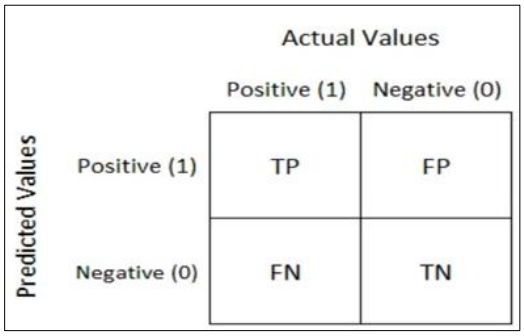

Figure 4 Confusion Matrix

The efficiency of a dataset may be evaluated using these metrics:

3.4.3. Accuracy

Accuracy, which determines the ratio of expected observations to total observations, is the most straightforward performance metric. The following mathematical expression can be found in equation (3):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad \text{……(3)}$$

3.4.4. Precision

Precision may be defined as the degree to which actual observations match theoretical predictions. The corresponding equation (4) is shown below:

$$Precision = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad \text{……(4)}$$

3.4.5. Recall

The capacity of the system to identify all current assaults is known as recall. Instead of counting real incursions, recall may be determined by looking at the total number of intrusions identified by the system. The equation (5) is shown below.

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad \text{…… (5)}$$

3.4.6. F1-Score

As a whole, recall and precision make up the F1 Score. Both positive and negative values may be assigned to this variable. F1 is sometimes more helpful than accuracy, particularly in cases when the distribution of classes is unequal, even if this does not seem to be an accuracy at first glance. Down below, you can see the matching equation:

$$F1 - score = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision+Recall}} \quad \text{.....(6)}$$

The following criteria are used to determine an algorithm's evaluation parameters, which include precision, recall, accuracy, and f1-score.

4. Experiment Results Analysis and Discussion

This section presents the performance outcomes of various network intrusion detection models applied in network security contexts. Table 2 compares the proposed RF model's effectiveness with other models using the CIS-CICIDS2017 dataset. Measures like as accuracy, precision, recall, and F1-score are the main focus of the examination, which highlights the RF model's outstanding performance on these parameters.

Table 2 Random forest model performance for network intrusion detection

Model	Accuracy	Precision	F1 – score	Recall
RF	99.9	97.78	97.41	97.08

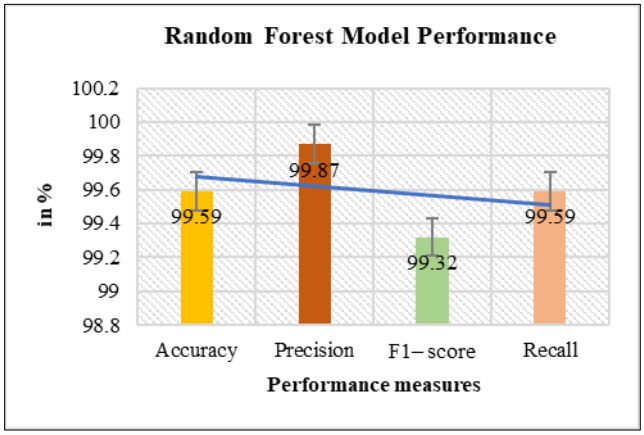

Figure 5 Random Forest (RF) Model performance

Figure 5 and Table 2 depicts the accuracy of Random Forest (RF) model for NIDS. The RF model provided very high-level performance with accuracy of 99.90 %, precision97.78 %, F1-score97.41 % and recall97.08%. The above indicators prove that the RF model is efficient and reliable in terms of detecting and categorizing intrusions while maintaining a good precision-recall balance.

