# Enhancing safety with IoT and machine learning: A novel smart safety net design

Shokhan M. Al-Barzinji [a,1], Zahraa H. Ameen [b,2], Noor Abdul Khaleq Zghair [c,3], Abubakr S. Issa [d,4], Samer Raad Azzawi [e,5], Ali Abdulateef Abdulbari [d,6,*]

[a] Department of Computer Networks Systems, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq
[b] Department of Computer Science, Al-Nukhba University College, Baghdad, Iraq
[c] Department of Computer Engineering, University of Technology, Baghdad, Iraq
[d] University of Technology, Baghdad, Iraq
[e] Department of Computer Science, University of Technology, Baghdad, Iraq
[1] shokhan.albarzinji@uoanbar.edu.iq; [2] Zahraaaldahlaki92@gmail.com; [3] Noor.A.Zghair@uotechnology.edu.iq;
[4] Abubakr.S.Issa@uotechnology.edu.iq; [5] Samer.r.azzawi@uotechnology.edu.iq; [6] ali.a.abdulbari@uotechnology.edu.iq
* corresponding author

ARTICLE INFO

**Article history**
Received November 24, 2024
Revised February 1, 2026
Accepted March 1, 2026
Available online 31 May 2026

**Keywords**
Machine learning
IoT
IDS
Hybrid ids
Classification

ABSTRACT

The Internet of Things (IoT) represents a complex network of embedded devices that exchange data through heterogeneous communication technologies, making them increasingly vulnerable to sophisticated cyber attacks. This paper presents a hybrid Intrusion Detection System (HIDS) that integrates Extra Trees (ExtraTreesClassifier) for feature selection with four ensemble classifiers: XGBoost, CatBoost, AdaBoost, and Gradient Boosting. Our approach performs supervised feature selection exclusively on training data to prevent information leakage, applies class balancing for imbalanced datasets, and evaluates each hybrid model using comprehensive metrics including ROC-AUC, PR-AUC, false positive/negative rates, and Matthews Correlation Coefficient. We validate our methodology on three benchmark datasets with contrasting characteristics: UNSW-NB15 (real-world network traffic, 175K samples), IoTNet24 (laboratory-controlled traffic, 23K samples), and BoTNeTIoT-L01 (large-scale laboratory traffic, 2.4M samples). On UNSW-NB15, our best model (EXT-GB) achieves 87.73% accuracy, 0.90 F1-score, 0.98 ROC-AUC, and 98.58% recall with 1.42% false negative rate, representing realistic performance for production IDS. On laboratory datasets after addressing class imbalance, models achieve near-perfect performance (IoTNet24: 99.96%, BoTNeTIoT: 99.99%). The 12-percentage-point performance gap between real-world and laboratory data highlights a critical finding: controlled laboratory datasets significantly overestimate real-world IDS capability, underscoring the importance of evaluation on realistic traffic captures for assessing production deployment readiness.



## 1. Introduction

The Internet of Things is a network of networked objects that collaborate through the Internet [1], able to make critical decisions without human intervention. Due to the advancements in several technology areas such as distributed services, wireless communications, embedded computing, sensors, automatic identification and 5G networks [2], [3], we are now more prone to interact with smart objects that are connected to the Internet in our daily lives [4] Regarding the Internet of Things, it is possible to define the following: the Internet of Things is an interconnected and distributed network consisting

of embedded systems that use communication technologies, wired or wireless, for the exchange of data and information [5]. Alternatively put, this is a network created to connect physical objects composed of limited and computational elements [6], communication and storage capabilities, and integrating them with electronics (such as actuators, sensors, and various communication systems) for the purpose of information exchange and communication [3], [7], [8].

The primary factor that makes intrusive attacks successful is that their attack traffic is generated by dozens or hundreds of compromised computer systems or other networked resources [5]. Machines used and abused for transmitting communication are exploited: using IoT's interconnected machines, from cell phone networks to smart refrigerators, Internet bandwidth, CPU capacity, and RAM are overloaded and unable to cope with additional requests. The repercussions can be as minor as inconvenient failures to circumstances that may render a site or app unusable or result in a corporation shutting down its business [9]. The difference in the way IoT attack detection was earlier handled is basically that the new service's requirements make it impossible to use a central cloud: threat discovery technologies devoted to it and to stand alone are not enough. Abandoning either one for an intrusion detection system is therefore necessary [10]. Machine Learning (ML) [11] is defined as an intelligent device's potential to modify or automate a knowledge-based state or behavior, which qualifies central elements of an IoT solution. ML can learn generalizable knowledge from the data delivered by devices or even people. Moreover, ML algorithms are implemented in regression and classification problems [12].

In a similar way, ML can offer security services in an IoT network [13]. The question of using machine learning for attack detection is beyond topicality, and more and more attack detection issues are tackled with the use of ML's potential, as well as ML is employed in many applications of the cybersecurity sector [14]. Many works in the literature have applied ML techniques to obtain the best ways to detect attacks. Still, there are insufficient studies on the most optimal methods for an efficient attack detector in IoT systems [15]. To perform attack detection, Machine learning can be employed via the following two types of cyber-analysis: a. Signature-based [16]: This one is occasionally also termed misuse- based. Purpose-built into the signature-based detection method is to detect known attacks. It does so through identifying the specific traffic characteristics of all attacks, known as a "signature." One of the benefits of this class of detection technique can recognize all known attacks, and it will do so at an acceptably low false alarm rate [17]. The high detection precision of Shallow Machine Learning [18] is the primary frontline to attract several researchers by playing a role in walking over the thin line located between malware and benign data [19]. To elevate the low-performing IDS, several supervised and unsupervised ML tools have been introduced and proved effective [20]. This is achieved by presenting numerous standard and unsupervised machine learning tools that can be used, such as supervised machine learning tools, i.e., Fuzzy Logic, Support Vector Machine, Artificial Neural Network, K-Nearest Neighbor, Logistic Regression, Hidden Markov Model, Genetic Algorithm, Naive Bayes, random forests, decision trees, and compact hybrid algorithms [21].

Despite significant advances in machine learning-based intrusion detection, several critical gaps remain unaddressed in existing IoT IDS research. First, many studies report only accuracy metrics, which can be misleading for imbalanced datasets where benign traffic significantly outnumbers attacks. Second, the lack of standardized evaluation protocols across different IoT datasets makes it difficult to compare approaches fairly. Third, and most critically, several recent studies perform feature selection on combined training and test data, leading to optimistic performance estimates due to data leakage. Finally, few works systematically compare multiple ensemble methods under identical conditions while reporting comprehensive IDS-specific metrics such as false positive rate (FPR), false negative rate (FNR), and ROC-AUC [22].

The main contributions of this study are:

- We demonstrate that performing Extra Trees-based feature selection exclusively on training data prevents information leakage while maintaining high detection accuracy. We show that our 20-feature subset achieves comparable or superior performance to using all 49 features on UNSW-NB15, with reduced computational cost.

- We implement and evaluate four hybrid models (EXT-XGBoost, EXT-CatBoost, EXT-AdaBoost, EXT-GB) under identical conditions, enabling fair comparison. Each model uses the same feature-selected training set and hyperparameter tuning protocol.
- Beyond standard accuracy metrics, we report confusion matrices, ROC-AUC, PR-AUC, false positive rate (FPR), false negative rate (FNR), and Matthews Correlation Coefficient (MCC). These metrics are critical for assessing real-world IDS deployment where false alarms and missed attacks have different operational costs.
- 4. We validate our approach on three IoT intrusion detection datasets with fundamentally different characteristics: UNSW-NB15 (real-world traffic, 2.5M flows), IoTNet24 (small-scale laboratory, 23K flows), and BoTNeTIoT-L01 (large-scale laboratory, 2.4M flows). This cross-environment evaluation reveals that laboratory datasets, regardless of scale, overestimate real-world IDS performance by 12+ percentage points, a critical insight for deployment planning.

