XGBoost Machine Learning Model Outperformed Competitors in Network Intrusion Detection

Peter O. Abaji
Lagos State University
Faculty of Computing and Information Technology
Department of Computer Science

Adedoyin T. Odumabo
Lagos State University of Science and Technology
College of Basic Science
Department of Computer Science

Benjamin S. Aribisala
Lagos State University
Faculty of Computing and Information Technology
Department of Computer Science

ABSTRACT
Intrusion Detection Systems (IDS) are essential components of modern cybersecurity infrastructures because they help identify unauthorized access, malicious activities, and cyber threats within network environments. However, traditional signature-based intrusion detection approaches often struggle to detect sophisticated and emerging attacks due to their reliance on predefined attack patterns. This study presents a machine learning-based intrusion detection framework using the UNSW-NB15 dataset to improve network attack detection accuracy and reliability. Five machine learning classifiers, namely Naive Bayes, Bagging, Random Forest, Multi-Layer Perceptron, and XGBoost, were implemented and comparatively evaluated for binary classification of network traffic into normal and malicious categories. Data preprocessing techniques such as feature scaling, label encoding, and train-test splitting were applied before model training and evaluation. The performance of the classifiers was assessed using accuracy, precision, recall, and F1-score metrics with weighted averaging to address class imbalance challenges within the dataset. Experimental results showed that ensemble learning approaches significantly outperformed individual classifiers. Among the evaluated models, XGBoost achieved the best overall performance with an accuracy of 90.07% and an F1-score of 89.77%, demonstrating strong capability in balancing precision and recall for intrusion detection tasks. The findings of this study indicate that XGBoost provide robust and reliable solutions for modern intrusion detection systems and can effectively improve cybersecurity defenses in contemporary network environments.

General Terms
Network Intrusion Detection

Keywords
Intrusion Detection Systems (IDS), Machine Learning, Ensemble Learning, Network Security, UNSW_NB15.

1. INTRODUCTION
This study is motivated by the growing need for intelligent and adaptive intrusion detection systems capable of overcoming the limitations of traditional cybersecurity approaches in modern network environments. The increasing complexity of cyber threats and network traffic patterns necessitates the development of robust machine learning techniques for accurate attack detection [1, 2].

The significance of this research lies in its contribution to improving intrusion detection performance through the comparative evaluation of conventional and ensemble machine learning classifiers using the UNSW-NB15 dataset. The study provides insights into the effectiveness and reliability of different machine learning approaches for detecting malicious network activities [3, 4].

The study contributes to existing literature by demonstrating the effectiveness of ensemble learning techniques, particularly XGBoost, in improving intrusion detection accuracy and reducing classification errors in complex network environments [5, 6]. Several researchers have applied supervised and ensemble learning approaches to intrusion detection with promising outcomes [2, 3]. Moustafa and Slay (2015) introduced the UNSW-NB15 dataset to address limitations associated with older benchmark datasets such as KDD99 and NSL-KDD by incorporating modern attack scenarios and realistic network traffic distributions [7]. . Recent studies have also demonstrated the effectiveness of ensemble-based models such as Random Forest, Bagging, and XGBoost for improving intrusion detection accuracy and reducing false positives [5, 4]. This study aims to comparatively evaluate multiple machine learning classifiers in order to identify a robust and reliable intrusion detection model suitable for modern network environments.

The remainder of this paper is organized as follows: Section 2 presents the literature review, Section 3 describes the methodology and experimental setup, Section 4 discusses the experimental results, and Section 5 Concludes the study.

2. LITERATURE REVIEW
Machine learning techniques have become increasingly important in network intrusion detection research because of their ability to automatically learn complex attack patterns and improve cyber threat detection accuracy. Recent advances in artificial intelligence and data-driven cybersecurity have significantly enhanced the capability of Intrusion Detection Systems (IDS) to detect sophisticated and evolving cyberattacks [1, 2]. Unlike traditional signature-based approaches, machine learning-based IDS frameworks can generalize from historical traffic patterns and identify previously unseen attacks.

