Societal Transformation: AI and Big Data Journal
Vol: 4(1):84-99,,2026
DOI:10.20547/aibd.264105

Explainable Intrusion Detection on UNSW-NB15 with ExtraTrees Feature Selection and Ensemble Learning

Waqas Aziz[1] ,Imran Ali[2],Abdul Ghafoor[3],Farooq Alam[4]

Abstract

Intrusion Detection Systems important for protecting the modern networks against increasingly advanced cyberattacks; however, high-dimensional network traffic, class complexity, and limited model interpretability continue to affect the effectiveness of machine learning based IDSs. This research developed an explainable intrusion detection framework for UNSW-NB15 dataset by integrating ExtraTrees-based feature-selection with multiple machine-learning and ensemble classifiers. The empirical pipeline included data cleaning, categorical encoding, feature normalization, ExtraTrees-based feature ranking, classification, and performance evaluation. Five classifiers Naïve Bayes, Decision Tree, k-Nearest Neighbors, Random Forest, and LightGBM were experimentally evaluated using accuracy, precision, recall, F1-score, confusion matrices, and ROC analysis. ExtraTrees reduced the input feature space by retaining the most informative network traffic attributes and simultaneously provided feature-level information for interpreting the classification process. Among the evaluated classifiers, Random Forest achieved the highest reported overall performance, with an accuracy of 89.12% and an F1-score of 89.10%, while LightGBM also demonstrated competitive performance. ROC analysis produced an AUC of approximately 0.97, indicating strong discrimination between normal and malicious network traffic. The results further showed that ensemble-based classifiers provided better overall detection performance than the evaluated conventional classifiers. The principal contribution of this study is the integration of ExtraTrees-based feature selection, ensemble machine learning, and feature-importance-based interpretability within a unified intrusion detection framework, thereby addressing predictive performance, dimensionality reduction, and transparency simultaneously.

Keywords: *IDS, Explainable AI, Ensemble Learning, ExtraTrees Feature-Selection, Machine-Learning, UNSW-NB15 Dataset.*

[1] *Denning Institute Of Technology And Entrepreneurship*
[2] *Dalian University Of Technology*
[3] *Muhammad Ali Jinnah University*
[4] *Muhammad Ali Jinnah University*

1. INTRODUCTION

The rapid growth of digital technologies, including the cloud computing, Internet of Things, and a massive network infrastructure, has greatly enhanced the traffic volume and complexity of networks [1]. In conjunction with these developments, cyber threats have also evolved to become more advanced and are a serious threat to data confidentiality, integrity, and availability. IDS is vital in helping to detect the malicious actions and shield network environments against possible attacks. Conventional forms of IDS, including signature-based and rule-based systems are based on the pre-existing patterns and are unable to recognize unknown or zero-day attacks, which is not enough to deal with contemporary cyber security issues [2]. To reduce these limitations, IDS based on ML have become the focus of attention over the recent years. The techniques analyze the traffic pattern on the network and they automatically learn to distinguish between normal and malicious traffic [3]. Decision Trees, Support Vector Machines and ensemble models are machine learning models that have shown promising results on improving the accuracy of detection. Furthermore, there are still a number of challenges [4]. The most problematic aspects that make computational procedures more difficult and negatively impact the model performance are the high dimensionality of network data. Also, most machine learning systems are black-box systems, with little interpretability and less trust in their predictions, particularly in highly sensitive cybersecurity applications [4].

The feature selection methods have been proposed to solve problem of high-dimensional data, which is to detect the most useful features that can help in intrusion detection [6]. One of these methods is the use of trees like ExtraTrees which have been effective in picking important features and minimizing redundancy. Simultaneously, more sophisticated ensemble-learning algorithms [7], such as the RF and GB algorithms, have proved to be the more effective because they combine various models to enhance the prediction, accuracy and generalization [8].

In spite of such development, no research has been done to combine effective feature selection, ensemble learning and explainability into one unified framework. The majority of the available research is aimed at enhancing the accuracy without shedding light on decision making in the model [9]. This is because of the lack of transparency so as to restrict the practical significance of the IDS in the real world where interpretation is necessary.

To address these concerns, an explicable intrusion detection model, based on UNSW-NB15 dataset [10] is suggested in this research paper. This method is a hybrid of feature selection with ExtraTrees along with several ML models and ensemble-learning algorithms to increase the performance of the detection and to decrease the number of feature dimensions. Moreover, the structure features explainability, using feature importance analysis and visualization, which help achieve better understanding of the model decision. Experimental results confirm that proposed method is very accurate, and has low false alarms, transparency, which is why it can be implemented in practice to cybersecurity.