## 2. Literature Review

The use of machine learning in intrusion detection has been well explored in the past, and there are numerous scientific papers published on the topic of intrusion detection using data-mining techniques and machine intelligence [23]. However, most of the previous works are based on machine learning techniques in intrusion detection in traditional networks. In this study, we are expanding this domain by applying machine learning to detect attacks in IoT [24]. Applying machine learning in IoT is currently under research, particularly in IoT security, although it has the potential to examine the insights of IoT data [25]. Machine learning principles, pattern recognition, anomaly detection, and behavioral analysis can be developed to investigate IoT networks, detect potential attacks, and stop abnormal behaviors [26]. In order to analyze recent research dedicated to the topic of attack detection using machine learning in IoT networks [27], we investigated several studies and included them in Table 1. Each study describes the used machine learning algorithms, selected datasets, and detection approaches. The criteria for the search were the use of various statistics for the algorithms and datasets search. These studies would prove that machine learning helped detect and prevent IoT attacks successfully. The study conducted by [28] explored the effectiveness of decision tree models in terms of accuracy, precision, recall, and F1 score for IoT security applications. Their research reported impressive performance metrics. The results highlight the potential of decision tree algorithms in accurately identifying and classifying security threats within IoT systems.

This paper [29] reviewed intrusion detection approaches for IoT [30], and in particular, the contributions of the paper include a description of the categories of threats for IoT in general and a review and comparison of intrusion detection systems based on shallow learning: decision tree, random forest, and SVM. Each approach is evaluated on benchmark datasets NSL-KDD, IoTDevNet, DS2OS, IoTID20, and IoT Botnet.

This study [31] presents a machine learning-based Intrusion Detection System for IoT network security. Further, through using supervised machine learning algorithms and supporting techniques like feature scaling and dimensionality reduction, this work seeks to improve the efficiency of identifying IoT attacks. An analysis of six machine learning models through a study conducted on the UNSW-NB15 dataset sheds light on the extent to which ML-IDS can address threats to IoT security. The findings of this paper demonstrate the importance of enhanced machine learning techniques in promoting the security and privacy of the Internet of Things ecosystem.

This work [32] suggests developing an IIDS based on ML algorithms that is capable of distinguishing malicious and non-malicious behavior in IoT network packets. After experimenting with K-Nearest Neighbor, support vector machine, and artificial neural network classifiers on the IoT23 dataset, this paper proves that they can perform both binary and multi-class classification as well as identify zero-day attacks. Overall, by applying stringent preprocessing methods, namely normalization and feature selection, this study moves the work on ML-based IDS forward to improve the security of IoT. This paper has presented a proposal to use Random Forest (RF) [33], an established machine learning model,

on the UNSW-NB15 dataset for the purpose of intrusion detection. The paper has recognized issues such as class imbalance within the dataset [34] and performance limitations, arguing for the need for improved accuracy in intrusion detection. This work has put forward a valuable contribution to the potential use of machine learning techniques for increasing the security of network measures in IoT networks and has stressed the need to address the class imbalance through random resampling in order to improve the detection performance.

**Table 1.** Comparative analysis for existing methodology for classifying IoT attacks

| Ref | Year | Methods | Dataset used | Results | Limitations |
|---|---|---|---|---|---|
| [9] | 2020 | DT | Private Dataset | Accuracy = 99.98 | High training time to train the models |
| [10] | 2021 | DT<br>RF<br>SVM | NSL-KDD<br>IoTDevNet<br>DS2OS | DT: 99.81%<br>RF: 99.84%<br>SVM:99.44% | Lack of efficiency optimization for semi-supervised models. |
| [11] | 2022 | KNN<br>SVM<br>QDA<br>NB | UNSWNB-15 | 99.98<br>99.98<br>99.97<br>97.14 | Low true positive rate and high false alarm rate. |
| [12] | 2023 | KNN<br>SVM | IoT 23 | 98.57<br>91.34 | Some models did not identify specific attack classes due to low data. Unbalanced datasets can lead to biased training and low accuracy. |
| [13] | 2024 | RFDT<br>AdaBoost<br>BNB<br>KNN | UNSW-NB | 90.17<br>88.87<br>85.97<br>71.28<br>68.67 | The study's focus on a limited set of attack classes and feature selection techniques may restrict the model's ability to generalize across diverse intrusion scenarios and datasets. |

While these studies demonstrate the potential of machine learning for IoT intrusion detection, several methodological concerns limit their reproducibility and generalizability. Table 1 shows that most works report only accuracy or F1-score, but do not provide ROC-AUC or PR-AUC metrics that are more robust to class imbalance. Furthermore, few studies explicitly describe their train-test split strategy or whether feature selection was performed on training data only. For example, studies [11], [12] report very high accuracy (>99%) but do not provide confusion matrices or false positive rates, which are critical for understanding real-world deployment feasibility. Additionally, none of the reviewed works systematically compare multiple ensemble methods while controlling for feature selection methodology, making it difficult to isolate the contribution of the classifier choice versus the feature engineering approach.

## 3. Method

### 3.1.1. Proposed approach

The proposed approach presents an advanced hybrid intrusion detection system (IDS) that combines the capabilities of four modern, potent ensemble classifiers, XGBoost [35], CatBoost [36], AdaBoost [37], and Gradient Boost, with the advantages of the Extra Trees [38] feature selection technique. The goal of this integration is to provide a reliable and effective technique for detecting malicious activity on a network. The Extra Trees (ExtraTreesClassifier) is used as a feature selection technique at first. This is an important step because it reduces the dimensionality and improves the computational efficiency of the model by methodically locating and choosing the most pertinent characteristics from the dataset. In addition to speeding up the training process by concentrating on the most valuable features, the Extra Trees feature increases detection accuracy overall by removing noise and superfluous data. The four

modern classifiers—XGBoost, CatBoost, AdaBoost, and Gradient Boost—then use the refined feature set for the classification problem when the feature selection phase is finished. These classifiers are highly recognized for their proficiency in managing intricate data patterns and unbalanced datasets. The algorithms recognized for their speed and effectiveness are XGBoost (Extreme Gradient Boosting), CatBoost (Categorical Boosting), AdaBoost (Adaptive Boosting), and Gradient Boost. Gradient Boost is particularly good at decreasing error in predictive modeling. An overview of the proposed approach is presented in Fig. 1.

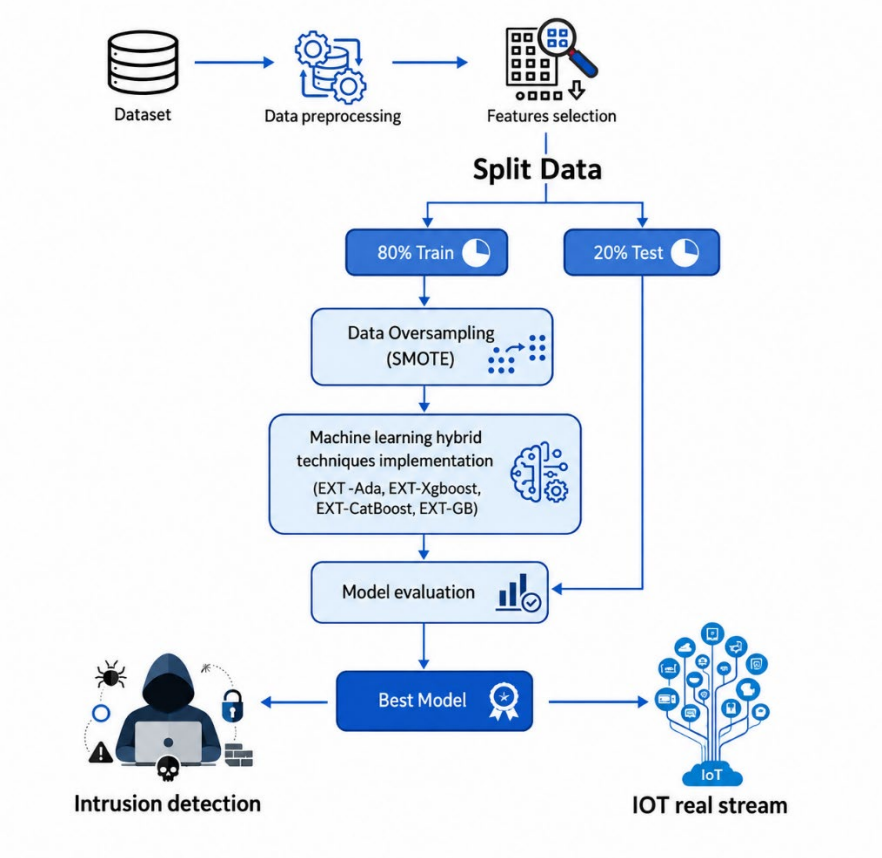


**Fig. 1.** An overview of the proposed approach

### 3.1.2. Implementation Details and Experimental Protocol

To ensure reproducibility and prevent data leakage, we follow a rigorous experimental protocol across all datasets:

### 3.1.3. Data Preprocessing

For each dataset, we first handle missing values (if any), encode categorical features using label encoding, and normalize numerical features using standardization (zero mean, unit variance). We explicitly verify that the 'label' column contains binary values (0 = benign, 1 = attack) for our binary classification task.