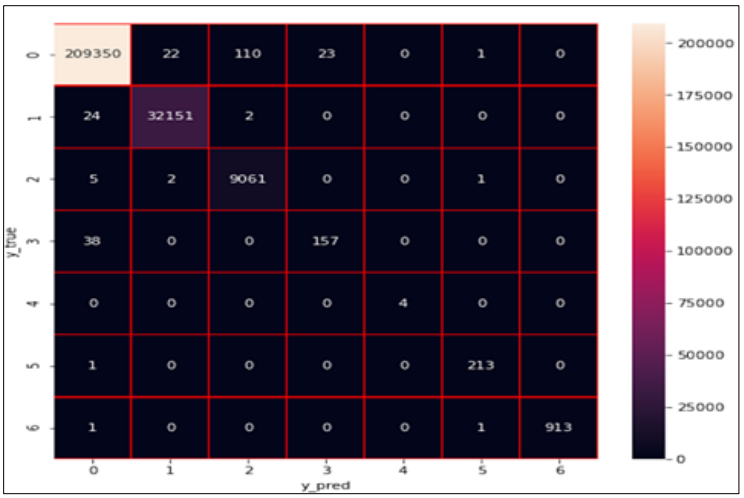

Figure 6 Random Forest (RF) Model performance as confusion matrix**.**

The figure 6 depicts a confusion matrix comparing actual `y_true` and predicted `y_pred` class labels. Diagonal values represent correctly classified instances, with class `0` showing the highest accuracy (`209,350`). Other classes, such as `1`, `2`, `5`, and `6`, also perform well, with minimal off-diagonal misclassifications. The color gradient highlights value density, and red gridlines improve clarity. Overall, the matrix indicates strong model performance with room for minor improvements in reducing misclassifications.

4.1.1. Comparative Analysis

Using metrics like accuracy, recall, precision, and F1-score, this section compares the outcomes of the suggested model with those of other ML models. The results are summarized in Table 3 and visualized in Figure 6 as a bar graph for better clarity and analysis.

Table 3 Comparision result using different ml models

Models	Accuracy	Precision	F1– score	Recall
DT [42]	99.59	99.87	99.32	99.59
Stacked LSTM [35]	99.01	96.71	98.58	97.64
AdaBoost	81.47	81.69	88.17	95.76
RF [43]	99.9	97.78	97.41	97.08

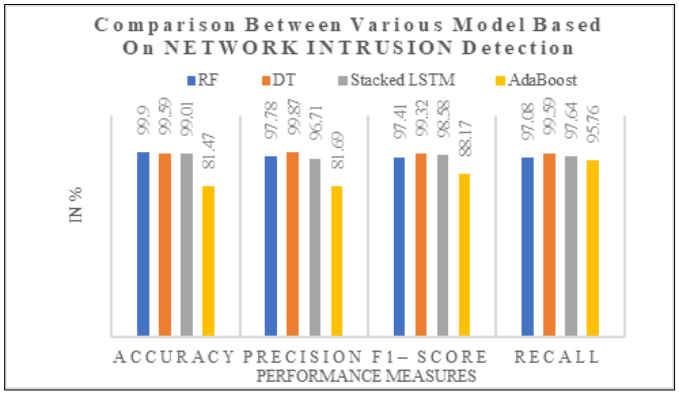

Figure 7 Comparison between different models based on network intrusion detection network security

In this comparative analysis of machine learning models for network intrusion detection using the CIS-CICIDS2017 dataset, as shown in Table 3 and illustrated in Figure 7, the Decision Tree (DT) model stands out with the highest precision of 99.87%, F1-score of 99.32%, and recall of 99.59%, demonstrating exceptional performance across all metrics. The Random Forest (RF) model achieves the highest accuracy of 99.90%, with strong precision 97.78%, recall 97.08%, and F1-score 97.41%, making it a highly competitive model. The Stacked LSTM model performs well, achieving an accuracy of 99.01%, a precision of 96.71%, recall of 97.64%, and an F1-score of 98.58%, showcasing a balanced performance. In contrast, the AdaBoost model exhibits the weakest overall performance, with an accuracy of 81.47%, precision of 81.69%, recall of 95.76%, and an F1-score of 88.17%, indicating issues with precision and accuracy despite its high recall. Overall, the DT and RF models emerge as the most reliable choices for network intrusion detection, with DT leading slightly due to its superior performance across all metrics.

