Archives available at journals.mriindia.com

International Journal of Recent Advances in Engineering and Technology

ISSN: 2347 - 2812

Volume 14 Issue 02s, 2025

Machine Learning Model for Efficient Botnet Attack Detection and Classification

[1]Sandhya C. Gaikwad, [2]Prof.R.H.Ambole

[1,2] Department of Computer Engineering (of Aff.) VidyaPratishthan's Kamal Nayan Bajaj Intstitute of Technology (of Aff.), Baramati,Pune Maharashtra,India

Email: sandhya.gaikwad242@gmail.com

Peer Review Information

Submission: 21 Oct 2025

Revision: 18 Nov 2025

Acceptance: 05 Dec 2025

Keywords

Botnet Detection, Machine Learning,XGBoost, Decision Tree, Logistic Regression, UNSW-NB15 Dataset, Cyber-security, Feature Selection, Attack Classification

Abstract

Botnets constitute a significant threat to cybersecu- rity, enabling large-scale malicious operations such as distributed denial-of-service attacks, data exfiltration, and unauthorized system access. This research presents a machine learning-based framework for the detection and classification of botnet attacks, utilizing Decision Tree, XGBoost, and Logistic Regression al- gorithms. The UNSW-NB15 dataset is employed, with distinct training and testing splits to ensure model generalization and to prevent overfitting. Feature selection techniques are applied to en-hance model performance and reduce computational complexity. Model evaluation is conducted using confusion matrices and Re- ceiver Operating Characteristic–Area Under Curve (ROC-AUC) metrics to provide a comprehensive assessment. Experimental results indicate that ensemble methods, particularly XGBoost, deliver superior performance in accurately detecting and cate- gorizing botnet traffic across various attack types. The findings highlight the effectiveness of machine learning approaches in improving the robustness and scalability of network intrusion detection systems.

INTRODUCTION

A botnet is a collection of compromised devices controlled by a hacker to execute cyber attacks like DDoS attacks (flooding a server with traffic),Spamming and phishing cam-paigns,Credential theft and financial fraud. Detecting botnets is challenging since they mimic normal network traffic. Ma- chine learning helps identify patterns in network behavior to detect botnet activity. There are currently available detection techniques for such stages, thus if a DDoS assault performed by an IoT Botnet has already taken place, identifying the DDoS attack, and the IoT Botnet network by itself at this point is not too challenging. The UNSW-NB15 dataset is used in this model which is a modern network intrusion dataset
created by the Australian Centre for Cyber Security (ACCS). The model is developed using Linear Regression,Decision Tree,XGBOOST Algorithms.

PROBLEM STATEMENT

The primary goal of this project is to develop an efficient, accurate, and scalable machine learning-based solution for detecting botnet attacks in IoT environments. Analyze network traffic data (e.g., packet size, frequency, IP address patterns,

protocol usage) and device behavior (e.g., communication patterns, data transfer rates) to identify features that can effectively distinguish between normal and malicious traffic. Build a machine learning model that can classify network traffic as either benign or indicative of a botnet attack.

MOTIVATION

The increasing sophistication of botnet attacks poses severe risks to network security, data integrity, and system availabil- ity. Traditional security mechanisms often struggle to detect complex, evolving botnet behaviors in real time. Therefore, there is a critical need for intelligent, data-driven solutions that can accurately detect and classify botnet activities. Leveraging machine learning offers the potential to enhance detection speed, accuracy, and adaptability, making it a promising ap- proach for building resilient cybersecurity defenses.