The development of benchmark intrusion detection datasets has also contributed significantly to IDS research progress. Moustafa and Slay (2015) introduced the UNSW-NB15 dataset to address the shortcomings of older datasets such as KDD99 and NSL-KDD by incorporating realistic network traffic and modern attack scenarios [7]. Recent studies have widely adopted the UNSW-NB15 dataset because it provides a more representative evaluation environment for contemporary intrusion detection systems [4].

Naive Bayes classifiers remain one of the most widely used conventional machine learning techniques in IDS research due to their simplicity, fast training speed, and low computational

requirements. However, several studies reported that the independence assumption of Naive Bayes limits its effectiveness when handling highly correlated network traffic features and complex attack patterns [1, 6].

Ensemble learning techniques have demonstrated superior intrusion detection performance by combining multiple weak learners into stronger predictive models. Bagging and Random Forest algorithms are particularly effective in reducing overfitting, improving classification stability, and enhancing generalization performance in cybersecurity applications [8, 9]. Recent comparative studies further confirmed that ensemble-based IDS models achieve higher detection accuracy and lower false alarm rates than conventional classifiers [5, 4].

Gradient boosting techniques, particularly XGBoost, have gained substantial attention in modern IDS research because of their high predictive capability and computational efficiency. Chen and Guestrin (2016) demonstrated that XGBoost optimizes gradient boosting through parallel tree boosting and regularization mechanisms, thereby improving model accuracy and scalability [10]. Recent cybersecurity studies have also shown that XGBoost consistently outperforms several traditional machine learning approaches in intrusion detection tasks involving high-dimensional and imbalanced datasets

Artificial Neural Networks (ANNs), especially Multi-Layer Perceptron (MLP) models, have also been extensively applied in intrusion detection systems because of their ability to capture nonlinear relationships within network traffic data. Shone et al. (2018) demonstrated that deep learning and neural-network-based intrusion detection approaches significantly improve attack detection performance [3]. However, neural network models often require extensive parameter tuning, high computational resources, and large training datasets [11].

Despite the significant progress achieved in machine learning-based intrusion detection, several research challenges remain unresolved. Many previous studies focused on a limited number of classifiers, neglected class imbalance problems, or failed to conduct comprehensive comparative evaluations using modern benchmark datasets such as UNSW-NB15 [12]. Furthermore, relatively few studies have simultaneously compared conventional machine learning classifiers with advanced ensemble learning techniques under identical experimental conditions using weighted evaluation metrics.

Therefore, this study addresses these research gaps by conducting a comprehensive comparative evaluation of five machine learning classifiers, namely Naive Bayes, Bagging, Random Forest, Multi-Layer Perceptron Neural Network, and XGBoost, using the UNSW-NB15 dataset [13, 14]. The study further incorporates weighted evaluation metrics, confusion matrix analysis, ROC curve analysis, and feature importance analysis to provide a more reliable and realistic assessment of classifier effectiveness for intelligent intrusion detection systems.

3. METHODOLOGY

This study adopted a machine learning-based experimental methodology for intelligent network intrusion detection using the UNSW-NB15 dataset. The proposed framework consists of data acquisition, preprocessing, feature transformation, classifier implementation, model training, performance evaluation, and result interpretation stages. The dataset was preprocessed through label encoding, feature normalization, and train-test partitioning before being supplied to the selected machine learning classifiers. Five classifiers, namely Naive Bayes, Bagging, Random Forest, Multi-Layer Perceptron Neural Network, and XGBoost, were implemented and comparatively evaluated using weighted performance metrics. Additional evaluation and interpretation techniques including confusion matrix analysis, ROC curve analysis, and feature importance analysis were also conducted to assess the effectiveness of the proposed intrusion detection framework

3.1 Dataset Description

The UNSW-NB15 dataset used in this study was developed by the Australian Centre for Cyber Security (ACCS) using the IXIA PerfectStorm tool. The dataset contains modern synthetic network traffic that combines normal activities with contemporary attack behaviors. It includes nine attack categories, namely:

- Fuzzers
- Analysis
- Backdoors
- DoS
- Exploits
- Generic
- Reconnaissance
- Shellcode
- Worms

The dataset consists of approximately 2.54 million records with 49 features extracted using the Argus and Bro-IDS tools. These features include flow characteristics, basic packet information, content features, and time-based statistical attributes.