5. Conclusion

Network intrusion detection is an important research direction of network security. The diversification of network intrusion mode and the increasing amount of network data make the traditional detection methods cannot meet the requirements of the current network environment. This study demonstrates the effectiveness of AI techniques, particularly the Random Forest (RF) model, in enhancing network intrusion detection. Through a comprehensive preprocessing pipeline and comparative analysis with other models, the RF model proved to be a reliable and robust approach for detecting and classifying network traffic. The results show that solutions powered by AI have the ability to improve network security.

Future Work

Future research will explore hybrid AI models, advanced feature engineering techniques, and validation on diverse datasets. Additionally, emerging methods such as transformer-based models and federated learning will be investigated to improve scalability, adaptability, and real-time detection capabilities for evolving cybersecurity challenges.

References

[1] Mani Gopalsamy, "Enhanced Cybersecurity for Network Intrusion Detection System Based Artificial Intelligence (AI) Techniques," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 12, no. 1, pp. 671–681, Dec. 2021, doi: 10.48175/IJARSCT-2269M.

[2] H. Sinha, "The Identification of Network Intrusions with Generative Artificial Intelligence Approach for Cybersecurity," *J. Web Appl. Cyber Secur.*, vol. 2, no. 2, pp. 20–29, Oct. 2024, doi: 10.48001/jowacs.2024.2220-29.

[3] M. Stampar and K. Fertalj, "Artificial intelligence in network intrusion detection," in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings*, 2015. doi: 10.1109/MIPRO.2015.7160479.

[4] H. Sinha, "Analysis of anomaly and novelty detection in time series data using machine learning techniques," *Multidiscip. Sci. J.*, vol. 7, no. 06, 2024, doi: https://doi.org/10.31893/multiscience.2025299.

[5] R. Arora, A. Kumar, A. Soni, and A. Tiwari, "AI-Driven Self-Healing Cloud Systems : Enhancing Reliability and Reducing Downtime through Event-Driven Automation," 2024, doi: 10.20944/preprints202408.1860.v1.

[6] B. Boddu, "Challenges and Best Practices for Database Administration in Data Science and Machine Learning," *https://www.ijirmps.org/research-paper.php?id=231461*, vol. 9, no. 2, p. 7, 2021.

[7] S. Bauskar, "AN PREDICTIVE ANALYTICS OR DATA QUALITY ASSESSMENT THROUGH ARTIFICIAL INTELLIGENCE TECHNIQUES," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 06, no. 09, pp. 3330–3337, 2024, doi: https://www.doi.org/10.56726/IRJMETS61568.

[8] H. S. Chandu, "Enhancing Manufacturing Efficiency : Predictive Maintenance Models Utilizing IoT Sensor Data," *IJSART*, vol. 10, no. 9, 2024.

[9] L. V Mohammed Kaif, Prajwal P, "A Study on Network Intrusion Detection System," *Int. J. Multidiscip. Res.*, vol. 6, no. 3, pp. 1–9, Jun. 2024, doi: 10.36948/ijfmr.2024.v06i03.20214.

[10] N. Abid, "A Climbing Artificial Intelligence for Threat Identification in Critical Infrastructure Cyber Security," *Int. J. Res. Anal. Rev.*, vol. 9, no. 4, 2022.

[11] Noman Abid, "A Review of Security and Privacy Challenges in Augmented Reality and Virtual Reality Systems with Current Solutions and Future Directions," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 3, no. 1, pp. 511–523,

Feb. 2023, doi: 10.48175/IJARSCT-8329A.

[12] N. G. Abhinav Parashar, "Asset Master Data Management: Ensuring Accuracy and Consistency in Industrial Operations," *IJNRD - Int. J. Nov. Res. Dev.*, vol. 9, no. 9, pp. 861-a867, 2024.