OBJECTIVE

The primary objective of this study is to develop an efficient and accurate botnet attack detection and classification sys- tem using machine learning algorithms on the UNSW-NB15 dataset. The specific objectives are: Analyze and preprocess the UNSW-NB15 dataset to extract relevant features for botnet detection. Implement and compare multiple machine learning
algorithms (Logistic Regression, Decision Tree, and XGBoost) to classify network traffic as DDoS,Mirai attacks, UDPFlood Attack etc. Evaluate model performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to determine the most effective approach

LITERATURE REVIEW

The detection of botnet attacks has evolved significantly with the adoption of machine learning (ML) techniques, addressing the limitations of traditional signature-based and anomaly-based detection systems. Early research focused on simple classifiers such as Decision Trees, valued for their inter- pretability but often prone to overfitting complex network traf- fic patterns. Ensemble methods, particularly Gradient Boosting algorithms like XGBoost, have shown superior performance by combining multiple weak learners and handling high- dimensional data more effectively, as highlighted in recent cybersecurity studies. Logistic Regression has been employed as a lightweight, interpretable model for binary classification but tends to underperform in capturing the non-linear behavior exhibited by modern botnet traffic. The UNSW-NB15 dataset, introduced by Moustafa and Slay (2015), has become a benchmark in intrusion detection research, offering realistic and diverse attack scenarios, including botnets. Feature selec- tion techniques have further improved detection models by reducing noise and focusing on the most predictive attributes. Despite these advancements, many studies report persistent challenges in balancing high detection accuracy with low false- positive rates, especially in dynamic and heterogeneous net- work environments, indicating the need for continued research into more adaptive and scalable ML-based solutions.

METHODOLOGY

The proposed methodology involves building a machine learning framework for botnet attack detection and classifica- tion using the UNSW-NB15 dataset. First, the data undergoes preprocessing steps, including handling missing values, encod- ing categorical features, and normalization. Feature selection techniques are applied to reduce dimensionality and retain only the most relevant attributes. Three models — Decision Tree, XGBoost, and Logistic Regression — are trained on the se- lected features using separate training and testing sets. Model performance is evaluated using confusion matrices, ROC-AUC scores, and other metrics such as accuracy, precision, recall, and F1-score. Finally, a comparative analysis is conducted to determine the most effective model for accurate and scalable botnet detection.

- The UNSW-NB15 dataset is a modern and comprehen- sive intrusion detection benchmark designed to capture realistic network traffic, including nine types of attacks and normal behavior. To prepare this dataset for effective botnet attack classification, several preprocessing steps are necessary. First, redundant and irrelevant columns such as IDs are removed, and any missing or infinite values are handled appropriately. Categorical features like protocol (proto), service (service), and connection

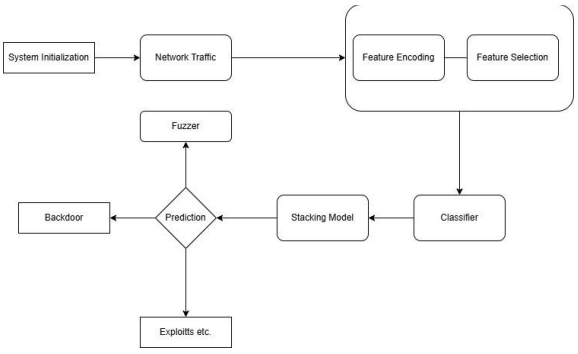

Fig. 1. Methodology

state (state) are then transformed using one-hot encoding, converting them into a numerical format suitable for machine learning algorithms.

Next, feature selection is applied to reduce dimensional- ity and improve model efficiency, often using statistical techniques like ANOVA F-score or feature importance from tree-based models to retain only the most relevant attributes. Given the class imbalance inherent in the dataset—where certain attack categories are underrep- resented—SMOTE (Synthetic Minority Over-sampling Technique) is used to generate synthetic examples of minority classes, ensuring balanced class distributions for more reliable training.

Once the data is preprocessed and balanced, multiple classifiers are trained, including Decision Tree, XGBoost, and Logistic Regression. Each offers unique strengths: Decision Trees are interpretable, XGBoost is powerful for complex patterns, and Logistic Regression serves as a strong linear baseline. To maximize classification performance, a stacking ensemble model is built by combining the predictions of these base learners, with a meta-classifier (such as Logistic Regression) making the final decision. This ensemble approach enhances accuracy and robustness, particularly for difficult-to-detect botnet and minority attack types, resulting in a comprehensive and effective botnet classification system.