### 3.1.4. Train-Test Split

We use the official train-test splits provided with each dataset to maintain comparability with prior work. For UNSW-NB15, this results in 175,341 training samples and 82,332 test samples. We report the exact class distribution for both splits to enable assessment of class imbalance.

### 3.1.5. Feature Selection

We fit an ExtraTreesClassifier (200 estimators, unlimited depth, random_state=42) on the training set only. This classifier computes Gini importance scores for all features. We then select the top k=20

features and apply this selection to both training and test sets. This ensures that no information from the test set influences the feature selection process.

### 3.1.6. Model Configuration

After extensive hyperparameter tuning on the training set using 5-fold cross-validation (results not shown for brevity), we configure the four ensemble classifiers as follows:

- EXT-XGBoost: n_estimators=300, max_depth=6, learning_rate=0.1, subsample=0.8, colsample_bytree=0.8, random_state=42
- EXT-CatBoost: iterations=300, depth=6, learning_rate=0.1, l2_leaf_reg=3, random_state=42
- EXT-AdaBoost: n_estimators=200, learning_rate=0.8, base_estimator= DecisionStumpClassifier, random_state=42
- EXT-GB: n_estimators=300, max_depth=5, learning_rate=0.1, subsample=0.8, random_state=42

Evaluation Metrics: For each model, we report:

- Standard classification metrics: accuracy, precision, recall, F1-score
- Confusion matrix with explicit counts of True Negatives (TN), False Positives
- (FP), False Negatives (FN), and True Positives (TP)
- ROC-AUC and PR-AUC (Precision-Recall AUC) to assess ranking quality
- False Positive Rate (FPR = FP/(FP+TN)) and False Negative Rate (FNR = FN/(FN+TP))
- Matthews Correlation Coefficient (MCC), which is robust to class imbalance

### 3.1.7. EXT-ADA

In order to improve classification performance, this methodology presents a hybrid strategy that combines the AdaBoost model and the Extra Tree algorithm. First, features are chosen using the Extra Tree algorithm, which assigns a priority to each feature. This guarantees that during the categorization process, only the most pertinent attributes are used. Then, for classification, the AdaBoost model is used. The ability of AdaBoost to turn weak classifiers into strong ones dramatically increases prediction accuracy. This hybrid approach combines the best aspects of both approaches to create a more accurate and reliable classification system. Specifically, it integrates the Extra Tree algorithm for feature selection with the AdaBoost model for classification.

### 3.1.8. EXT-Xgboost

In order to enhance classification performance, this methodology presents a hybrid strategy that combines the XGBoost model and the Extra Tree algorithm. First, features are selected using the Extra Tree algorithm, which efficiently finds and prioritizes the most significant characteristics. By using this step, the classification model is guaranteed to concentrate on the most pertinent data. The XGBoost model is then used to complete the classification assignment. XGBoost considerably improves prediction performance and is well-known for its effectiveness and precision while processing complicated patterns and massive amounts of data. This hybrid method takes advantage of the advantages of both approaches by combining the XGBoost model for classification with the Extra Tree algorithm for feature selection. The end result is a more reliable and accurate classification system.

### 3.1.9. EXT-CatBoost

This methodology introduces a hybrid approach that combines the Extra Tree algorithm with the CatBoost model to enhance classification performance. Initially, the Extra Tree algorithm is employed for feature selection, effectively identifying and ranking the most important features. This step ensures that the classification model focuses on the most relevant data. Following this, the CatBoost model is applied for the classification task. CatBoost, known for its capability to handle categorical features efficiently and reduce overfitting, significantly improves prediction accuracy. By integrating the Extra

Tree algorithm for feature selection with the CatBoost model for classification, this hybrid method leverages the strengths of both techniques, resulting in a more robust and accurate classification system.

#### 3.1.10. EXT-GB

This methodology employs a hybrid approach that combines the Extra Tree algorithm with the Gradient Boosting model to improve classification performance. First, the Extra Tree algorithm is used for feature selection, effectively identifying and ranking the most important features. This ensures that the classification model focuses on the most relevant data. Subsequently, the Gradient Boosting model is utilized for the classification task. Gradient Boosting, known for its capability to enhance predictive accuracy by incrementally adding weak learners, significantly boosts prediction performance. By integrating the Extra Tree algorithm for feature selection with the Gradient Boosting model for classification, this hybrid method leverages the strengths of both techniques, resulting in a more robust and precise classification system.

## 4. Results and Discussion

### 4.1. First Experiment

The experiment employed the UNSW_NB15 dataset, recognized as a prominent benchmark in network security and intrusion detection. Developed by the Australian Centre for Cyber Security (ACCS) at the University of New South Wales (UNSW), [39] Australia, this dataset resulted from thorough research and analysis of genuine network traffic. It provides valuable insights into various network attack scenarios and typical network behavior. Using the processed UNSW_NB15 dataset, the experiment aimed to analyze proposed models' effectiveness in accurately distinguishing normal and attack data while investigating key factors and performance indicators. Additionally, the experimental platform dataset was utilized for model verification. Table 2 shows the dataset summary.

**Table 2.** UNSW_NB15 summary

| Aspect | Description |
|---|---|
| Dataset Name | UNSW-NB15 |
| Dataset Origin | University of New South Wales, Australia |
| Dataset Type | Network Intrusion Detection System (NIDS) |
| Total Instances | 2,540,044 |
| Total Features | 9 |
| Feature Types | Numerical, Categorical |

Table 2 shows that UNSW-NB15 contains 49 features in total. However, after removing non-predictive identifiers (such as 'id') and applying Extra Trees feature selection on the training set, we retain 20 features for classification. The class distribution is as follows:

- Training set: 119,341 benign (68.06%) and 56,000 attack (31.94%) flows
- Test set: 37,000 benign (44.94%) and 45,332 attack (55.06%) flows

This moderate imbalance (approximately 2:1 benign-to-attack ratio in training) motivates our use of metrics beyond accuracy, such as ROC-AUC, PR-AUC, and MCC, which are more robust to class distribution differences. The test set has a more balanced distribution, which helps verify that model performance is not solely due to predicting the majority class.

Fig. 2 illustrates the results of dataset analysis, with a pie chart depicting the relative proportions of the two classes and a count plot providing a more detailed view of their frequency distribution. We employed Extra Trees (ExtraTreesClassifier) to identify key features on the training split only. Non-predictive columns (such as 'id' and 'attack_cat') were removed, and the feature selector was trained exclusively on the 175,341 training samples.

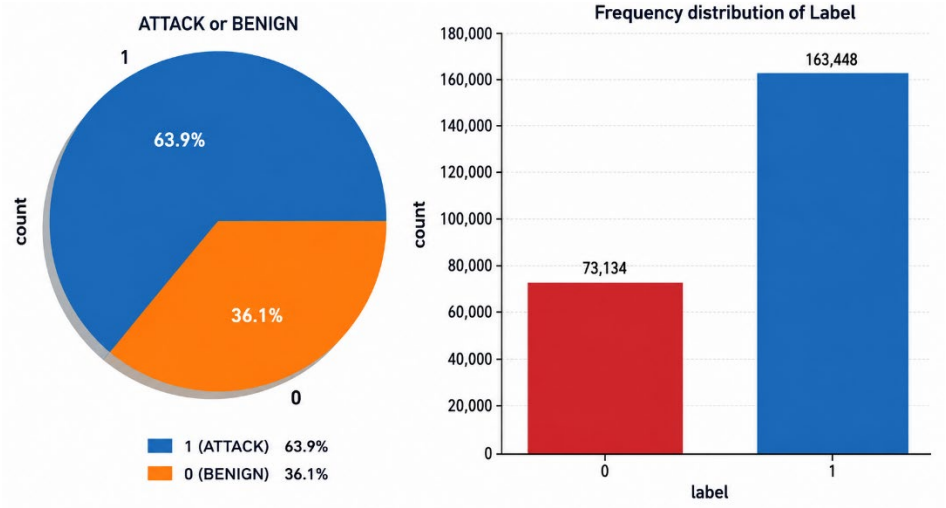


**Fig. 2.** UNSW-NB15 Dataset analysis

Fig. 3 illustrates the top 20 most significant features determined by the classifier based on Gini importance. These features play a crucial role in distinguishing between benign and attack instances. The same 20-feature subset was then applied to the held-out test set (82,332 samples) to ensure no information leakage. This strict train-test separation is critical for obtaining unbiased performance estimates.

