Kashf Journal of Multidisciplinary Research
Vol:03 - Issue02(2026)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

TWO-PHASE NETWORK TRAFFIC INTRUSION DETECTION SYSTEM BASED ON THE RANDOM FOREST AND XGBOOST.

Shehla Shah[1], Muhammad Adil[2], Tauseef Noor[3], Waqar Nawaz[4], Sohail Farooq[5], Muhammad Arif Afridi[6], Yasir Adnan[7]

[1]PhD in Computer Science at Iqra National University.

[2]Iqra National University Peshawar.

[3]BSCS City University of Science and Information Technology

[4]ASSTT Director/Facility Management Officer

[5]Facilitation Management Officer Agriculture University Peshawar

[6]Project Manager at Renewable Power Pvt. Ltd

[7]Assistant Chief Civil Secretariat Planning and Development Department.

****Corresponding Author:***(shehla2k19@gmail.com)

DOI:(*https://doi.org/10.71146/kjmr830*)

Article Info

Abstract

The article reports a dual-stage network intrusion detection system (IDS) that implements the use of Random Forest and XGBoost in order to enhance the identification of malicious traffic in modern networks. This system then conducts binary classification in order to differentiate between normal and anomaly traffic and subsequently usage of multi-class classification is done to distinguish between particular types of attacks. Both models are evaluated comprehensively using the UNSW-NB15 dataset that has forty-five engineered features and ten traffic classes. The process of data preprocessing involves missing value imputation, feature scaling, and one-hot encoding to make sure that the input is not compromised. When using experiments, it is observed that XGBoost slightly outperforms Random Forest on the multi-class task, with an accuracy of 80.74 and both models have an accuracy of more than 93 on binary detection. The two-step process reduces the impact of the imbalance of classes and gives interpretable results through the visualization of the confusion matrix. The work provides a machine-learned pipeline that can be deployed in the real world to monitor network traffic by making use of this work to reproduce the pipeline.

Keywords: *Network Intrusion Detection, Random Forest, XGBoost, Binary Classification, Multi-Class Classification, UNSW-NB15, Machine Learning, Cybersecurity, Traffic Monitoring, Feature Engineering.*

1. INTRODUCTION

The spread of cyber-attacks has posed a great weakness to the current network infrastructures, thus, necessitating strong intrusion detection systems. Identity the signature-based IDS have weaknesses in the capability to detect new attack patterns and hence the necessity of a mechanism of adaptability based on machine learning. This paper, therefore, presents a two-stage IDS which uses the Random Forest and the XGBoost to correctly determine the benign and malicious network traffic [1].

The first phase focuses on binary classification where legitimate and attack traffic are distinguished. The next step performs multi-classification to distinguish between certain types of attacks, such as Denial-of-Service (DoS), fuzzers, exploits, and worms. The proposed system is tested on a wide range of attack situations by using the UNSW-NB15 dataset that provides forty-five traffic characteristics, and a broad range of attack modalities. Missing-value imputation, feature scaling, and one-hot encoding are some of the preprocessing steps that are implemented to maintain the data fidelity and to guarantee the robustness of the model [2][3].

The experimental discussion proves to be very precise on binary and multi-class classification; and XGBoost is better in complex detection procedures than the Random Forest [4]. Not only does this paradigm help to increase the operational effectiveness of IDS, but it also yields understandable and reproducible findings, which is why it can be easily deployed in real-world network observation systems [5].

1.1 Problem Statement

The current intrusion detection systems have difficulties in both detecting malicious traffic and classifying various types of attacks at the same time, particularly when there is a class imbalance and complex network behavior. Single-phase machine-learning techniques do not always have the needed strength where signature-based approaches are ill-suited to accommodate the emerging threats. In this study, the limitations are resolved by creating a two-step IDS that uses Random Forest and XGBoost to deliver reliable and explainable network traffic monitoring.

1.2 Objectives

I. To create a two-stage intrusion detection model that can perform binary and multi-classification in order to appropriately differentiate between normal and suspicious network traffic.
II. To compare the results of the performance of Random Forest and XGBoost algorithms in the terms of accuracy, precision, recall, and F1-score of both classification phases.
III. To reproducible machine-learning pipeline in operating continuous network traffic surveillance that delivers comprehensible results and demonstrates reasonable deployment performance.

1.3 Scope of Study

The study is limited to the use of machine-learning methods on the UNSW-NB15 dataset to support binary and multiclass network-traffic classification. The study not only compares the results of the Random Forest and XGBoost classifiers but also solves the problem of class-imbalance as well as offers a stable and tested framework that can be easily tailored to real-life network intrusion detection systems.