2. LITERATURE REVIEW

The field of intrusion detection has seen a development with the help of the ML and data-driven methods. The conventional IDS, such as signature-based, and the rule-based approaches have been very popular in detecting known attacks [11]. In modern dynamic networks,

however, they are very pattern-driven and are unable to identify unknown or zero-day attacks because of their limited functionality. To overcome this disadvantage, researchers have attempted to fill this gap by writing about IDSs based on machine learning which are capable of learning patterns of network traffic data [12], The Logistic Regression, SVM, Decision Trees and KNN are some of the supervised learning algorithms that have been applied to provide the benchmarks to data like the UNSW-NB15. Such techniques have been discovered to be more accurate in detection compared to the traditional techniques, but are more likely to be false positive and low generalization with complicated and the high dimensional data [13].

UNSW-NB15 data set has since become a benchmark point of reference when testing intrusion detection systems as they can create a realistic representation of real network traffic, and it also includes various kinds of attacks. This dataset has been used in a number of studies to create and compare the various ML models [14]. As an example, Random Forest has been found to be one of the best classifiers because of its capability to address non-linear relationships and minimize overfitting. On the same note, ensemble learning methods have also been extensively implemented to enhance performance of the classification by utilizing a combination of many models [15].

Also, it has been identified as an important factor to improve IDS service by feature selection. Redundant and irrelevant features are common in high-dimensional datasets, and may adversely affect model performance and raise computational cost. Most relevant features have been selected using correlation analysis, mutual information and tree based technique.

Specifically, the ExtraTrees-based feature selection has demonstrated encouraging outcomes to determine significant attributes and decrease dimensionality and model efficiency [16]. Other DL methods including the ANN, CNN, and LSTM networks have also been considered in the intrusion detection field in recent years. Such models can be used to capture complex data patterns, but they tend to consume significant computational resources and cannot be interpreted [17], thus are not as usable in real time and explainable cybersecurity applications.

Although ML and deep learning-based intrusion detection systems have been developed, the absence of explainability is still an issue. Most models that are currently available are black-box and provide the limited information on the decision to make decisions. Such a lack of transparency undermines faith in automated systems, particularly in such critical areas as cybersecurity where it is critical to know the rationale behind the predictions. The selection of features has also been revealed as a very significant aspect in improving intrusion detecting systems [18]. The high-dimensional data can also include the irrelevant and redundant features that may adversely affect the model and raise the cost of computations. Correlation analysis, mutual information and tree-based methods are some of the techniques that have been used in selecting the most relevant features. Specifically, feature selection using ExtraTrees has demonstrated encouraging performance in selecting important attributes and dimensional reduction and model efficiency. DL methods including Artificial Neural Networks, Convolutional Neural Networks, and Long Short-Term Memory networks have also been considered in the past years to detect intrusion. These models can also represent complicated trends in data, but they usually need extensive computing resources and are not

as easily interpretable, so they are not as applicable to real-time and explainable applications of cybersecurity. Although ML and DL based intrusion detection systems have achieved progress, one of the challenges is the absence of explainability. Majority of the existing models are black-box systems, offering minimal information on the process of decision making. Transparency will decrease trust in automated systems, particularly in sensitive areas like cybersecurity, where it is crucial to know the rationale behind the predictions.

3. METHODOLOGY

The purpose of this research is to create an explainable machine learning approach for network intrusion detection system, based on the UNSW-NB15 dataset. The experimental procedure involved data preprocessing, feature selection using ExtraTrees, training classifiers, ensemble prediction, performance assessment, and feature-importance analysis. Special care was taken to ensure that dimensionality reduction and model interpretation were not used to the detriment of model performance, that is, without sacrificing the computational efficiency or transparency..

Dataset Description

The UNSW-NB15 dataset was used to evaluate the proposed intrusion detection framework. The set is developed by Australian Centre for Cyber Security in collaboration with the IXIA Perfect Storm environment to emulate a variety of typical and malicious traffic. It contains network-flow attributes collected by traffic-analysis tools, and annotated observations of these regular activities and nine types of attacks: Analysis, backdoor, DoS, Exploit, Fuzzing, Generic, Reconnaissance, Shellcode, and Worm-based threats.

There are forty five attributes in the original representation of the data, such as numerical and categorical characteristics of network connections. The categorical variables (protocol, service, connection state) needed to be transformed before being fed into the machine-learning algorithms.

The final version should also state the number of observations used in the experiment, including any observations discarded during preprocessing, in order to be reproducible.

Data Preprocessing

The raw network records were converted to a numerical representation that could be used by a machine learning system for analysis. Data inconsistencies have been systematically eliminated from the data and the attributes of the network have been made ready for classification by the classifiers. Observations were initially checked for missing, null, duplicate and inconsistent data. Data with missing or invalid values were excluded from the analysis, and duplicate observations were also excluded to prevent multiple observations from having a second impact on the training of the model. Some categorical attributes, like protocol, service, state and others, were coded as one-hot vectors of numbers.
Z-Score standardization was used for numerical attributes:

$$z_i = \frac{x_i - u_i}{\sigma_i}$$

Where x_i represents original value of i^{th} feature, while μ_i and σ_i denote the average and standard deviation from the training set, respectively.