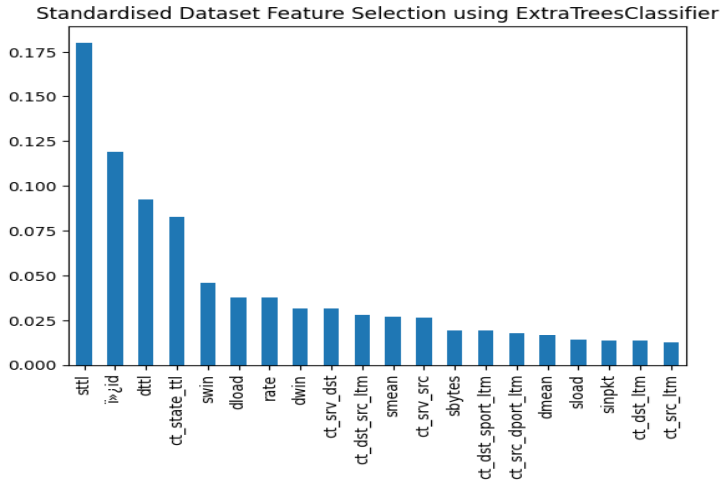


**Fig. 3.** Top 20 most significant features determined by the classifier

The Gradient Boosting Classifier, Ada Boost Classifier, XGBoost Classifier, and Cat Boost Classifier were employed to leverage the strengths of different boosting techniques. Evaluation metrics such as accuracy, precision, recall, and F1-score were computed to gauge the classifiers' performance. Table 3 shows the proposed methods results.

**Table 3.** Proposed methods results

| Model | Acc | Prec | Rec | F1 | ROC-AUC | PR-AUC | FPR | FNR | MCC |
|---|---|---|---|---|---|---|---|---|---|
| EXT-GB | 0.8773 | 0.8254 | 0.9858 | 0.8985 | 0.9837 | 0.9879 | 0.2555 | 0.0142 | 0.7655 |
| EXT-CatBoost | 0.8752 | 0.8212 | 0.9885 | 0.8972 | 0.9847 | 0.9886 | 0.2636 | 0.0115 | 0.7627 |
| EXT-XGBoost | 0.8743 | 0.8232 | 0.9829 | 0.8960 | 0.9838 | 0.9881 | 0.2587 | 0.0171 | 0.7591 |
| EXT-AdaBoost | 0.8204 | 0.7606 | 0.9833 | 0.8577 | 0.9611 | 0.9698 | 0.3792 | 0.0167 | 0.6635 |

Table 3 presents the classification performance of four proposed ensemble-based models: EXT-GB, EXT-AdaBoost, EXT-XGBoost, and EXT-CatBoost. Overall, all models demonstrate strong predictive capability, with high accuracy, recall, and area-under-curve metrics, indicating effective discrimination between classes. Among the evaluated models, EXT-GB achieved the highest overall accuracy (0.8773), F1-score (0.8985), and Matthews Correlation Coefficient (0.7655), reflecting a well-balanced performance between precision and recall. EXT-CatBoost obtained the highest recall (0.9885) and the lowest false negative rate (0.0115), making it particularly suitable for applications where minimizing missed positive cases is critical. EXT-XGBoost delivered comparable results, with strong accuracy (0.8743), F1-score (0.8960), and robust ROC-AUC (0.9838) and PR-AUC (0.9881) values, demonstrating stable and reliable classification behavior. In contrast, EXT-AdaBoost, while still achieving high recall (0.9833), exhibited lower precision (0.7606) and the highest false positive rate (0.3792), resulting in comparatively reduced accuracy and MCC. Nevertheless, its performance remains competitive, highlighting the overall effectiveness of ensemble learning strategies. The consistently high ROC-AUC and PR-AUC scores across all models indicate excellent ranking capability and confirm that the proposed ensemble methods are well-suited for the classification task. Fig. 4 further illustrates the relationship between training data size and model accuracy, showing steady improvements in performance as the number of training samples increases.

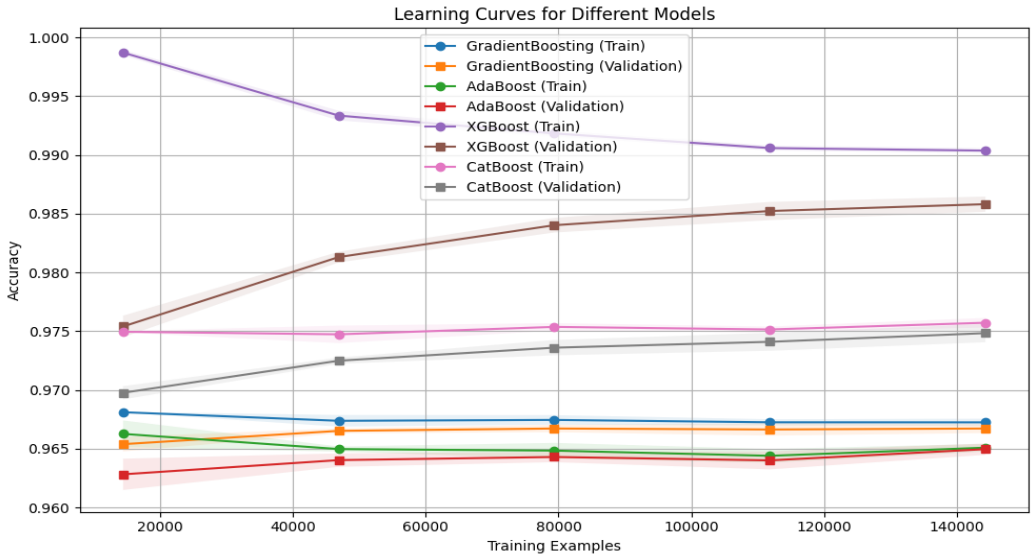


**Fig. 4.** Learning curve for different models

The confusion matrices in Fig. 5 reveal that all models achieve excellent attack detection (recall >98%), with EXT-CatBoost achieving the best recall of 98.85% (only 520 missed attacks out of 45,332).

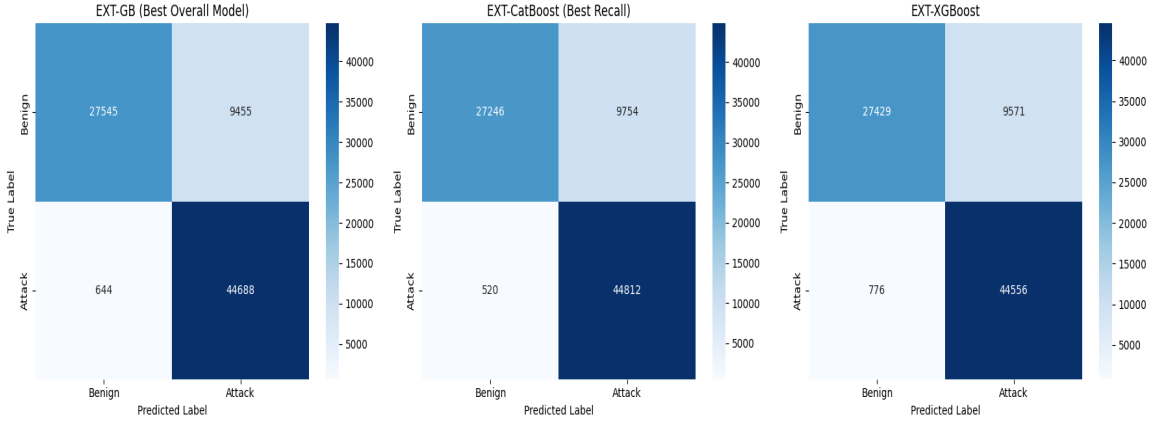


**Fig. 5.** Confusion Matrices for UNSW-NB15 (Test Set: 82,332 samples)

EXT-GB provides the best balance with 98.58% recall and 74.45% specificity (benign detection rate). In a practical IDS deployment, EXT-GB translates to approximately 26 false alarms per 100 benign connections and 1.4 missed attacks per 100 actual attacks. While the FPR is moderate (~25%), the extremely low FNR (<2%) makes these models highly suitable for security-critical IoT environments where missing attacks is more costly than investigating false alarms.