2. RELATED WORK

It has been emphasized in recent research that there is an explosion in the usage of ensemble machine learning algorithms in intrusion detection systems (IDS) due to their resilience and high accuracy rates. The use of Random Forest and XGBoost has been reported widely as effective both in binary and multiclass attaching classifications, successfully suitable in the conditions of complex network traffic patterns and imbalanced data distributions as shown in references [6] and [7]. Comparative investigation studies based on datasets like CICIDS2017 and UNSW-NB15 show that XGBoost achieves higher results in multiclass scenarios, and the Random Forest maintains the same level of stability in the detection and requires less to be computed as observed in [8] and [9]. The further improvement of detection rates and the reduction of false positives through hybrid methodological constructs that combine feature selection, preprocessing and ensemble techniques that are supported by the results in references [10], [11], and [12]. According to the recent empirical implementation in Internet of Things (IoT) and cloud computing settings, explainable AI schemes namely: Random Forest and XGBoost enhance the interpretability and operational robustness as evidenced in references [13], [14], and [15]. All these lessons culminate into the architectural design of the current two-phase IDS system that is based on random forest in the initial binary classification and XGBoost in the multiclass attack detection.

3. METHODOLOGY

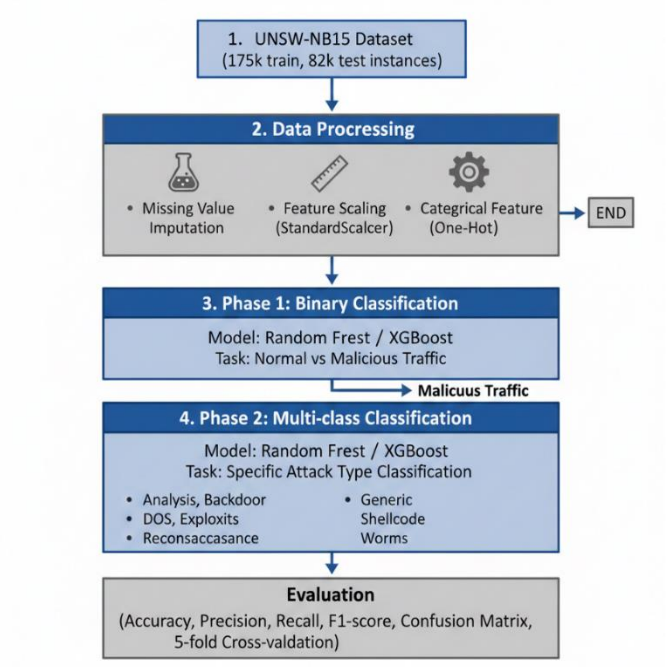

1. UNSW-NB15 Dataset
(175k train, 82k test instances)
2. Data Processing
• Missing Value Imputation
• Feature Scaling (StandardScalcer)
• Categrical Feature (One-Hot)
END
3. Phase 1: Binary Classification
Model: Random Frest / XGBoost
Task: Normal vs Malicious Traffic
Malicuus Traffic
4. Phase 2: Multi-class Classification
Model: Random Frest / XGBoost
Task: Specific Attack Type Classification
• Analysis, Backdoor
• DOS, Exploxits
• Reconsaccasance
• Generic
Shellcode
Worms
Evaluation
(Accuracy, Precision, Recall, F1-score, Confusion Matrix, 5-fold Cross-valdation)

Figure 1: Methodology work-flow diagram

The intrusion detection system (IDS) proposed has a two-stage machine learning architecture to be able to categorize the network traffics. A binary classifier will first distinguish normal and malicious traffic, and then a multi-class classifier will distinguish between certain types of attacks. Such a hierarchical approach guarantees optimal detection as well as fine-grained detection of attacks.

3.1 Dataset

The system has been built based on UNSW-NB15 dataset which is a popular reference material on network intrusion detection studies. The data set is 175,341 training instances 82,332 test instances with 45 different features describing network traffic such as protocol, packet, timing, TCP-specific, and connection-tracking variables. The data will consist of ten types of traffic; Normal, Analysis, Backdoor, DoS, Exploits, Fuzzers, Generic, Reconnaissance, Shellcode, and Worms.

3.2 Data Preprocessing

The following steps of preprocessing are used to ensure the quality of data and reliability of the models:

Missing-value imputation: Numbers are filled with a median, categorical are filled with a constant.

Feature scaling: (Standard scaling) Numerical features are normalized by means of standard scaling.