I. Algorithms for Botnet Attack Classification

The goal of botnet attack detection and classification is to identify and categorize malicious network traffic from legitimate traffic, specifically detecting botnet-related activities in the dataset. The following algorithms are commonly applied to this task, each with their own strengths and theoretical foundations:

A. Decision Tree Classifier

The **Decision Tree** is a supervised learning algorithm that splits data into branches based on the most significant features, which makes it an ideal model for botnet attack detection.

- **Tree Structure:** A Decision Tree is constructed from the input features, creating binary decisions at each node.

These decisions are based on feature values that partition the dataset to minimize the impurity of the nodes.

- **Splitting Criterion:** The *Gini Impurity* or *Entropy* is often used as the splitting criterion to determine the best feature to split the data at each node. The Gini Impurity is given by:

$$G = 1 - \sum_{k=1}^{K} p_k^2$$

where p_k is the proportion of instances of class k at a node.

- **Training:** The tree is recursively split based on feature values that maximize the reduction in impurity. This process continues until leaf nodes are reached, where each leaf node corresponds to a classification label (e.g., Botnet or Normal).

- **Softmax Function:** After the final ensemble of trees is built, the output is converted into probabilities using the *softmax function*:

$$p(y = k|x) = \frac{\exp(\hat{y}_k)}{\sum_{j=1}^{K} \exp(\hat{y}_j)}$$

This allows for multi-class classification, where the class with the highest probability is chosen.

Example for Botnet Detection:: For a sample network traffic

- **Prediction:** Once the tree is built, new instances are classified by traversing the tree based on feature values. The final classification is the label at the leaf node, which represents the predicted class.

Example for Botnet Detection:: Consider a situation where network traffic features like sbytes (source bytes) and dbytes(destination bytes) are used. If a packet has sbytes < 1000 and dbytes < 100, the Decision Tree may classify it as **Botnet**.

B. XGBoost Classifier

XGBoost (Extreme Gradient Boosting) is an ensemble learning algorithm based on gradient boosting. It builds a series of decision trees where each subsequent tree attempts to correct the errors made by the previous ones.

- **Gradient Boosting:** XGBoost builds an ensemble of trees by fitting a new tree to the residual errors of the previous trees. The trees are added in a sequential manner, where each tree minimizes the loss function using the *gradient descent* method.
- **Loss Function:** The loss function to be minimized in XGBoost is typically a combination of a *logistic loss* (for classification) and a *regularization term* that prevents overfitting.

$$L(\theta) = \sum_{i=1}^{N} L(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(T_m)$$

where $L(y_i, \hat{y}_i)$ is the loss between the true label y_i and the predicted label $\hat{y}_i$, and $\Omega(T_m)$ is the regularization term for the m-th tree.

- **Boosting Process:** After each round of training, the prediction is updated as:

$$\hat{y}^{(t)} = \hat{y}^{(t-1)} + \eta \cdot T_t(x)$$

where η is the learning rate, and $T_t(x)$ is the prediction from the t-th tree.

instance, XGBoost might predict probabilities like:

$p(\text{Botnet}) = 0.75, \quad p(\text{Normal}) = 0.10, \quad p(\text{DDoS}) = 0.15$

The predicted class would be **Botnet** as it has the highest probability.

C. Softmax Regression

Softmax Regression is a generalized version of logistic regression used for multi-class classification problems. It is particularly effective when the output classes are mutually exclusive, as is the case in botnet attack detection.

Theory::

- **Logistic Regression:** Softmax Regression extends logis- tic regression to multiple classes. For each class k, a linear function is applied:

$$\hat{y}_k = w^T x_k + b_k$$

where w_k and b_k are the weight vector and bias term for class k, and x is the input feature vector.

- **Softmax Function:** The model computes a probability distribution over the classes using the *softmax function*, which normalizes the scores into probabilities:

$$p(y = k|x) = \frac{\exp(\hat{y}_k)}{\sum_{j=1}^{K} \exp(\hat{y}_j)}$$

This function ensures that the sum of probabilities across all classes equals 1, and the class with the highest probability is selected as the prediction.

Example for Botnet Detection:: Given a set of network traffic features, Softmax Regression might output the following probabilities:

$p(\text{Botnet}) = 0.70, \quad p(\text{Normal}) = 0.20, \quad p(\text{DDoS}) = 0.10$

The class with the highest probability, **Botnet**, would be selected as the predicted class.