For this study, the dataset was obtained from the official UNSW-NB15 repository on January 15, 2026: https://research.unsw.edu.au/projects/unsw-nb15-dataset

3.2 Data Preprocessing

The following preprocessing procedures were applied before model training:

1. Missing value checking and removal of duplicate records.
2. Label encoding of categorical attributes into numerical representations using LabelEncoder.
3. Feature normalization using StandardScaler to improve classifier convergence.
4. Dataset partitioning into training and testing subsets using a 70:30 ratio.
5. Weighted evaluation metrics were adopted to address class imbalance.

3.3 Experimental Setup

The experiments were conducted using Python programming language with libraries including pandas, NumPy, scikit-learn, imbalanced-learn, matplotlib, and xgboost.

The experiment was executed on a computer system with the following specifications:

- Processor: 1.8 GHz Dual-Core Intel Core i5
- Ram: 8GB SSD:256
- Operating System: macOS Sonoma
- Python Version: Python 3.x

To improve reliability and reduce experimental bias, each classifier was executed five times and the average performance metrics were recorded.

The following classifiers were implemented:

- Gaussian Naive Bayes
- Bagging Classifier
- Random Forest Classifier
- Multi-Layer Perceptron Classifier
- XGBoost Classifier

Performance evaluation and interpretation techniques was conducted using:

- Accuracy
- Precision
- Recall
- F1-Score
- ROC Curve
- Confusion Matrix

These metrics were computed using weighted averaging to ensure fair evaluation under imbalanced class distributions.

4. RESULTS AND DISCUSSIONS

4.1 Experimental Results Presentation

4.1.1 Comparative Performance of Machine Learning Classifiers

The results presented in Table 1 show the comparative performance of the evaluated machine learning classifiers based on weighted accuracy, precision, recall, and F1-score metrics. The evaluation demonstrates that ensemble learning approaches significantly outperform conventional classifiers in intrusion detection tasks.

Table 2 further highlights the superiority of ensemble learning models, with XGBoost achieving the highest classification accuracy among all implemented techniques.

The chart shows that XGBoost achieved the highest overall performance among the evaluated classifiers, with an accuracy of 90.07% and an F1-score of 89.77% (See Figure 2). Bagging and Random Forest also demonstrated strong performance, indicating the effectiveness of ensemble learning techniques for network intrusion detection tasks. The MLP Neural Network produced competitive results with an accuracy of 88.69%, while Naive Bayes recorded the lowest performance due to its limitation in handling complex relationships among network traffic features. Overall, the results confirm that ensemble-based machine learning models provide more reliable and accurate intrusion detection performance compared to conventional classification approaches.

4.1.2 Confusion Matrix for XGBoost

Figure 3. indicate that the model achieved strong classification performance across most network traffic classes, as shown by the high number of correctly classified instances along the diagonal of the matrix. Classes 5 and 6 recorded the highest correct predictions with 5,574 and 11,097 instances, respectively, demonstrating the effectiveness of the classifier in detecting major attack categories. Some misclassifications were observed between related classes, particularly among Classes 2, 3, and 4, due to similarities in network traffic characteristics. However, the number of incorrect predictions remained relatively low compared to the correctly classified instances. Overall, the confusion matrix confirms that the XGBoost classifier provides robust and reliable intrusion detection performance with high classification accuracy and minimal false predictions, making it suitable for intelligent network intrusion detection systems.

4.1.3 ROC Curve for XGBoost

Figure 4 presents the Receiver Operating Characteristic (ROC) curves and Area Under Curve (AUC) scores for the XGBoost classifier. The ROC analysis demonstrates excellent classification capability across all evaluated classes. Most classes achieved AUC values ranging from 0.95 to 1.00, indicating outstanding discrimination performance between attack and non-attack classes. Classes 5, 6, 8, and 9 achieved perfect AUC scores of 1.00, signifying near-perfect classification capability for those categories. The high AUC values confirm that the XGBoost classifier possesses strong predictive power and maintains a favorable balance between true positive rate and false positive rate. These findings further validate the suitability of ensemble boosting techniques for intelligent network intrusion detection systems.