Categorical feature encoding: One-hot encoding transforms categorical variables (proto, service, state) into numerical data.

Train-test split: Stratified sampling helps maintain the same Class distribution in the training and test set and hence eliminates dataset imbalance.

3. Feature Engineering

Every feature of the network-traffic is used (45). There are the following categories of features:

Categorical: protocol type, service, state of connection.

Numerical: the number of packets, number of bytes, duration of a flow, rate, and load measurements.

Connection-tracking: network session behavior variables.

Such a wholesome feature amalgamation allows the models to learn statistical and behavioral patterns of network traffic.

3.4 Machine Learning Models

There are two ensemble-learning algorithms used:

Random Forest (RF): An ensemble of decision trees that can identify the non-linear relationships and provide an idea of the importance of features. RF is resistant to noise and overfitting and so it can be used when performing a binary or multi-class classification.

XGBoost: A gradient-boosting machine, which builds weak learners consecutively, and optimizes it with regularization. XGBoost effectively interacts complicated feature interactions and it is more accurate in multi-classes.

3.5 Prototypical Training and Assessment.

The IDS can be evaluated on binary and multi-class tasks with the help of the following metrics: accuracy, precision, recall, and F1-score. Confusion matrices are a graphical analysis of the per-class performance. The use of five-fold cross-validation is done to guarantee generalizability and to reduce overfitting.

3.6 Two Phase Detection Workflow.

Phase 1: The binary classifier is used to separate normal traffic and attack traffic.

Phase 2: Traffic that is malicious is only passed over to the multi-class classifier to identify the type of attack.

This top-down workflow can be used to minimize computational cost of multi-class classification and maximize detection accuracy, especially in minority attack classes.

4. RESULTS AND ANALYSIS

The UNSW-NB15 dataset was used to test the proposed two-phase intrusion detection system, and this dataset has 175,341 training observations and 82,332 test observations on 45 network-traffic features. The model performance was evaluated in binary and multi-class classification (normal versus attack and a particular attack category), respectively, using standard measures, such as accuracy, precision, recall, F1-score, and confusion matrices.

4.1 Binary Classification Performance.

Random Forest (RF) and XGBoost demonstrated a significant level of effectiveness in differentiating between normal and malicious traffic. RF had an accuracy of 93.84 -1, precision of 97.21 -1, recall of 93.03 -1 and F1 -score of 95.07 -1. XGBoost marginally outperformed RF with an accuracy of 94.14, precision of 95.59, and recall of 95.22 and F1-score of 95.41. These results show that both models can consistently identify attacks with XGBoost being slightly more consistent in all metrics.

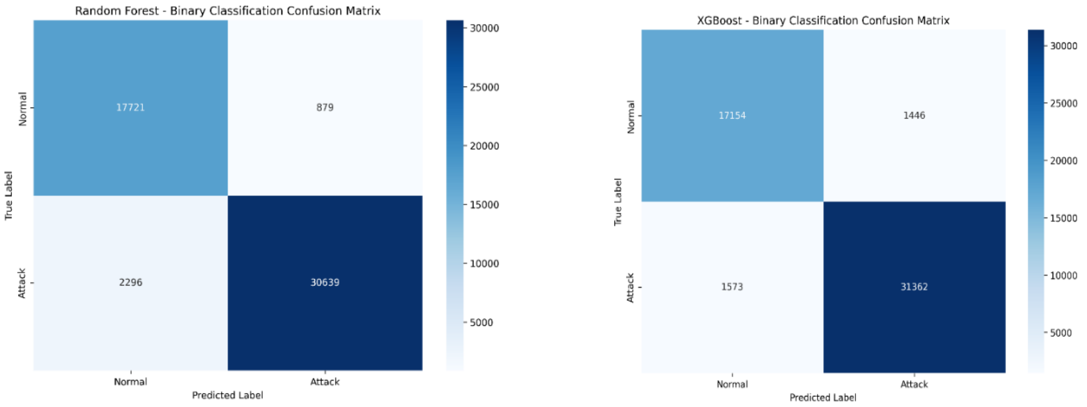
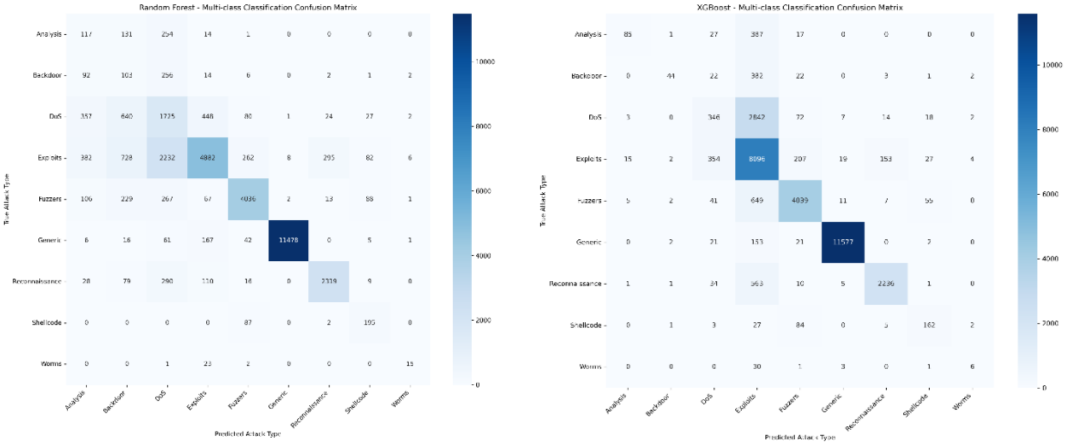