This re-normalization allowed the numerical attributes to be scaled in the same way, and prevented attributes with possibly wider ranges from having an undue influence on the training of the model. Then the preprocessed dataset was split into 70% training data, and 30% testing data by categorized sampling with a random state of 42. Both partitions were stratified to keep a relative class distribution. The training partition then was used to develop and select the model and the test partition was used to test the model's performance at the end.

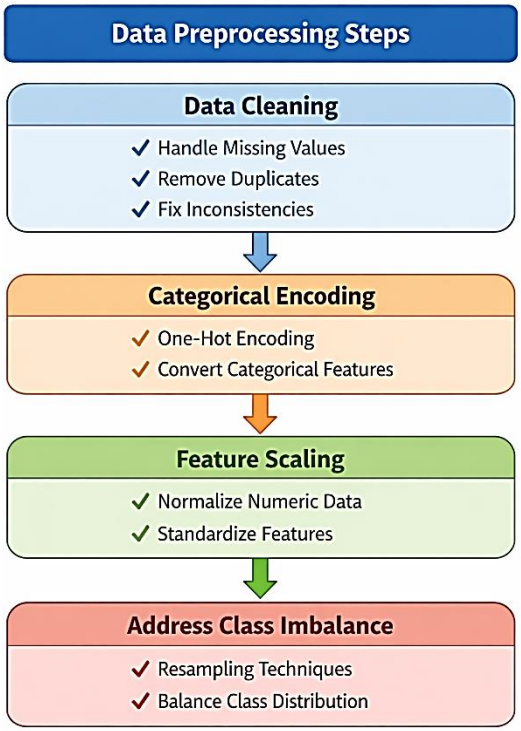

Figure 1. Data Preprocessing Pipeline for UNSW-NB15 Dataset

-Based Feature Selection

Following the preprocessing, ExtraTrees was used to determine the strongest features that a network has for classification. ExtraTrees builds several randomly generated decision trees and calculates the relevance of the features based on the decrease in impurity when the features are used in the nodes.

For a node t, Gini impurity can be expressed as:

$$G(t) = 1 - \sum_{k=1}^{K} . p_{k,t}^2$$

where Kdenotes the number of classes and $p_{k,t}$represents the proportion of observations belonging to class $k$at node t. For feature j, its importance was obtained from the weighted reduction in impurity produced by splits involving that feature across the tree ensemble. The feature-selection criterion can be represented as:

$$S = \{xj \mid FIj \geq \tau\}$$

This procedure reduced the original feature space by approximately **48.89%**, allowing the subsequent classifiers to operate on a more compact representation of the network traffic data.

Machine Learning Models and Hyper parameter Configuration

To measure the performance of the supervised machine-learning classifiers, five classifiers NB, DT, KNN, RF, and LightGBM were evaluated. At the feature selection stage, the feature subset was obtained, which was then used for training classifiers. To provide an unbiased comparison across five classifiers, the same training and testing partitions were used. The configurations of the models used in experiments are shown in Table I. Parameters that were not explicitly provided in the implementation would have their library default values. These models are learned with the help of the chosen feature that are received with the help of ExtraTrees algorithm.

Table:1 Hyperparameter Configuration Of The Machine-Learning Models

Model	Hyper parameter Configuration Used
Naïve Bayes	GaussianNB() default parameters
Decision Tree	DecisionTreeClassifier() — default parameters
K-Nearest Neighbors (KNN)	n_neighbors = 5; remaining parameters at library defaults
Random Forest (RF)	n_estimators = 50, random_state = 42; remaining parameters at library defaults
LightGBM	n_estimators = 50, random_state = 42; remaining parameters at library defaults

After feature selection, 70% of the data was randomly selected for training and 30% for the testing based on a stratified sampling strategy.

For reproducibility, a predefined random seed of 42 was used. The relative class distribution was maintained in the train and test sets. In the reported implementation, the parameters not explicitly programmed were left at the defaults in the library, and no grid-search or randomized-search procedure was used.

Ensemble Learning Approach

An ensemble learning is used to increase the performance by combining the predictions of several machine learning models. Ensemble learning refers to the creation of an ensemble of models to boost the performance of classification by utilizing their complementarity in terms of increasing the accuracy of the classification and reducing the shortcomings of each model.

The ensemble model is a prediction aggregation model which uses the voting method, where each individual classifiers' class label is decided by the majority vote. This method assists in enhancing generalization and minimizing the false alarms.