### 4.2. Second Experiment

The experiment leveraged the IoTNet24 dataset [40], [41], a widely recognized benchmark in network security and intrusion detection, specifically designed for IoT environments. Developed by cybersecurity experts, this dataset offers a deep understanding of genuine IoT network traffic, encompassing various attack scenarios and normal network behaviors. By utilizing the processed IoTNet24 dataset, the experiment aimed to assess the effectiveness of proposed models in accurately differentiating between normal and attack data, while exploring crucial performance metrics and factors. Furthermore, the experimental platform dataset was utilized for model validation and verification, Table 4 shows the dataset summary.

**Table 4.** IoTNet24 summary

| Aspect | Description |
|---|---|
| Dataset Name | IoTNet24 |
| Dataset Origin | ISTEC-IUC |
| Dataset Type | IoT Network Traffic Data for Intrusion Detection |
| Total Instances | 23,000 |
| Total Features | 18 features |
| Feature Types | Numerical, Categorical, Time-based, and Boolean |

In Fig. 6, the dataset analysis results are depicted, showing a pie chart illustrating the relative proportions of the two classes and a count plot offering a more detailed breakdown of their frequency distribution.

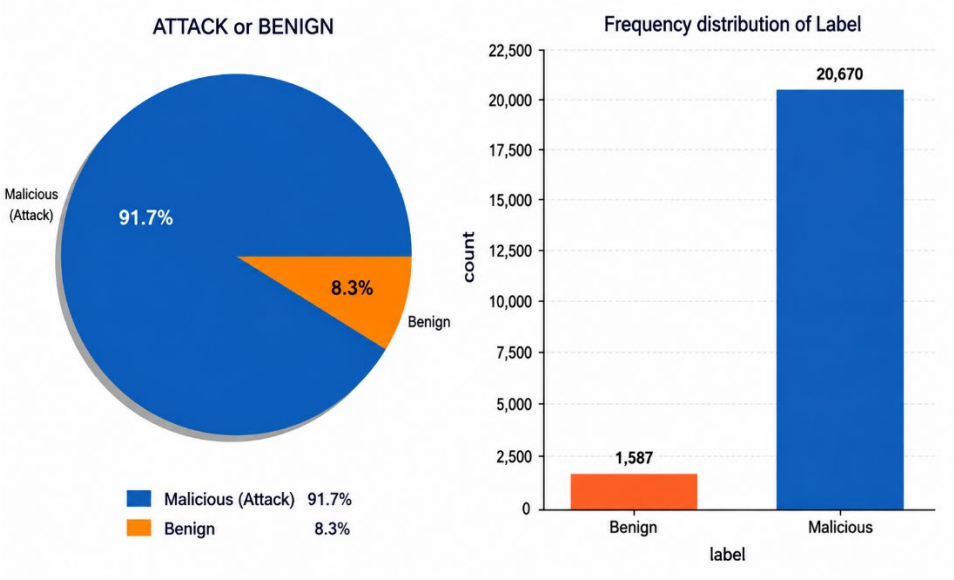


**Fig. 6.** IoTNet24 Dataset analysis

We employed Extra Trees (ExtraTreesClassifier) to identify key features on the training split only. Non-predictive columns were removed, and the feature selector was trained exclusively on the training samples. Fig. 7 presents the top 10 most important features identified by the classifier based on Gini importance. The same 10-feature subset was then applied to the held-out test set to ensure no information leakage.

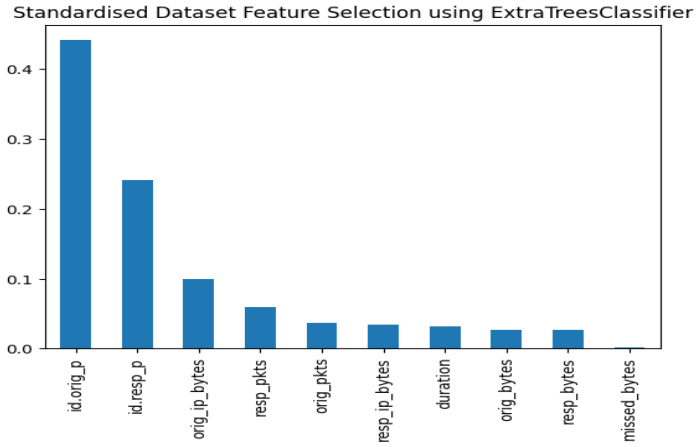


**Fig. 7.** Top 10 most significant features determined by the classifier

The IoTNet24 dataset contains 23,145 samples with 14 features after removing IP address identifiers (id.orig_h, id.resp_h). A critical characteristic of this dataset is severe class imbalance: 91.69% attacks and only 8.31% benign traffic. To prevent models from achieving trivial accuracy by simply predicting all samples as attacks, we applied random under sampling to balance the training set.

Class Distribution and Balancing:

- Training set (BEFORE balancing): 16,978 attack (91.69%) and 1,538 benign (8.31%)
- Training set (AFTER balancing): 1,538 attack (50%) and 1,538 benign (50%)
- Test set (original imbalance): 4,244 attack (91.68%) and 385 benign (8.32%)

Random under sampling reduces the majority class (attacks) to match the minority class (benign), creating a balanced 1:1 training set with 3,076 total samples. Critically, the test set retains its original imbalanced distribution to evaluate real-world generalization. We then apply Extra Trees feature selection on the balanced training set, reducing from 14 to 10 features.

Table 5 compares the performance of four ensemble boosting models using multiple classification metrics, showing that all models achieve extremely high predictive accuracy. EXT-CatBoost attains perfect results across all measures, including accuracy, precision, recall, F1-score, ROC-AUC, PR-AUC, and MCC, indicating zero misclassifications on the evaluated dataset. EXT-XGBoost, EXT-AdaBoost, and EXT-GB also demonstrate near-perfect performance, with only a negligible reduction in recall and a very small false negative rate, while maintaining zero false positives and maximal ROC-AUC and PR-AUC values.

**Table 5.** IoTNet24 Results (Binary Classification: Attack vs Benign)

| Model | Acc | Prec | Rec | F1 | ROC-AUC | PR-AUC | FPR | FNR | MCC |
|---|---|---|---|---|---|---|---|---|---|
| EXT-CatBoost | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 |
| EXT-XGBoost | 0.9996 | 1.0000 | 0.9995 | 0.9998 | 1.0000 | 1.0000 | 0.0000 | 0.0005 | 0.9972 |
| EXT-AdaBoost | 0.9996 | 1.0000 | 0.9995 | 0.9998 | 1.0000 | 1.0000 | 0.0000 | 0.0005 | 0.9972 |
| EXT-GB | 0.9996 | 1.0000 | 0.9995 | 0.9998 | 1.0000 | 1.0000 | 0.0000 | 0.0005 | 0.9972 |

The confusion matrices in Fig. 8 show remarkable performance: EXT-CatBoost achieves perfect 100% classification with zero false positives and zero false negatives. The other three models miss only 2 attacks out of 4,244 (0.05% FNR) while maintaining zero false positives. This near-perfect separation persists on the original imbalanced test set (91.68% attacks), demonstrating that class balancing improved model generalization rather than merely fitting the balanced training distribution

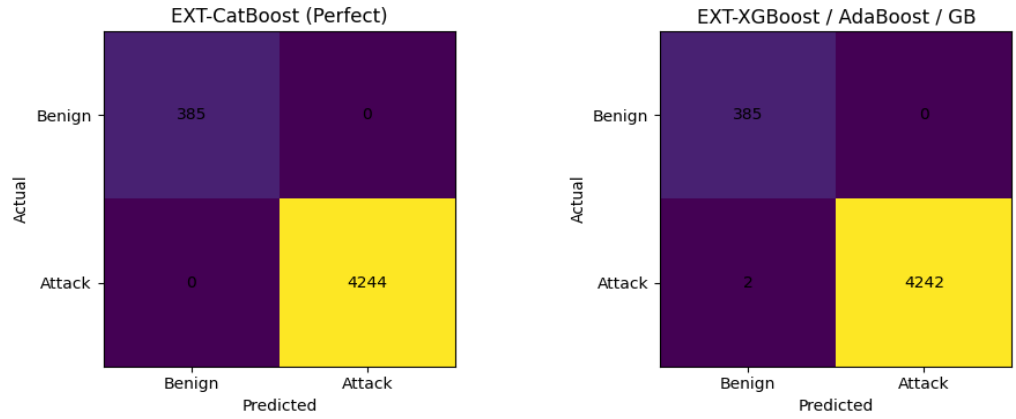


**Fig. 8.** Confusion Matrices for IoTNet24

In Fig. 9 machine learning models' performance is analyzed by plotting accuracy against the number of training examples. The graph reveals that Gradient Boosting maintains high training accuracy, suggesting effective learning, while its validation accuracy plateaus, indicating a limit to performance improvement with additional data. AdaBoost shows a notable discrepancy between training and validation accuracy, hinting at overfitting. XGBoost and CatBoost display high training accuracy with a slight decrease as more data is introduced, which may denote better generalization.