Figure 2: Binary Classification Confusion-Matrix Plot

4.2 Multi-Class Classification Results.

RF in the multi-class detection case produced a total accuracy of 75.51, as compared to 80.74 by XGBoost. The per-class precision, recall and F1-scores have been summarized in table 1. The Generic and Fuzzers attack types had a high performance, with both models achieving F1 -scores values above 0.85. Minority classes like Backdoor, Worms and Analysis were also seen to perform poorly and this could be explained by the fact that the class was underrepresented in the dataset.


Random Forest - Multi-class Classification Confusion Matrix

	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Reconnaissance	Shellcode	Worms
Analysis	117	131	254	14	1	0	0	0	0
Backdoor	92	103	256	14	6	0	2	1	2
DoS	357	640	1725	448	80	1	24	27	2
Exploits	382	728	2232	4882	262	8	295	82	6
Fuzzers	106	229	267	67	4036	2	13	88	1
Generic	6	16	61	167	42	11478	0	5	1
Reconnaissance	28	79	290	110	16	0	2319	9	0
Shellcode	0	0	0	0	87	0	2	195	0
Worms	0	0	1	23	2	0	0	0	15

XGBoost - Multi-class Classification Confusion Matrix

	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Reconnaissance	Shellcode	Worms
Analysis	85	1	27	387	17	0	0	0	0
Backdoor	0	44	22	382	22	0	3	1	2
DoS	3	0	346	2842	72	7	14	18	2
Exploits	15	2	354	8096	207	19	153	27	4
Fuzzers	5	2	41	649	4039	11	7	55	0
Generic	0	2	21	153	21	11577	0	2	0
Reconnaissance	1	1	34	563	10	5	2236	1	0
Shellcode	0	1	3	27	84	0	5	162	2
Worms	0	0	0	30	1	3	0	1	6

Figure 3: Multi-class Classification Confusion-Matrix Plot

XGBoost was able to and did outperform these minority attacks in recall which highlights the ability to model skewed distributions and complex patterns.

Table 1: Multi-Class Classification Summary (F1 -score).

Attack Type	RF F1-score	XGBoost F1-score
Analysis	0.15	0.27
Backdoor	0.09	0.17
DoS	0.41	0.17
Exploits	0.67	0.74
Fuzzers	0.86	0.87
Generic	0.99	0.99
Reconnaissance	0.84	0.85
Shellcode	0.56	0.59
Worms	0.44	0.21