4.1.4 Top 10 Important Features in XGBoost

Figure 5 illustrates the top 10 most important features identified by the XGBoost classifier. The feature importance analysis indicates that the swin feature contributed most significantly to the classification process, followed by service_dns, dttl, and sttl. These features play critical roles in distinguishing between normal and malicious network traffic patterns. The dominance of protocol- and service-related features suggests that traffic behavior characteristics are highly informative for intrusion detection tasks. The results demonstrate that XGBoost effectively identifies relevant network attributes and prioritizes features that provide the greatest discriminatory power for attack detection. The feature importance analysis further enhances the interpretability of the proposed intrusion detection framework by providing insight into the network traffic attributes most relevant to cyberattack classification.

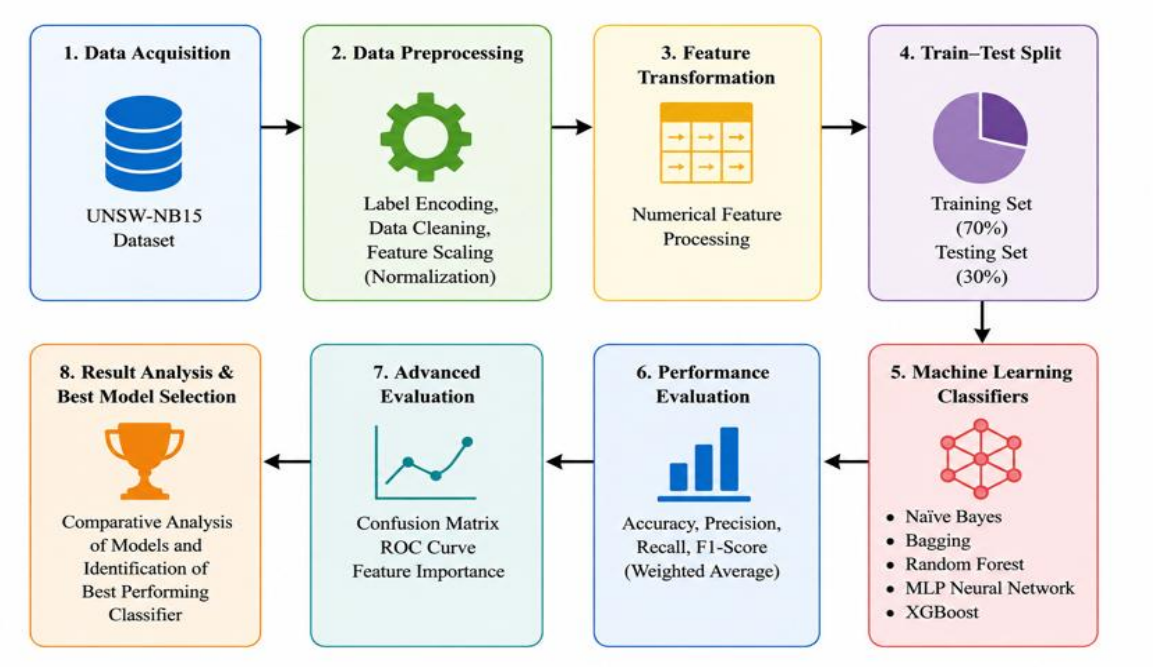

Figure 1: Proposed Intrusion Detection Framework

Table 1: Performance Evaluation of Machine Learning Classifiers

Classifier	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Naive Bayes	57.27	69.45	57.27	57.63
Bagging	89.40	89.45	89.40	89.40
Random Forest	89.22	88.96	89.22	89.03
MLP Neural Network	88.69	88.26	88.69	88.19
XGBoost	90.07	89.82	90.07	89.77

Table 2: Comparative Ranking of Classifiers Based on Accuracy

Rank	Classifier	Accuracy (%)
1	XGBoost	90.07
2	Bagging	89.40
3	Random Forest	89.22
4	MLP Neural Network	88.69
5	Naive Bayes	57.27