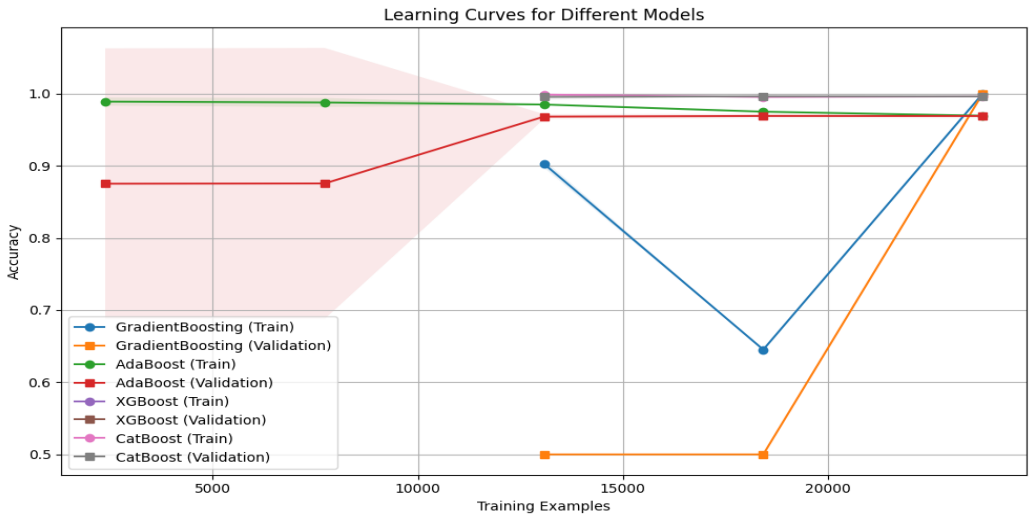


**Fig. 9.** Learning Curve of different models

### 4.3. Third Experiment

This experiment utilized the BoTNeTIoT-L01 dataset as shown in Table 6, a comprehensive dataset that integrates data from various IoT devices originally found in the BoTNeTIoT dataset. This enhanced version minimizes redundancy by selecting features based on a 10-second time window. This dataset, created by the Cyber Range Lab of UNSW Canberra, simulates realistic IoT network traffic, including both normal and various types of malicious traffic such as botnet attacks.

**Table 6.** BoTNeTIoT-L01 dataset summary

| Aspect | Description |
|---|---|
| Dataset Name | BoTNeTIoT-L01 |
| Dataset Origin | UNSW Canberra Cyber Range Lab |
| Dataset Type | Network Traffic Data for Intrusion Detection (IOT) |
| Total Instances | Approximately 7.6 million instances |
| Total Features | 27 features |
| Feature Types | Numerical, Categorical, Time-based, and Boolean |

The dataset analysis results are shown in Fig. 10, where a count plot offers a more in-depth look of the frequency distribution of the two classes and a pie chart shows the relative proportions of the two classes.

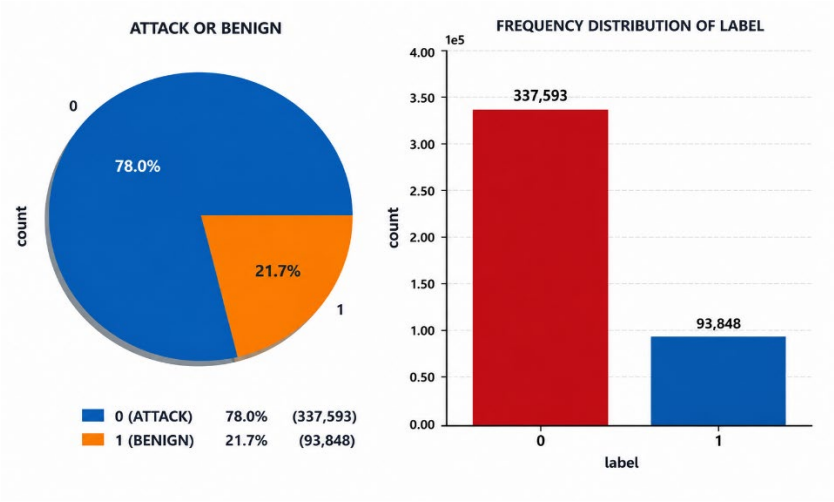


**Fig. 10.** BoTNeTIoT-L01 dataset analysis

We employed Extra Trees (ExtraTreesClassifier) to identify key features on the training split only. Non-predictive columns were removed, and the feature selector was trained exclusively on the training samples. Fig. 11 presents the top 10 most important features identified by the classifier based on Gini importance. The same 20-feature subset was then applied to the held-out test set to ensure no information leakage.

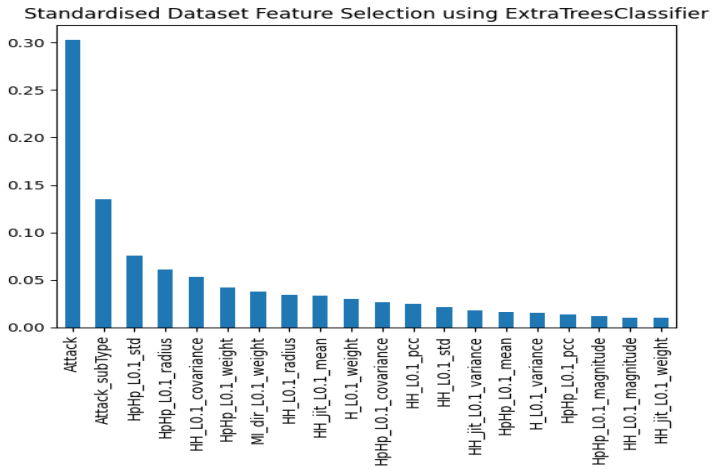


**Fig. 11.** Top 20 most significant features determined by the classifier

The BoTNeTIoT-L01 dataset is a large-scale IoT botnet traffic dataset containing 2,426,574 samples with 23 features (after removing the index column). The dataset exhibits moderate class imbalance with 78.84% benign traffic and 21.16% attack traffic—less severe than IoTNet24 but still requiring class balancing to prevent majority class bias.

Class Distribution and Balancing:

- Training set (BEFORE balancing): 1,530,461 benign (78.84%) and 410,798 attack (21.16%)
- Training set (AFTER balancing): 410,798 benign (50%) and 410,798 attack (50%)
- Test set (original imbalance): 382,616 benign (78.84%) and 102,699 attack (21.16%)

Random under sampling reduces the majority class (benign) to match the minority class (attack), creating a balanced 1:1 training set with 821,596 total samples. The test set retains its original imbalanced

distribution to evaluate real-world generalization. We then apply Extra Trees feature selection on the balanced training set, reducing from 23 to 11 features (52% reduction).