[13] B. Boddu, "Ensuring Data Integrity and Privacy: A Guide for Database Administrators," *https://www.ijfmr.com/research-paper.php?id=10880*, vol. 4, no. 6, p. 6, 2022.

[14] Sahil Arora and Pranav Khare, "AI/ML-Enabled Optimization of Edge Infrastructure: Enhancing Performance and Security," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 6, no. 1, pp. 1046–1053, 2024, doi: 10.48175/568.

[15] A. Goyal, "Enhancing Engineering Project Efficiency through Cross-Functional Collaboration and IoT Integration," *Int. J. Res. Anal. Rev.*, vol. 8, no. 4, pp. 396–402, 2021.

[16] J. Thomas, J. Vummadi, and R. Shah, "Machine Learning Integrated Supplier Management Device," 2024 [Online]. Available: https://www.researchgate.net/publication/384291069_Machine_Learning_Integrated_Supplier_Management_Device

[17] P. Bosco, "Intrusion Detection and Prevention Systems Cheat Sheet: Choosing the Best Solution, Common Misconfigurations, Evasion Techniques, and Recommendations.," *SANDS*, 2016.

[18] V. S. Thokala, "A Comparative Study of Data Integrity and Redundancy in Distributed Databases for Web Applications," *Int. J. Res. Anal. Rev.*, vol. 8, no. 4, pp. 383–389, 2021.

[19] M. Gopalsamy, "Predictive Cyber Attack Detection in Cloud Environments with Machine Learning from the CICIDS 2018 Dataset," *IJSART*, vol. 10, no. 10, 2024.

[20] S. G. Jubin Thomas, Piyush Patidar, Kirti Vinod Vedi, "Predictive Big Data Analytics For Supply Chain Through Demand Forecastin," *Int. J. Creat. Res. Thoughts*, vol. 10, no. 06, pp. h868–h873, 2022.

[21] P. Khare, "Transforming KYC with AI: A Comprehensive Review of Artificial Intelligence-Based Identity Verification," *J. Emerg. Technol. Innov. Res.*, vol. 10, no. 12, pp. 525–530, 2023.

[22] V. V. Kumar, A. Sahoo, and F. W. Liou, "Cyber-enabled Product Lifecycle Management: A Multi-agent Framework," *Procedia Manuf.*, vol. 39, pp. 123–131, 2019, doi: 10.1016/j.promfg.2020.01.247.

[23] P. P. Angelov, E. A. Soares, R. Jiang, N. I. Arnold, and P. M. Atkinson, "Explainable artificial intelligence: an analytical review," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 2021, doi: 10.1002/widm.1424.

[24] A. K. Sah and V. K, "Anomaly-Based Intrusion Detection in Network Traffic using Machine Learning: A Comparative Study of Decision Trees and Random Forests," in *2024 2nd International Conference on Networking and Communications (ICNWC)*, 2024, pp. 1–7. doi: 10.1109/ICNWC60771.2024.10537451.

[25] X. Hu, X. Meng, S. Liu, and L. Liang, "An Improved Algorithm for Network Intrusion Detection Based on Deep Residual Networks," *IEEE Access*, vol. 12, pp. 66432–66441, 2024, doi: 10.1109/ACCESS.2024.3398007.

[26] S. Arshad, W. Ashraf, S. Ashraf, I. Hassan, and F. S. Masoodi, "Comparative Study of Machine Learning Techniques for Intrusion Detection on CICIDS-2017 Dataset," in *Proceedings of the 17th INDIACom; 2023 10th International Conference on Computing for Sustainable Global Development, INDIACom 2023*, 2023.

[27] J. Akoto and T. Salman, "Machine Learning vs Deep Learning for Anomaly Detection and Categorization in Multi-cloud Environments," in *Proceedings - 2022 IEEE Cloud Summit, Cloud Summit 2022*, 2022. doi: 10.1109/CloudSummit54781.2022.00013.