DATASET

The UNSW-NB15 dataset is a comprehensive and widely used cybersecurity dataset designed for evaluating network intrusion detection systems. It was created by the Australian Centre for Cybersecurity (ACCS) at the University of New South Wales, and it contains a wide variety of network traffic data. The dataset includes 2.5 million network traffic records and is designed to simulate real-world network traffic with both normal and attack data. These records are classified into 9 attack categories, which include various types of cyberattacks
such as DoS (Denial of Service), scanning, backdoors, and shellcode injection, among others.

The dataset is composed of 49 features, including flow features like packet size, duration, and protocol type, as well as more detailed network activity data. The attack categories in the dataset are based on the classification of real-world network threats, making it a valuable resource for developing and testing machine learning models in intrusion detection systems. The dataset is well-structured for both binary (normal vs. attack) and multi-class classification tasks, providing an opportunity for researchers and practitioners to explore various machine learning algorithms and techniques to improve the detection and classification of different types of attacks.

Table I: ATTACK CATEGORIES IN THE UNSW-NB15 DATASET

Attack Type	Description
Normal	Legitimate network traffic.
Fuzzers	Sending malformed or random data to crash a service.
Analysis	Host scan, OS fingerprinting, and data theft activities.
Backdoor	Installing malicious software for remote control.
DoS (Denial of Service)	Overloading a system to disrupt services.
Exploits	Using system vulnerabilities to gain unauthorized access.
Generic	Cryptographic attacks on ciphers.
Reconnaissance	Scanning networks for vulnerabilities.
Shellcode	Injecting and executing malicious code via exploits.
Worms	Self-replicating malware spreading across networks.

```
Decision Tree Accuracy: 0.796237046669062
XGBoost Accuracy: 0.8533201019727275
Logistic Regression Accuracy: 0.7646186573590888

Classification Report (Decision Tree):
              precision    recall  f1-score   support

           0       0.20      0.15      0.17      2000
           1       0.33      0.09      0.14      1746
           2       0.31      0.56      0.40     12264
           3       0.77      0.44      0.56     33393
           4       0.67      0.76      0.71     18184
           5       0.92      0.99      0.95     40000
           6       1.00      1.00      1.00     56000
           7       0.84      0.75      0.79     10491
           8       0.14      0.44      0.21      1133
           9       0.42      0.48      0.45       130

    accuracy                           0.80    175341
   macro avg       0.56      0.57      0.54    175341
weighted avg       0.82      0.80      0.80    175341
```

Fig. 2. Classification report of decision Tree

RESULT

The Decision Tree classifier achieved an accuracy of 79.6percent on the UNSW-NB15 dataset, performing well on majority classes like Normal and Generic with high preci- sion and recall. However, it showed weaker performance on minority classes such as Analysis, Backdoor, Shellcode, and Worms due to class imbalance. The model is effective and interpretable but may need enhancements like class balancing or ensemble methods for better detection of less frequent attacks.

The XGBoost classifier achieved 85 percent accuracy on the UNSW-NB15 dataset, effectively detecting major classes like Normal and Generic with high F1-scores. However, it strug- gled with minority classes such as Analysis and Backdoor due to class imbalance. Despite a lower macro F1-score (0.59), the high weighted F1-score (0.83) indicates strong performance on dominant classes. Overall, XGBoost is effective for botnet detection but could benefit from class balancing for better minority class detection.

The Logistic Regression model showed weaker performance on the UNSW-NB15 dataset, struggling with multi-class dis- tinctions, especially for attacks like Exploits, Fuzzers, and rare types such as Worms and Shellcode. While it classified Normal and Generic traffic well, its linear nature and inability to capture complex patterns make it less suitable compared to advanced models like XGBoost.