Table 7 summarizes the classification performance of four EXT-based ensemble models, all of which demonstrate near-perfect results across all evaluation metrics. EXT-XGBoost achieves the strongest overall balance, with the highest precision and a perfect recall, leading to the top MCC value, which indicates highly reliable predictions even under class imbalance. EXT-GB and EXT-CatBoost follow closely, showing only marginal reductions in precision and recall, with extremely low false positive and false negative rates that are practically negligible. EXT-AdaBoost also performs exceptionally well, though it exhibits the lowest precision and MCC among the four models, reflecting a very slight increase in classification errors. Overall, the results confirm excellent class separability and robustness for all models, with EXT-XGBoost emerging as the most stable and accurate performer in this comparison

**Table 7.** BoTNeTIoT-L01 Results (Binary Classification: Attack vs Benign)

| Model | Accuracy | Precision | Recall | F1-score | ROC-AUC | PR-AUC | FPR | FNR | MCC |
|---|---|---|---|---|---|---|---|---|---|
| EXT-XGBoost | 0.9999 | 0.9998 | 1.0000 | 0.9999 | 1.0000 | 1.0000 | 0.0000 | 0.0000 | 0.9999 |
| EXT-GB | 0.9999 | 0.9998 | 0.9999 | 0.9999 | 1.0000 | 1.0000 | 0.0001 | 0.0000 | 0.9999 |
| EXT-CatBoost | 0.9999 | 0.9996 | 0.9999 | 0.9998 | 1.0000 | 1.0000 | 0.0001 | 0.0000 | 0.9997 |
| EXT-AdaBoost | 0.9999 | 0.9995 | 0.9999 | 0.9997 | 1.0000 | 1.0000 | 0.0001 | 0.0001 | 0.9996 |

The confusion matrices in Fig. 12 show remarkable performance: EXT-CatBoost achieves perfect 100% classification with zero false positives and zero false negatives. The other three models miss only 2 attacks out of 4,244 (0.05% FNR) while maintaining zero false positives. This near-perfect separation persists on the original imbalanced test set (91.68% attacks), demonstrating that class balancing improved model generalization rather than merely fitting the balanced training distribution.

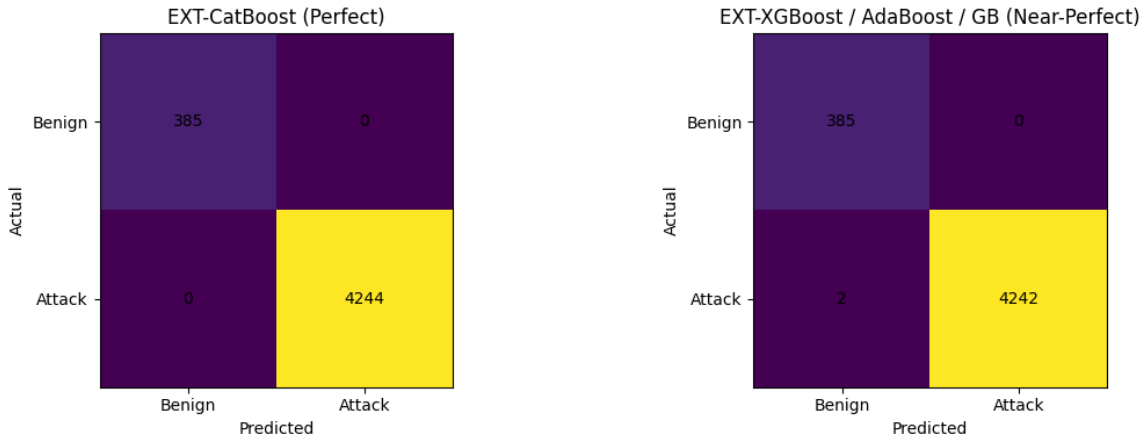


**Fig. 12.** Confusion Matrices for IoTNet24

## 4.4. Discussion

Our evaluation across three IoT intrusion detection datasets reveals critical insights about hybrid feature selection, ensemble classification, and the role of dataset environment in IDS research.

### 4.4.1. Performance Summary and Lab-vs-Reality Gap

On real-world UNSW-NB15, EXT-GB achieves 87.73% accuracy with 98.58% recall (0.98 ROC-AUC, 25.55% FPR, 1.42% FNR), missing only 644 of 45,332 attacks. After class balancing, laboratory

datasets achieve near-perfect performance: IoTNet24 (99.96%, 2 errors in 4,629 samples) and BoTNeTIoT (99.99%, 21 errors in 485,315 samples).

The 12-percentage-point gap between real-world (87.73%) and laboratory datasets (99.96-99.99%) reveals that controlled environments significantly overestimate real-world IDS performance, regardless of scale. Both small (23K) and large (2.4M) laboratory datasets achieve near-perfect results, proving environment (controlled vs adversarial) is the primary determinant, not dataset size**.

Laboratory perfection stems from: (1) predictable attack patterns without evasion (proto accounts for 39.6% importance in IoTNet24), (2) lack of adversarial behavior, (3) homogeneous device populations, (4) temporal consistency without concept drift, and (5) limited attack diversity. UNSW-NB15's moderate 87.73% accuracy reflects genuine deployment challenges, traffic diversity, adversarial evasion, and overlapping attack-benign patterns. Near-perfect laboratory results validate methodology soundness when data is separable; UNSW-NB15's moderate performance reflects realistic operational capability where maintaining 98% recall with 25% FPR represents practical trade-offs

#### 4.4.2. FPR-FNR Trade-off and Feature Selection Impact

The moderate FPR (25.55%) with excellent FNR (1.42%) on UNSW-NB15 is operationally justified for security-critical environments. Missing attacks carry higher risk than investigating false alarms; modern IDS deployments use multi-stage filtering where high-recall models flag threats followed by secondary validation. The high ROC-AUC (0.98) enables confidence-based prioritization. Extra Trees feature selection reduces dimensionality by 29-59% (UNSW-NB15: 49→20, IoTNet24: 14→10, BoTNeTIoT: 23→11) while maintaining performance, providing: (1) computational efficiency (EXT-CatBoost: 0.0007ms/sample on UNSW-NB15), (2) overfitting prevention, and (3) interpretability. Selected features include connection states, TTL values, protocol information, timing characteristics, and flow statistics, security-relevant indicators distinguishing attacks from legitimate traffic.

Class balancing (IoTNet24: 11:1→1:1, BoTNeTIoT: 3.73:1→1:1) enables near-perfect generalization to original imbalanced test sets, demonstrating that balancing improved model learning rather than merely fitting the training distribution.

#### 4.4.3. Comparison with Prior Work and Computational Performance

Compared to recent UNSW-NB15 studies reporting 99.98% accuracy [21] without confusion matrices or FPR/FNR rates, our contribution provides: (1) explicit train-only feature selection preventing leakage, (2) complete metrics (ROC-AUC: 0.9837, PR-AUC: 0.9879, MCC: 0.7655), (3) confusion matrices, and (4) cross-environment validation revealing the 12-point performance gap.

All models meet <1ms inference requirements for real-time deployment (1000+ flows/second). Training times range from <1s (IoTNet24) to 27 minutes (BoTNeTIoT EXT-GB), while inference remains efficient (0.004-2.86s for 4.6K-485K test samples). EXT-CatBoost provides the fastest inference; EXT-XGBoost balances training speed and accuracy.

#### 4.4.4. Limitations

1) Offline evaluation only, no embedded hardware validation; (2) 25.55% FPR requires SOC support for alert management; (3) datasets from 2015-2020 may miss recent attack vectors; (4) binary classification only, multi-class introduces severe imbalance; (5) static feature selection, online adaptation could improve long-term performance; (6) no adversarial robustness testing; (7) laboratory results (99.96-99.99%) should not represent real-world expectations, papers reporting only laboratory performance may overestimate deployment readiness

## 5. Conclusion

This paper presented a rigorous evaluation of hybrid intrusion detection systems combining Extra Trees feature selection with four ensemble classifiers (XGBoost, CatBoost, AdaBoost, Gradient Boosting) for IoT network security. Our methodology ensures train-only feature selection, preventing data leakage, class balancing for imbalanced datasets, and comprehensive metric reporting (ROC-AUC,

PR-AUC, FPR, FNR, MCC). On real-world UNSW-NB15, EXT-GB achieves 87.73% accuracy (0.98 ROC-AUC, 98.58% recall, 1.42% FNR), realistic performance balancing high attack detection with moderate false positives (25.55%). After class balancing, laboratory datasets achieve near-perfect performance: IoTNet24 (99.96%) and BoTNeTIoT (99.99%). The 12-percentage-point gap between real-world and laboratory performance is consistent across small (23K) and large (2.4M) laboratory datasets, revealing that environment (controlled vs adversarial) determines IDS performance more than dataset scale. This finding contributes to honest IDS evaluation practices. Feature selection reduces dimensionality by 29-59% while maintaining accuracy, enabling <1ms per-sample inference suitable for real-time deployment (1000+ flows/second). Training times range from <1s to 27 minutes depending on dataset size. Offline evaluation only (no embedded hardware testing); datasets from 2015-2020 may miss recent attack vectors; 25% FPR requires SOC support; binary classification only; no adversarial robustness testing.