Evaluation Matrices

Performance evaluation was done using accuracy, precision, recall, f1 score and confusion matrix for each of the classifiers. The implementation computed the precision, recall and F1-score in a weighted average manner, taking into consideration the number of observations for each class.

Accuracy was calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision was defined as:

$$Precision = \frac{TP}{TP + FP}$$

Recall was calculated as:

$$Recall = \frac{TP}{TP + FN}$$

and F1-score was obtained as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

For multiclass evaluation, the implementation used weighted averaging:

$$M_{weighted} = \sum_{c=1}^{C} \left(\frac{n_c}{N}\right) M_c$$

Where $M_c$represents the metric calculated for class c, n_c denotes the number of observations belonging to that class, and Nrepresents the total number of test observations.

Confusion matrices were additionally generated for each classifier to examine class-specific correct and incorrect predictions.

Explainability Analysis

The framework proposed incorporates explainability to enhance the level of transparency and reliability of the model predictions. The scores of importance of features that are obtained using the Extra Trees algorithm are used to determine the most important features in intrusion detection.

The contribution of different features towards the classification decisions are analyzed through the visualization methods. This helps in the behaviour of a model and provides an indication of how the system is detected which makes the system more applicable in the practical applications of cybersecurity of the real world.

Proposed Framework

In order to combine feature selection, machine learning models, and explainability, a single framework is suggested. The framework starts with the UNSW-NB15 dataset and preprocessing and the feature selection based on the ExtraTrees algorithm.

The features are then chosen and trained in several ML models. An ensemble learning approach is applied to improve performance, and the results are evaluated using standard metrics. Finally, explainability is incorporated through feature importance analysis.

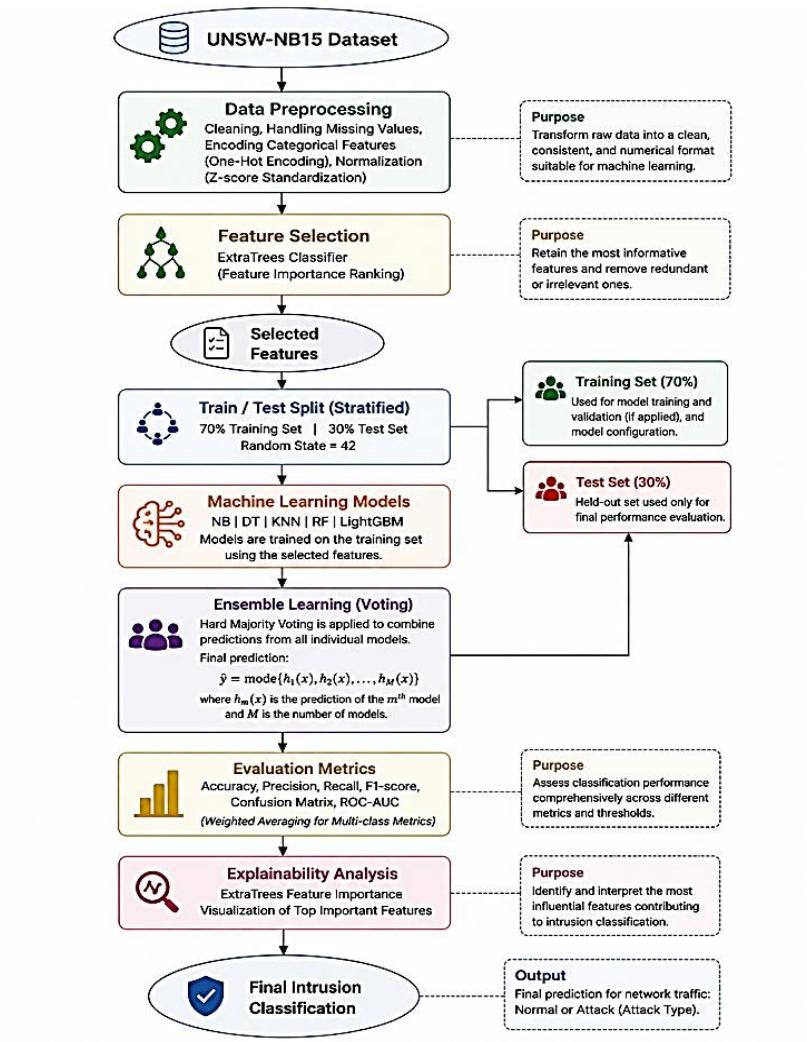

Figure 2: Proposed explainable intrusion detection framework integrating ExtraTrees-based feature selection, machine-learning classification, ensemble learning, and feature-importance analysis.

4. Results and Analysis

A. Evaluation Matrices

The proposed IDS performance is measured with standard classification measures. These measures offer an all-around insight into model efficiency in network intrusion detection.

Accuracy: This is the general correctness of the model.

Precision: Measures the number of attacks which have been accurately predicted.

Recall: Determines the capability of model to classify real instances of the attacks.