The Hybrid model achieved 80.43 percent accuracy and a weighted F1-score of 0.81,

showing strong performance

Classification Report (XGBoost):

	precision	recall	f1-score	support
0	0.60	0.00	0.00	2000
1	0.82	0.07	0.14	1746
2	0.42	0.07	0.12	12264
3	0.62	0.86	0.72	33393
4	0.76	0.87	0.81	18184
5	0.98	0.99	0.99	40000
6	1.00	1.00	1.00	56000
7	0.86	0.77	0.81	10491
8	0.70	0.58	0.63	1133
9	0.69	0.63	0.66	130
accuracy			0.85	175341
macro avg	0.75	0.58	0.59	175341
weighted avg	0.84	0.85	0.83	175341

Fig. 3. Classification Report of Xgboost

overall. However, a low macro F1-score of 0.58 indicates poor classification of minority classes like 'analysis', 'backdoor', and 'worms', which were often misclassified as dominant classes such as 'dos' or 'exploits'. This highlights the effect of class imbalance, suggesting a need for techniques like over- sampling, class-weighted loss, or advanced ensemble methods to improve balance and accuracy.

✅ Accuracy: 0.8043388269385009
🎯 Macro F1 Score: 0.5816462545789346

📊 Classification Report:

	precision	recall	f1-score	support
analysis	0.22	0.18	0.19	803
backdoor	0.19	0.19	0.19	699
dos	0.36	0.75	0.49	4906
exploits	0.82	0.57	0.67	13358
fuzzers	0.64	0.64	0.64	7274
generic	1.00	0.98	0.99	17661
normal	0.93	0.89	0.91	27900
reconnaissance	0.90	0.79	0.84	4196
shellcode	0.39	0.90	0.54	453
worms	0.24	0.71	0.36	52
accuracy			0.80	77302
macro avg	0.57	0.66	0.58	77302
weighted avg	0.84	0.80	0.81	77302

Fig. 4. Classification Report of Linear Regression

Classification Report (Logistic Regression):

	precision	recall	f1-score	support
0	0.04	0.01	0.01	2000
1	0.05	0.37	0.09	1746
2	0.35	0.01	0.02	12264
3	0.68	0.61	0.64	33393
4	0.66	0.73	0.69	18184
5	0.79	0.99	0.88	40000
6	1.00	1.00	1.00	56000
7	0.69	0.40	0.50	10491
8	0.00	0.00	0.00	1133
9	0.00	0.00	0.00	130
accuracy			0.76	175341
macro avg	0.43	0.41	0.38	175341
weighted avg	0.76	0.76	0.75	175341