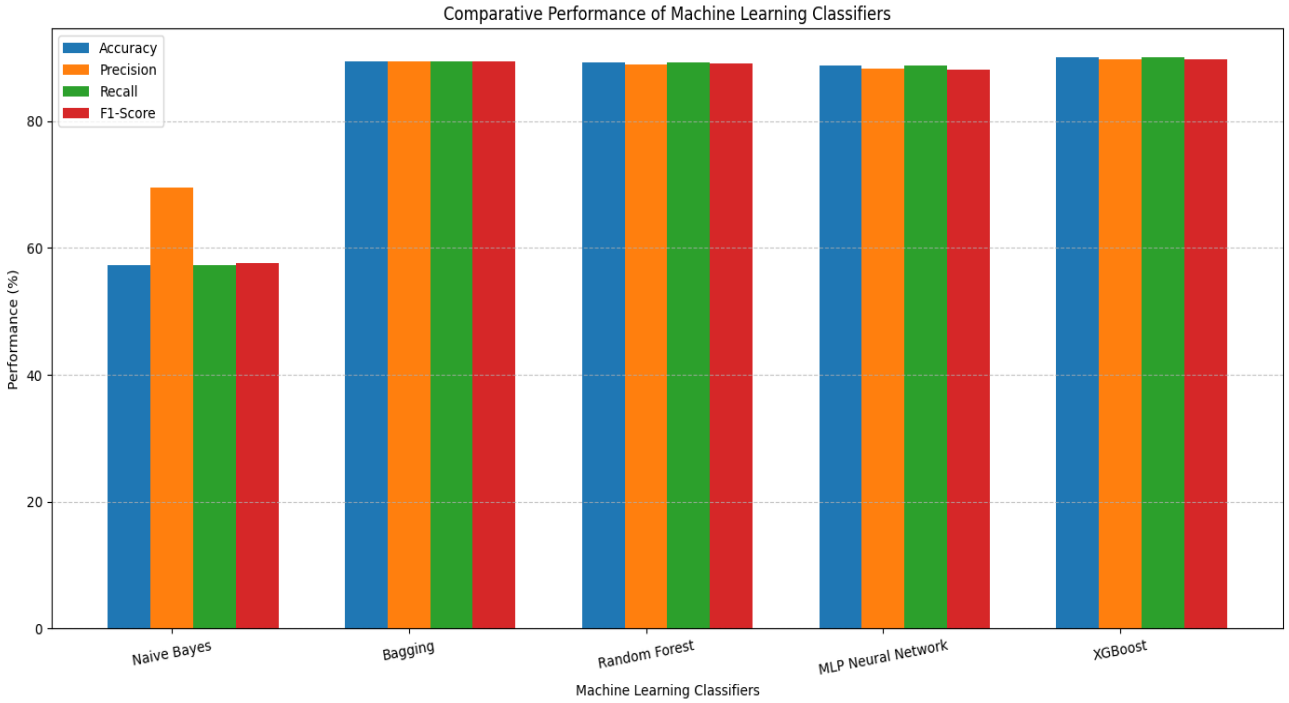

Figure 2: Comparative Performance of Machine Learning Classifiers

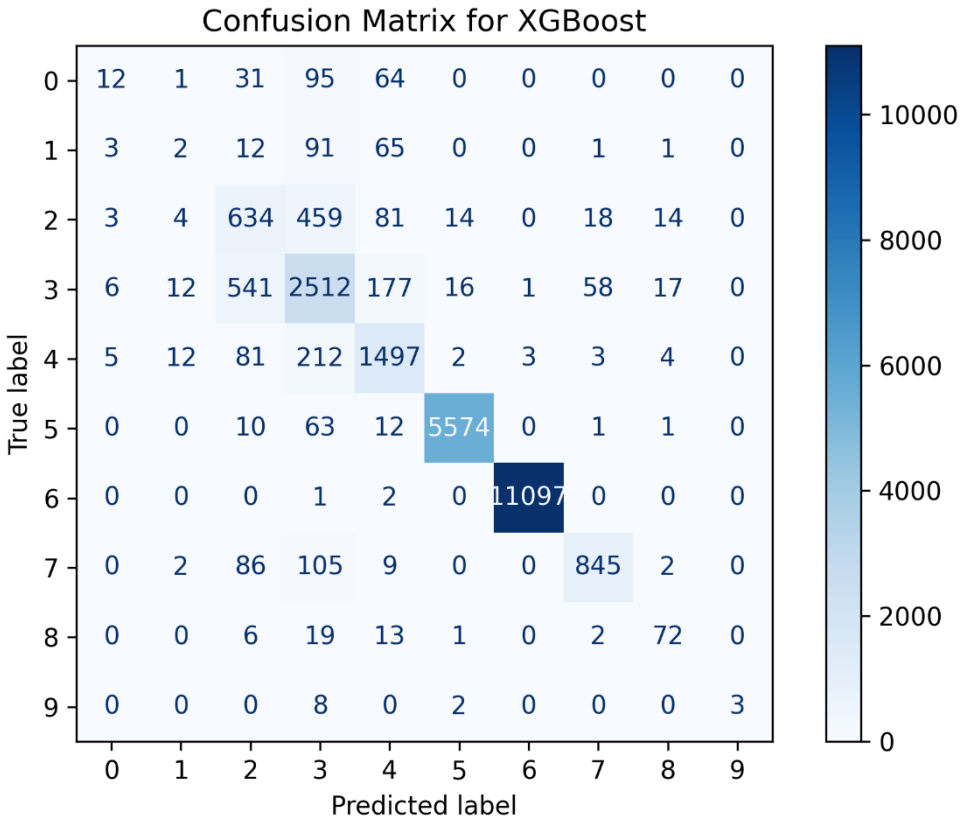

Figure 3: Confusion Matrix for XGBoost

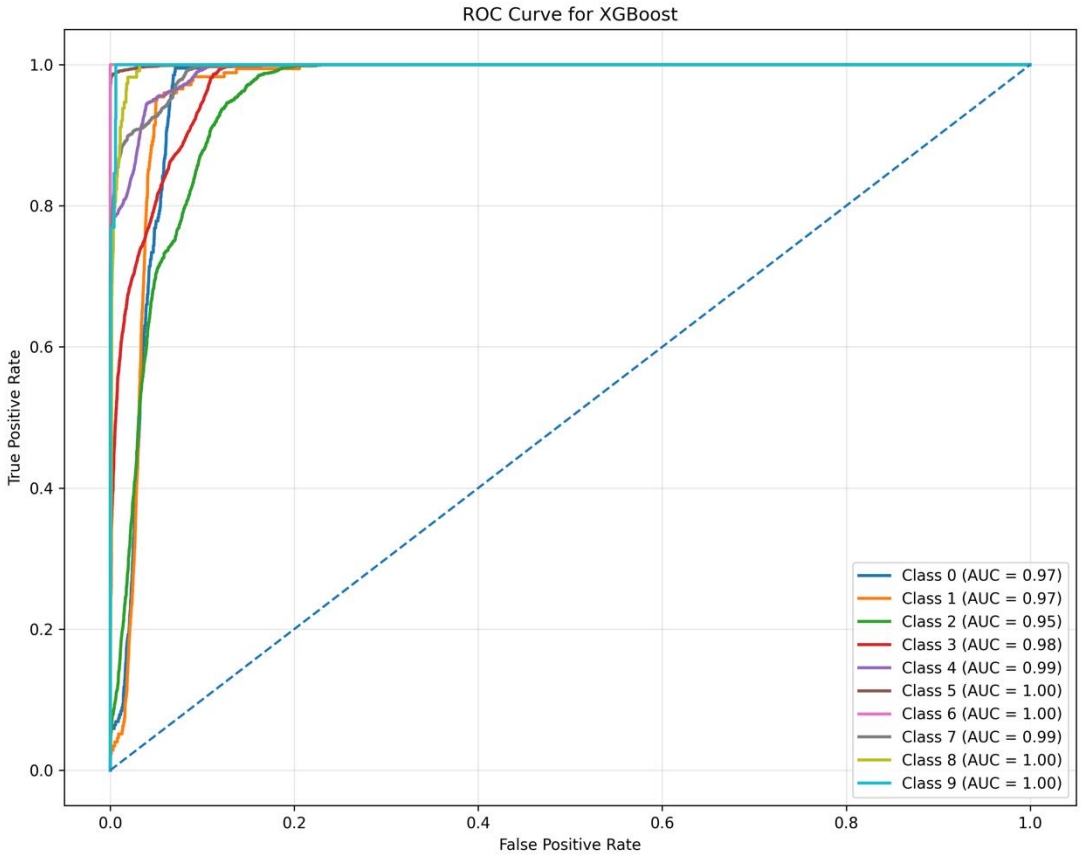