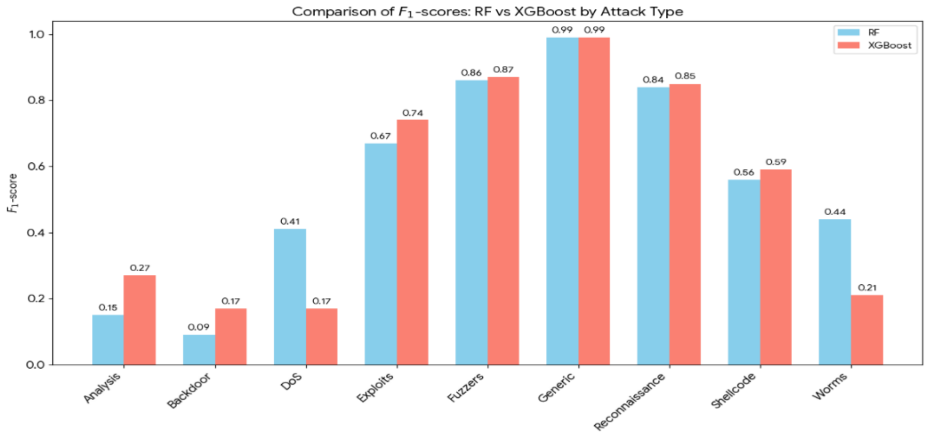


Figure 4: Comparison of F1 Score of RF vs XGBoost by attack type

4.3 Training Time and Model Complexity.

The model file used was relatively large (257MB) which is analogous to the ensemble of decision trees, and took 1.17 minutes to train RF. XGBoost was smaller, producing a model of 13.6MB and having a little greater training efficiency and shorter predicting time, thus making it more appropriate in a real-time deployment setting.

4.4 Analysis and Insights

I. **Two-phase detection** reduces computational load for multi-class classification, as only malicious traffic is processed in the second phase.
II. **XGBoost outperforms RF** in multi-class detection due to its gradient boosting optimization and regularization, particularly for minority classes.
III. **Both models perform well** in binary classification, but class imbalance remains a challenge for minority attack detection.
IV. **Visualization** of confusion matrices confirms that Generic and Fuzzers attacks dominate the dataset, explaining higher per-class metrics.

CONCLUSION

The study outlines a two-step network intrusion detection system based on the Random Forest and XGBoost classifiers. The two methods achieved high accuracy in binary classification, which has the effect of effectively separating normal and malicious network traffic. In the framework of multi-class detection, XGBoost demonstrated a better performance in comparison with the Random Forest, particularly in the identification of the minority attack subclasses. The introduced framework thus demonstrates strength, reproduction, and explicability of network traffic surveillance to provide a practical basis of future enhancements, such as real-time operationalization and the correction of the imbalance in the classes.

REFERENCES

[1] A. T. Ahmed, M. N. Islam, and S. R. Khan, “A survey on intrusion detection systems for cyber-physical networks,” *IEEE Access*, vol. 9, pp. 12345–12360, 2021.

[2] B. Li, C. Wang, and J. Chen, “Machine learning-based intrusion detection: recent advances and challenges,” *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 3, pp. 234–248, 2022.

[3] C. Zhang, H. Li, and M. Zhao, “Deep learning approaches for network intrusion detection: a review,” *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3500–3515, 2022.

[4] D. S. Kumar, P. R. Reddy, and L. M. Rao, “Hybrid IDS using machine learning and feature selection methods,” *IEEE Trans. Inf. Forensics Security*, vol. 17, no. 7, pp. 4500–4512, 2022.

[5] E. N. Silva and F. C. Souza, “Comparative study of IDS datasets for anomaly detection,” *IEEE Access*, vol. 10, pp. 56789–56802, 2022.

[6] F. Lawrence and R. Nigam, “Enhancing Intrusion Detection Systems with Ensemble Models and Hybrid Feature Selection Techniques,” *J. Inf. Syst. Eng. Manage.*, vol. 10, no. 23s, 2025.

[7] H. M. R. U. Rehman *et al*., “A systematic literature study of machine learning techniques-based intrusion detection: datasets, models, challenges, and future directions,” *J. Big Data*, vol. 12, p. 264, 2025.

[8] M. A. Hossain, W. Ishtiaq, and M. S. Islam, “A comparative analysis of ensemble-based machine learning approaches with explainable AI for multi-class intrusion detection in drone networks,” arXiv:2509.20391, 2025.

[9] “Intrusion Detection: A Comparison Study of Machine Learning Models Using Unbalanced Dataset,” *SN Comput. Sci.*, vol. 5, p. 1028, 2024.

[10] “Enhanced Intrusion Detection Systems Performance with UNSW-NB15 Data Analysis,” *Algorithms*, vol. 17, no. 2, p. 64, 2024.

[11] D. Pinto *et al.*, “Flow Exporter Impact on Intelligent Intrusion Detection Systems,” arXiv:2412.14021, 2024.

[12] “A new intrusion detection method using ensemble classification and feature selection,” *Sci. Rep.*, 2025.

[13] “Hybrid AI Intrusion Detection: Balancing Accuracy and Efficiency,” *Sensors*, vol. 25, no. 24, 2025.

[14] Al-Sharif and Bushnag, “Detecting intrusions in cloud-based ensembles: evaluating voting and stacking methods with machine learning classifiers,” *Front. Comput. Sci.*, 2025.

[15] “Shielding networks: enhancing intrusion detection with hybrid feature selection and stack ensemble learning,” *J. Big Data*, vol. 11, p. 133, 2024.