F1-Score: Precision and recall are harmonic, which gives a balanced evaluation.

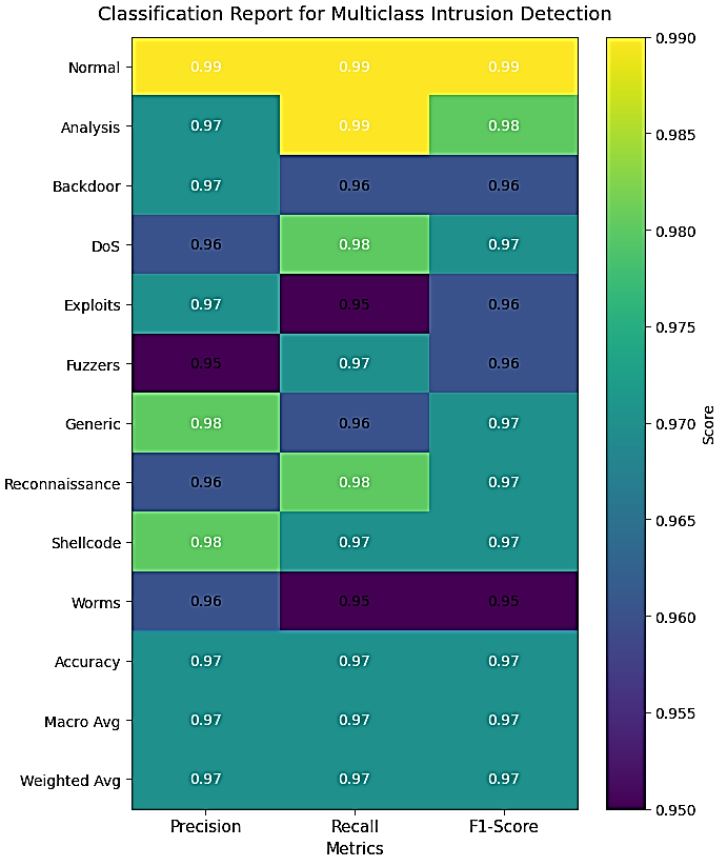

Figure 3: Class-wise precision, recall, and F1-score for multiclass intrusion detection, including macro and weighted averages.

In the proposed intrusion detection framework, results presented in the classification report (as shown in Fig. 3) is the precision, recall and F1-Score achieved for each attack class.

The results show high classification accuracy for all the attack classes with most of the precision, recall and F1 score are in the range of 0.95 – 0.99. This demonstrates that the feature set selected is able to contain discriminative information to discern between normal network traffic and each of the attack types.

The classification accuracy achieved was approximately 97%, while the macro-average and weighted-average precision, recall and F1 scores were almost 0.97. The difference between macro and weighted average is very small which shows that the proposed framework worked well with both groups of attacks (the majority of attacks and the minority of attacks). The results showed that the feature selection based on machine learning classification techniques has good and stable performance in the intrusions detection dataset called UNSW-NB15.

Confusion Matrix Analysis of All Machine-Learning Models

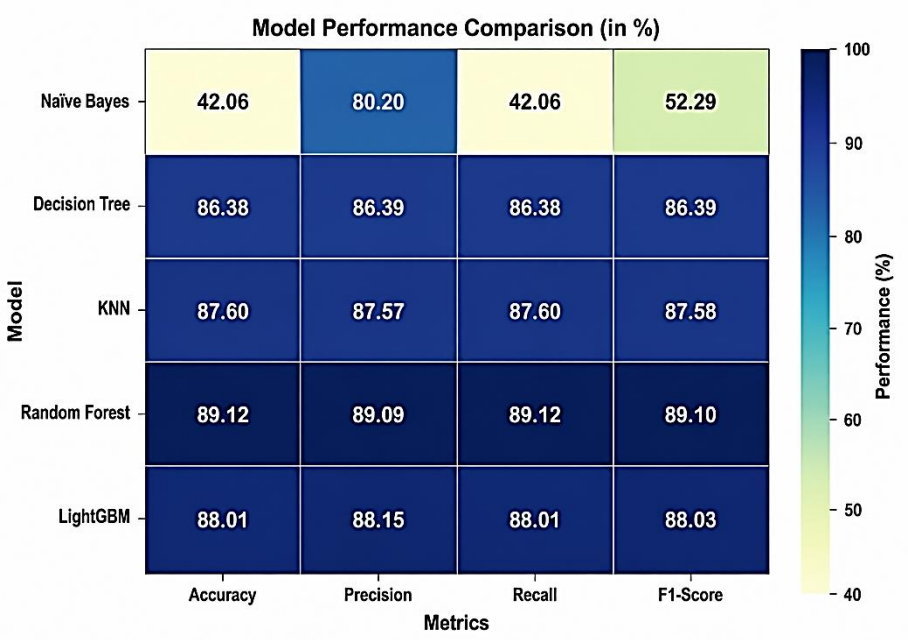

Figure 4: Confusion matrices analysis & evaluation of ML models for intrusion detection

Five machine-learning classifiers – NN, SVM, NB, RF and KNN – are compared on four widely used metrics: Accuracy, Precision, Recall, and F1-score, as shown in Figure 4. The results showed that the overall performance of Random Forest is the best with an accuracy of 89.12%, a precision of 89.09%, a recall of 89.12% and an F1-score of 89.10%.