Figure 4: ROC Curve for XGBoost.

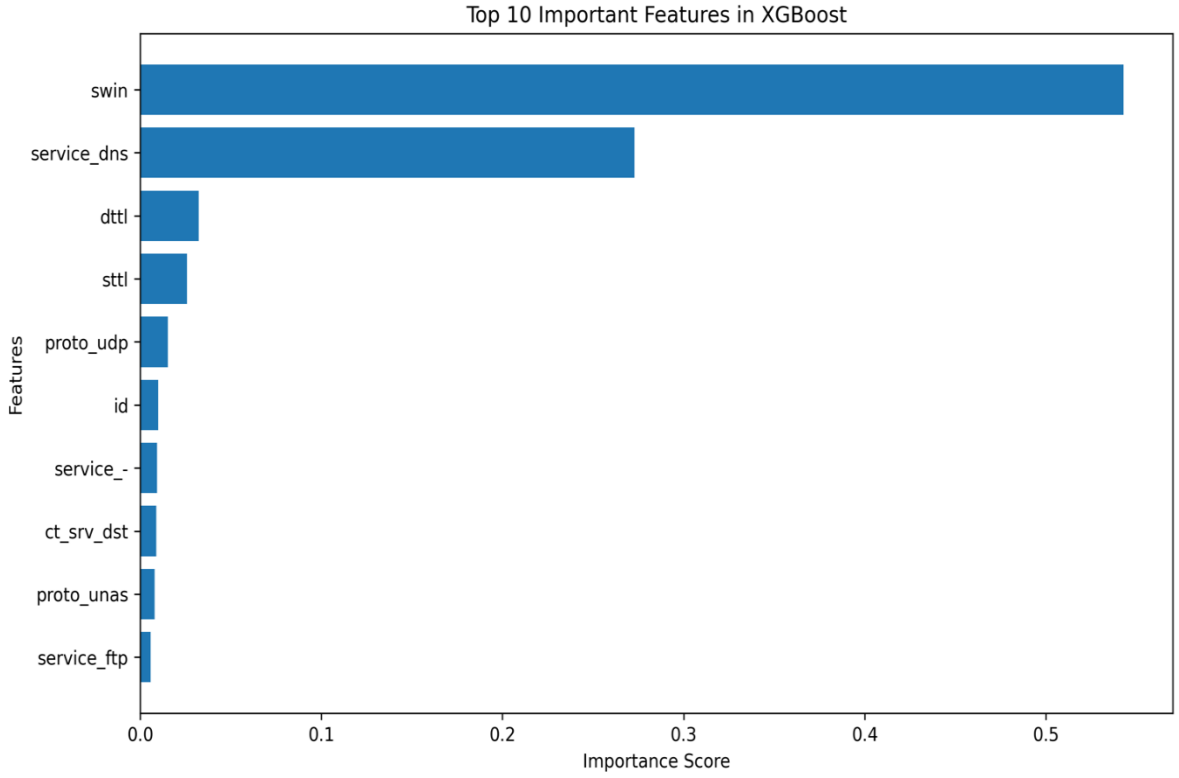

Figure 5: Top 10 Important Features in XGBoost

4.2 Experimental Reliability

To ensure consistency and reliability, the experiments were repeated multiple times under the same conditions, and the average results were reported. Weighted evaluation metrics were used because the UNSW-NB15 dataset contains imbalanced attack distributions that may bias conventional accuracy measurements.