Fig. 5. Classification report of Hybrid Model

Fig. 6. Confusion Matrix for decision tree

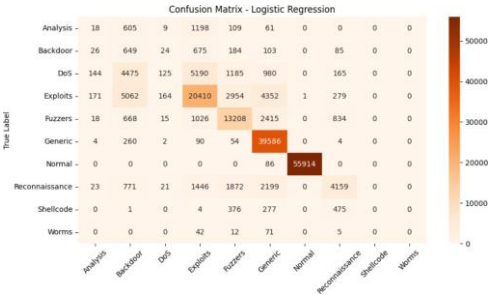

Fig. 7. Confusion Matrix for Xgboost

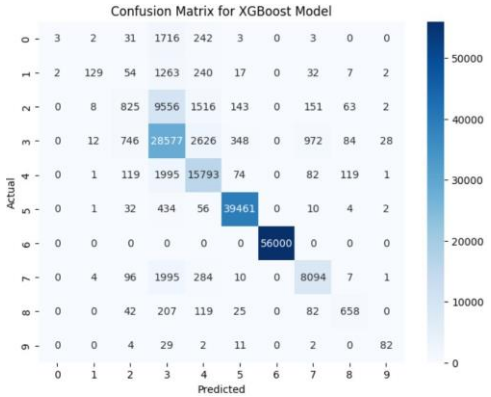

Fig. 8. Confusion Matrix for Linear Regression

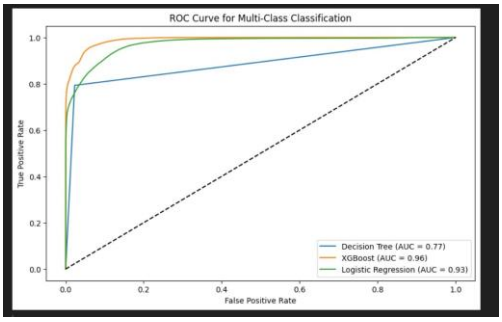

Fig. 9. ROC,AUC

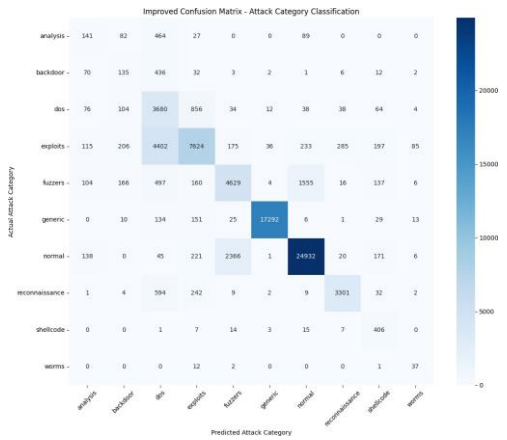

Fig. 10. ROC,AUC for hybrid model

Conclusion

In this research, an effective framework for botnet at- tack detection and classification was developed using the UNSW-NB15 dataset. The data underwent extensive preprocessing including one-hot encoding, feature selec- tion, and class balancing through SMOTE to address the challenges of high dimensionality and class imbal- ance. Multiple machine learning algorithms—Decision Tree, XGBoost, and Logistic Regression—were applied and evaluated for their performance in identifying and classifying various network attacks. To enhance clas- sification robustness, a stacking ensemble method was implemented, combining the strengths of individual clas- sifiers. The results demonstrated that ensemble models significantly improve the detection accuracy and classi- fication of both frequent and rare botnet attacks. This approach highlights the importance of data preprocessing and ensemble learning in building reliable and scalable intrusion detection systems for modern networks.

REFERENCES

"Botnet Attack Detection in IoT Using Machine Learning" Hindawi Computational Intelligence and Neuroscience Volume 2022, Article ID 4515642

"Botnet Attack Detection in SDN-Enabled IoT Using Machine Learn- ing" Sensors 2022, 22, 9837. https://doi.org/10.3390/s22249837

" BOTNET DETECTION USING VARIOUS MACHINE LEARNING

ALGORITHMS" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 09 Issue: 12 — Dec 2022 www.irjet.net

Network Traffic Visualization Coupled With Convolutional Neural Net- works for Enhanced IoT Botnet Detection DAVID ARNOLD , (Member, IEEE), MIKHAIL GROMOV, (Member, IEEE)

Hybrid Machine Learning Model for Efficient Botnet Attack Detection in IoT Environment MUDASIR ALI SULTAN ALFARHOOD

Latent Semantic Analysis and Graph Theory for Alert Correlation: A Proposed Approach for IoT Botnet Detection MOEMEDI LEFOANE 1 (Member,IEEE), IBRAHIM GHAFIR1, SOHAG KABIR 1, IRFAN- ULLAH AWAN1, KHALIL EL HINDI 2, AND ANAND MAHEN-

DRAN3

A. Alharbi and K. Alsubhi, "Botnet Detection Approach Using Graph-Based Machine Learning," in IEEE Access, vol. 9, pp.99166 99180,2021,10.1109/ACCESS.2021.3094183.

D. Nanthiya, P. Keerthika, S. B. Gopal, S. B.Kayalvizhi, T.Raja and

R. S. Priya, "SVM Based DDoS Attack Detection in IoT Using Iot-23 Botnet Dataset," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT),

N. Moustafa, B. Turnbull and K. -K. R. Choo, "An Ensemble Intrusion Detection Technique Based on Proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things,"in IEEE Internet of Things Journal, vol. 6, no. 3, pp. 4815-4830,June 10.1109/JIOT.2018.2871719.

A STUDY OF MACHINE LEARNING CLASSIFIERS FOR ANOMALY-BASED MOBILE BOTNET DETECTION Ali Feizollah1,

Nor Badrul Anuar2, Rosli Salleh3, Fairuz Amalina4, Ra'uf Ridzuan Ma'arof5,Shahaboddin Shamshirband