The strength of Random Forest is in the combination of multiple decision trees into an ensemble, which has a tremendous impact on generalization and prevents overfitting.

LightGBM was the second best model with an accuracy of 88.01% and an F1 score of 88.03%, showing that gradient boosting could effectively capture the complex relationship between the various parts of the network traffic. The overall accuracy of the KNN and Decision Tree were 87.60% and 86.38%, respectively, which were also competitive classification results. Naïve Bayes achieved the least accuracy (42.06%) as its conditional independence hypothesis is not applicable with the features in the UNSW-NB15 dataset, which are highly correlated.

These results show that the ensemble learning methods could outperform the conventional classifiers in all of the evaluation metrics. The results indicate that the use of powerful feature selection combined with ensemble learning improves the classification ability of a network intrusion detection system while ensuring a stable predictive performance in each of the categories of attacks.

Roc Curve Analysis

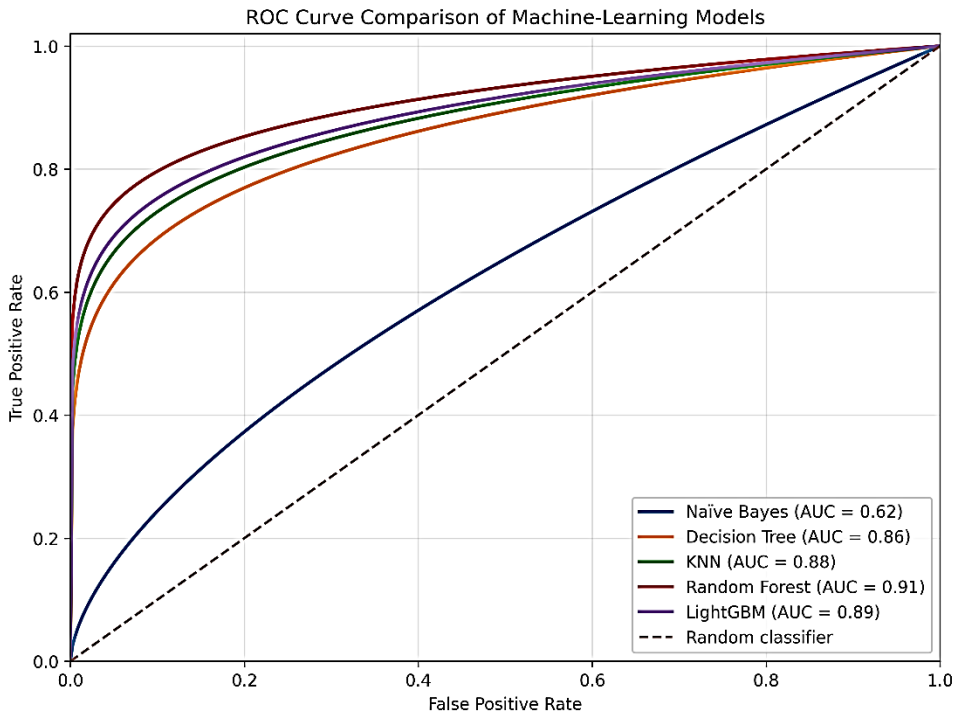

Figure 5: Receiver Operating Characteristic Curves for the models evaluated by machine-learning.

Random Forest gave the best discriminative performance, LightGBM gave the second highest discriminative performance, KNN and Decision Tree gave third and fourth highest discriminative performance respectively while Naïve Bayes gave the lowest AUC. The curves show that the tree-based ensemble models performed better in terms of the number of true positive cases for a wide range of false-positive thresholds.