[28] K. Atefi, H. Hashim, and M. Kassim, "Anomaly analysis for the classification purpose of intrusion detection system with K-nearest neighbors and deep neural network," in *Proceeding - 2019 IEEE 7th Conference on Systems, Process and Control, ICSPC 2019*, 2019. doi: 10.1109/ICSPC47137.2019.9068081.

[29] H. He, X. Sun, H. He, G. Zhao, L. He, and J. Ren, "A Novel Multimodal-Sequential Approach Based on Multi-View Features for Network Intrusion Detection," *IEEE Access*, 2019, doi: 10.1109/ACCESS.2019.2959131.

[30] M. Gopalsamy, "Evaluating the Effectiveness of Machine Learning (ML) Models in Detecting Malware Threats for Cybersecurity," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, 2023, doi: : https://doi.org/10.14741/ijcet/v.13.6.4.

[31] R. Goyal, "AN EFFECTIVE MACHINE LEARNING BASED REGRESSION TECHNIQUES FOR PREDICTION OF HEALTH INSURANCE COST," *Int. J. Core Eng. Manag.*, vol. 7, no. 11, 2024.

[32] M. Al Lail, A. Garcia, and S. Olivo, "Machine Learning for Network Intrusion Detection—A Comparative Study,"

Futur. Internet, 2023, doi: 10.3390/fi15070243.

[33] M. Mujahid *et al.*, "Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering," *J. Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-024-00943-4.

[34] B. Boddu, "Essential Cybersecurity Measures for Databases to Mitigate Cyber Attacks," *https://www.ijirmps.org/research-paper.php?id=231460*, vol. 11, no. 6, p. 8, 2023.

[35] J. Figueiredo, C. Serrão, and A. M. de Almeida, "Deep Learning Model Transposition for Network Intrusion Detection Systems," *Electron.*, 2023, doi: 10.3390/electronics12020293.

[36] B. Boddu, "DevOps for Database Administration: Best Practices and Case Studies," *https://jsaer.com/download/vol-7-iss-3-2020/JSAER2020-7-3-337-342.pdf*, vol. 7, no. 3, p. 5, 2020.

[37] S. Pandey and S. Pandey, "Integrating ' Approval to Post ' into HCM Systems : Enhancing Efficiency and Transparency through Workday Solutions," *North Am. J. Eng. Res.*, vol. 2, no. 2, 2021.

[38] R. Tandon, "The Machine Learning Based Regression Models Analysis For House Price Prediction," vol. 11, no. 3, 2024.

[39] D. Xhemali, C. J. Hinde, and R. G. Stone, "Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages," *Int. J. Comput. Sci.*, 2009.

[40] N. M. vikas kumar, "Warranty Failure Analysis in Service Supply Chain A Multi-agent Framework Warranty Failure Analysis in Service Supply Chain," *ResearchGate*, no. May, pp. 1–7, 2011.

[41] S. Pandey, "TRANSFORMING PERFORMANCE MANAGEMENT THROUGH AI: ADVANCED FEEDBACK MECHANISMS, PREDICTIVE ANALYTICS, AND BIAS MITIGATION IN THE AGE OF WORKFORCE OPTIMIZATION," *Int. J. Bus. Quant. Econ. Appl. Manag. Res.*, vol. 6, no. 7, pp. 13–19, 2020.

[42] T. H. Chua and I. Salam, "Evaluation of Machine Learning Algorithms in Network-Based Intrusion Detection Using Progressive Dataset," *Symmetry (Basel).*, 2023, doi: 10.3390/sym15061251.

[43] A. Yulianto, P. Sukarno, and N. A. Suwastika, "Improving AdaBoost-based Intrusion Detection System (IDS) Performance on CIC IDS 2017 Dataset," in *Journal of Physics: Conference Series*, 2019. doi: 10.1088/1742-6596/1192/1/012018.