The experimental results demonstrate notable performance differences among the evaluated classifiers. Ensemble-based models consistently outperformed individual learning algorithms. XGBoost achieved the highest overall performance across all evaluation metrics, followed closely by Bagging and Random Forest classifiers. These results highlight the effectiveness of ensemble learning in improving detection accuracy and robustness.

In contrast, the Naive Bayes classifier exhibited comparatively lower performance due to its simplifying assumptions, which limit its ability to model complex feature interactions. Balanced Bagging and Easy Ensemble approaches showed improved detection of minority attack classes, underscoring the importance of addressing class imbalance in intrusion detection tasks.

5. CONCLUSION

This study proposed a machine learning-based framework for network intrusion detection using the UNSW-NB15 dataset. Five machine learning classifiers were comparatively evaluated using weighted performance metrics, including accuracy, precision, recall, and F1-score. The experimental findings revealed that ensemble learning methods consistently outperformed individual classifiers. Among all evaluated models, XGBoost achieved the best overall performance with an accuracy of 90.07% and an F1-score of 89.77%, demonstrating its effectiveness and robustness for intrusion detection applications. Random Forest also produced strong results, particularly in handling imbalanced attack distributions. The study contributes to cybersecurity research by providing a comparative evaluation of conventional, ensemble, and imbalance-aware machine learning techniques under the same experimental conditions. The findings further emphasize the importance of ensemble learning approaches like XGBoost in improving intrusion detection performance within modern network environments.

The future work may focus on integrating deep learning and hybrid ensemble techniques to further improve intrusion detection accuracy and real-time attack detection capability. Additional studies may also investigate feature selection optimization, real-world deployment scenarios, and evaluation using larger and more diverse cybersecurity datasets. Furthermore, incorporating explainable artificial intelligence (XAI) techniques could improve the interpretability and transparency of intrusion detection models in practical cybersecurity applications.

6. REFERENCES

[1] Ahmad, Z., Khan, A. S., Wai Shiang, C., Abdullah, J., and Ahmad, F. 2021. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. Transactions on Emerging Telecommunications Technologies, vol. 32, no. 1.

[2] Ferrag, M. A., Maglaras, L., Moschoyiannis, S., and Janicke, H. 2020. Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications, vol. 50.

[3] Shone, N., Ngoc, T. N., Phai, V. D., and Shi, Q. 2018. A deep learning approach to network intrusion detection. IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41–50.

[4] Tama B. A., and Lim, S. 2020. Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation. Computer Science Review, vol. 39.

[5] Alazzam, H., Sharieh, A., and Sabri, K. E. 2022. A feature selection algorithm for intrusion detection system based on Pigeon Inspired Optimizer. Expert Systems with Applications, vol. 148.

[6] Vinayakumar, R., Soman, K. P., Poornachandran, P., and Akarsh, S. 2019. Evaluating deep learning approaches to characterize and classify network traffic. Journal of Intelligent & Fuzzy Systems, vol. 36, no. 5, pp. 4775–4785.

[7] Moustafa N., and Slay, J. 2015. UNSW-NB15: A comprehensive data set for network intrusion detection systems. In Proc. MilCIS, pp. 1–6.

[8] Breiman, L. 1996. Bagging predictors. Machine Learning, vol. 24, no. 2, pp. 123–140.

[9] Breiman, L. 2001. Random forests. Machine Learning, vol. 45, no. 1, pp. 5–32.

[10] Chen T., and Guestrin, C. 2016. XGBoost: A scalable tree boosting system. In Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

[11] Haykin, S. 1999. Neural Networks: A Comprehensive Foundation, 2nd ed. Prentice Hall.

[12] Lemaître, G., Nogueira, F., and Aridas, C. K. 2017. Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning. Journal of Machine Learning Research, vol. 18, no. 17, pp. 1–5.

[13] Liu, X. Y., Wu, J., and Zhou, Z. H. 2009. Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 39, no. 2, pp. 539–550.

[14] Rish, I. 2001. An empirical study of the Naive Bayes classifier. IJCAI Workshop on Empirical Methods in Artificial Intelligence, pp. 41–46.