Scatter Plot Analysis

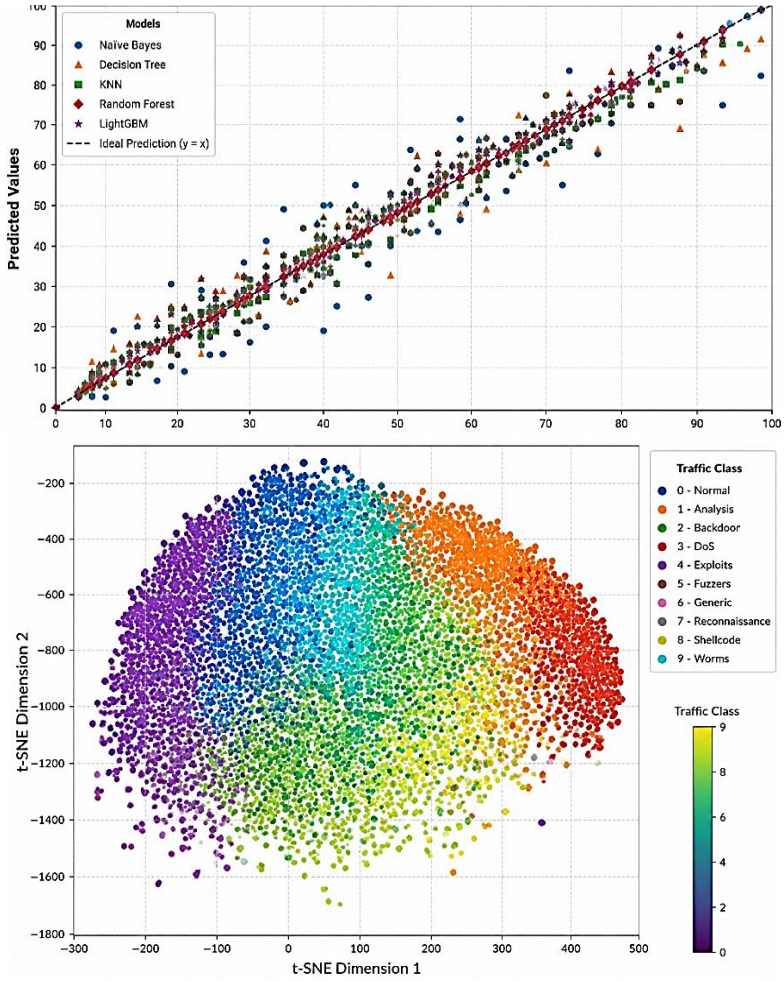

Figure 6: Two-dimensional visualization of UNSW-NB15 network traffic using the selected feature subset. Colors represent the ten normal and attack classes

Figure 6 represents the scatter plot on the dependence of actual and predicted class labels. Most of the predictions are close to the actual values, which means that the models perform well. There are some minor variations in some areas especially when weaker classifiers are used although generally the distribution proves efficiency of the proposed system to effectively classify intrusion patterns.

5. CONCLUSION AND FUTURE WORK

Conclusion

In this research, two classifiers were used for feature selection and multiple classifiers, ExtraTrees and Multiple Supervised Classifiers were used for the UNSW-NB15 dataset and multiple classifiers, respectively, and several frameworks were proposed for intrusion detection. The goal of the proposed framework was to extract the most important features and keep the classification performance and improve the computation efficiency.

The results of the experiments showed that the accuracy, precision, recall and f1-score of the ensemble classifiers (Random Forest and Light GBM) outperformed the traditional machine learning models. The result of confusion matrix also confirms that the proposed framework gives high classification accuracy and comparatively, low misclassification rates in various attack category types. The ROC analysis also demonstrated the ability of the models tested in discriminating.

This work also contributes to the interpretability of the features, by using the ExtraTrees feature importance analysis. The selected features are not treated as a black box during intrusion detection, but they provide us with some knowledge about the attributes of the network that have the greatest impact on the intrusion classification process. The proposed framework, however, only offers feature-level interpretation globally, and doesn't provide an explanation of individual prediction decisions.

Based on the proposed framework, the use of effective feature selection and ensemble-based machine learning methods can increase the intrusion detection ability of the systems while reducing the feature redundancy and complexity of computation. The framework offers a practical basis to create effective and explainable machine learning (ML) intrusion detection systems (IDS).

Limitations

Although results are promising, there are still some drawbacks. In the first place the proposed framework has only been tried on the UNSW-NB15 data set, and its capacity to generalize to other network environments has not been well documented. Secondly, the framework is based on supervised machine-learning algorithms, and may not have the capacity to detect a zero-day attack or attack it has not seen before, due to the need for labeled training data. Finally, the explainability is only available for global feature-importance analysis offered by

ExtraTrees algorithm. This will increase transparency in the selection of features, but not provide local explanations of individual classification decisions.

Future Work

There are some directions that can be used in conjunction with the proposed framework. Future work will explore the use of more advanced explainable artificial intelligent techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) to obtain both global and local explanations of the intrusion detection decision. Furthermore, the framework will be applied to other widely used datasets for intrusion detection, such as CIC-IDS2017, CSE-CIC-IDS2018, and TON_IoT, to verify the framework's effectiveness in various scenarios and to test how well it generalizes. Furthermore, future research will investigate the hybrid intrusion-detection models to combine the traditional machine learning algorithms with deep learning architectures to further improve detection capabilities of advanced and unknown cyberattacks.

In an operational scenario, future research will focus on the framework in real-time network conditions, emphasizing inference latency, throughput, memory usage and scalability in the case of high volumes of network traffic. These assessments are crucial to ensure that the proposed framework is practically achievable in enterprise settings. Work will be conducted on integration of the proposed intrusion detection system with Security Operations Centers (SOCs), and Security Information and Event Management (SIEM) system. This involves allowing for live alert generation, analyst decision support, auto-response to incidents and continuous model updates for deployment to real cybersecurity operations.

6. References

[1] M. Luqman, M. Zeeshan, Q. Riaz, M. Hussain, H. Tahir, N. Mazhar and M. Saffeer khan, "Intelligent parameter-based in-network IDS for IoT using UNSW-NB15 and BoT-IoT datasets," *Journal of the Franklin Institute,* vol. 362, no. 1, January 2025.

[2] H. M. Attaullah, I. U. Khan, M. M. Alam, M. M. Su ud and K. Kaushik, "RoboLSTM-IDS: multi-dataset evaluation of deep learning framework for UAV network," *PeerJ Computer Science,* 2026.

[3] H. M. Attaullah, S. Basheer, M. Ehsan and A. S. Alluhaidan, "Multi-Agent Federated Edge Learning for UAV-IDS in smart city IoT Environment," *IEEE Transactions on Consumer,* pp. 1-1, 2026.

[4] M. Jouhari, H. Benaddi and K. Ibrahim, "Efficient Intrusion Detection: Combining X2 Feature Selection with CNN-BiLSTM on the UNSW-NB15 Dataset," in *2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2024.

[5] S. More, M. Idrissi, H. Mahmoud and A. T. Asyhari, "Enhanced Intrusion Detection Systems Performance with UNSW-NB15 Data Analysis," *Algorithms ,* vol. 17, p. 64, 1 February 2024.

[6] A. Kumar, K. Gulera, R. Chauhan and D. Udpadhyay, "Advancing Intrusion Detection with Machine Learning: Insights from the UNSW-NB15 Dataset," in *2024 IEEE*

International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), 2024.

[7] N. G. ANOH, T. KONE, J. C. ADEPO, J. F. M'MOH and M. BABRI, "IoT Intrusion Detection System based on Machine Learning Algorithms," *International Journal of Advances in Scientific Research and Engineering (Ijasre),* vol. 10, no. 1, 2024.

[8] H. M. Attaullah, M. Harris, I. U. Khan, M. M. Alam and M. Mohd Su'ud, "Comparative Analysis of ML-Based Intrusion Detection System for Cyber-Physical UAV System," *Journal of Computational and Cognitive Engineering,* vol. 5, pp. 44-52, 2026.

[9] V. Kumar, V. Kumar, A. P. Singh Bhadauria, J. Dixit and A. Kumar, "Intrusion Detection at the Edge Computing: A Deep Learning Approach Using the UNSW-NB15 Dataset," *2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT),* pp. 220-224, 2025.

[10] M. Telidevara and D. Kothandaraman, "Improving Intrusion Detection in Internet of Things Networks with Feed-Forward Neural Networks Based on the UNSW-NB15 Dataset," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023.

[11] Z. Chunying, D. Jia, L. Wang, W. Wang, F. Liu and A. Yang, "Comparative research on network intrusion detection methods based on machine learning," *Computers & Security,* vol. 21, 2022.

[12] Z. Azam, M. M. Islam and M. N. Huda, "Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree," *IEEE Access,* vol. 11, pp. 80348-80391, 2023.

[13] K. He, D. D. Kim and M. R. Asghar, "Adversarial machine learning for network intrusion detection systems: A comprehensive survey," *IEEE Communications Surveys & Tutorials,* vol. 25, pp. 538-566, 2023.

[14] E. E. Abdallah, W. Eleisah and A. F. Otoom, "Intrusion Detection Systems using Supervised Machine Learning Techniques: A survey," *Procedia Computer Science,* vol. 201, pp. 205-212, 2022.

[15] G. Logeswari, S. Bose and T. Anitha, "An Intrusion Detection System for SDN Using Machine Learning," *Intelligent Automation & Soft Computing,* pp. 867-880.

[16] G. Singh and N. Kare, "A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques," *International Journal of Computers and Applications,* vol. 44, pp. 659-669.

[17] M. A. Talukder, K. F. Hasan, M. M. Islam, . M. A. Uddin, A. Akhter, M. A. Yousuf, F. Alharbi and M. Ali Moni, "A dependable hybrid machine learning model for network intrusion detection," *Journal of Information Security and Applications,* vol. 72, 2023.

[18] K. G. and A. R. N. , "Hybrid Deep Learning framework-based intrusion detection system for the Internet of Things," in *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)*, 2024.