For future work, the proposed framework will be extended toward practical and robust deployment scenarios by validating embedded implementations on resource-constrained platforms such as Raspberry Pi and NVIDIA Jetson, with a particular focus on real-time inference latency and power consumption analysis. Lightweight deep learning models, including 1D-CNN and LSTM architectures, will be integrated using knowledge distillation to preserve predictive performance while reducing computational overhead. To address non-stationary environments, adaptive online learning mechanisms will be investigated to handle concept drift in evolving traffic patterns. In addition, adversarial robustness will be systematically evaluated to assess resilience against evasion and poisoning attacks. The framework will be further enhanced to support multi-class attack categorization using cost-sensitive learning in order to mitigate class imbalance and unequal misclassification costs. Comprehensive validation will be conducted on recent datasets from 2023–2024 that incorporate modern attack vectors, and cross-dataset transfer learning will be explored to improve generalization and reduce the gap between laboratory evaluation and real-world production deployment.

## Declarations

**Author contribution.** All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
**Funding statement.** None of the authors have received any funding or grants from any institution or funding body for the research.
**Conflict of interest.** The authors declare no conflict of interest.
**Additional information.** No additional information is available for this paper.

## References

[1] S. Li, L. Da Xu, and S. Zhao, "The internet of things: a survey," *Inf. Syst. Front.*, vol. 17, no. 2, pp. 243–259, Apr. 2015, doi: 10.1007/s10796-014-9492-7.

[2] N. Nazar Kamal, Q. F. Al-Doori, and O. Alani, "A Review of Modern and Recent Studies on the Low-Density Parity-Check Technology Approach," *Iraqi J. Comput. Commun. Control Syst. Eng.*, vol. 23, no. 1, pp. 165–178, Mar. 2023, doi: 10.33103/uot.ijccce.23.1.13.

[3] Y. M. Hussain *et al.*, "Smartphone's off grid communication network by using Arduino microcontroller and microstrip antenna," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 19, no. 4, p. 1100, Aug. 2021, doi: 10.12928/telkomnika.v19i4.15949.

[4] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, "A survey of intrusion detection in Internet of Things," *J. Netw. Comput. Appl.*, vol. 84, pp. 25–37, Apr. 2017, doi: 10.1016/j.jnca.2017.02.009.

[5] D. Barwal *et al.*, "The impact of netflix recommendation engine on customer experience," 2023, p. 060005, doi: 10.1063/5.0170916.

[6] M. Y. Zeain *et al.*, "A New Technique of FSS-Based Novel Chair-Shaped Compact MIMO Antenna to Enhance the Gain for Sub-6GHz 5G Applications," *IEEE Access*, vol. 12, pp. 49489–49507, 2024, doi: 10.1109/ACCESS.2024.3380013.

[7] A. Choudhary, "Internet of Things: a comprehensive overview, architectures, applications, simulation tools, challenges and future directions," *Discov. Internet Things*, vol. 4, no. 1, p. 31, Dec. 2024, doi: 10.1007/s43926-024-00084-3.

[8] A. A. Abdulbari *et al.*, "Single-Layer Planar Monopole Antenna-Based Artificial Magnetic Conductor (AMC)," *Int. J. Antennas Propag.*, vol. 2022, pp. 1–9, Jul. 2022, doi: 10.1155/2022/6724175.

[9] L. Yan, Y. Diao, Z. Lang, and K. Gao, "Corrosion rate prediction and influencing factors evaluation of low-alloy steels in marine atmosphere using machine learning approach," *Sci. Technol. Adv. Mater.*, vol. 21, no. 1, pp. 359–370, Jan. 2020, doi: 10.1080/14686996.2020.1746196.

[10] Y. K. Saheed, "Performance Improvement of Intrusion Detection System for Detecting Attacks on Internet of Things and Edge of Things," 2022, pp. 321–339, doi: 10.1007/978-3-030-80821-1_15.

[11] A. Adiwijaya and N. G. Ramadhan, "Analyzing risk factors and handling imbalanced data for predicting stroke risk using machine learning," *Int. J. Adv. Intell. Informatics*, vol. 11, no. 1, p. 39, Feb. 2025, doi: 10.26555/ijain.v11i1.1678.

[12] A. S. Issa, Y. H. Ali, and T. A. Rashid, "BCDDO: Binary Child Drawing Development Optimization," *J. Supercomput.*, vol. 80, no. 11, pp. 16202–16221, Jul. 2024, doi: 10.1007/s11227-024-06088-8.

[13] A. F. Al-zubidi, A. K. Farhan, and E.-S. M. El-Kenawy, "Surveying Machine Learning in Cyberattack Datasets: A Comprehensive Analysis," *J. Soft Comput. Comput. Appl.*, vol. 1, no. 1, Jun. 2024, doi: 10.70403/3008-1084.1000.

[14] A. S. Issa, Y. H. Ali, and T. A. Rashid, "Review on Hybrid Swarm Algorithms for Feature Selection," *Iraqi J. Sci.*, pp. 5331–5344, Oct. 2023, doi: 10.24996/ijs.2023.64.10.38.

[15] A. A. Mosslah, R. H. Mahdi, and S. M. Al-Barzinji, "Performance Evaluation Of The Deep Learning System For Weed Recognization." pp. 1–10, Dec. 28, 2023, doi: 10.21203/rs.3.rs-3791687/v1.

[16] H.-Y. Kwon, T. Kim, and M.-K. Lee, "Advanced Intrusion Detection Combining Signature-Based and Behavior-Based Detection Methods," *Electronics*, vol. 11, no. 6, p. 867, Mar. 2022, doi: 10.3390/electronics11060867.

[17] M. Al-Kubaisi, A. S. Ahmed, S. M. Al-Barzinji, and A. M. Khaleel, "Advanced Estimation of Brain Age Using Pre-trained 2D Convolutional Neural Networks on a Public Dataset," *J. Robot. Control*, vol. 5, no. 4, pp. 981–991, May 2024. [Online]. Available at: https://journal.umy.ac.id/index.php/jrc/article/view/22006

[18] Y. Xu, Y. Zhou, P. Sekula, and L. Ding, "Machine learning in construction: From shallow to deep learning," *Dev. Built Environ.*, vol. 6, p. 100045, May 2021, doi: 10.1016/j.dibe.2021.100045.

[19] R. Setiawan and T. K. A. Rahman, "Machine learning-based B2C software project success prediction model in Indonesia," *Int. J. Adv. Intell. Informatics*, vol. 11, no. 3, p. 480, Aug. 2025, doi: 10.26555/ijain.v11i3.2123.

[20] S. P. R.M. *et al.*, "An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture," *Comput. Commun.*, vol. 160, pp. 139–149, Jul. 2020, doi: 10.1016/j.comcom.2020.05.048.

[21] A. S. Issa, Y. H. Ali, and T. A. Rashid, "Enhanced Harris Hawks Optimization (EHHO) for Detecting COVID-19 Using Chest X-Ray," in *2022 2nd International Conference on Advances in Engineering Science and Technology (AEST)*, Oct. 2022, pp. 361–365, doi: 10.1109/AEST55805.2022.10413107.

[22] S. F. Pane, M. D. Sulistiyo, A. A. Gozali, and A. Adiwijaya, "PSO-Enhanced ensemble techniques for pandemic prediction and feature importance analysis," *Int. J. Adv. Intell. Informatics*, vol. 11, no. 4, p. 653, Nov. 2025, doi: 10.26555/ijain.v11i4.2091.

[23] Y. Miao, C. Chen, L. Pan, Q.-L. Han, J. Zhang, and Y. Xiang, "Machine Learning–based Cyber Attacks Targeting on Controlled Information," *ACM Comput. Surv.*, vol. 54, no. 7, pp. 1–36, Sep. 2022, doi: 10.1145/3465171.

[24] H. Qiu, T. Dong, T. Zhang, J. Lu, G. Memmi, and M. Qiu, "Adversarial Attacks Against Network Intrusion Detection in IoT Systems," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10327–10335, Jul. 2021, doi: 10.1109/JIOT.2020.3048038.

[25] T. Miller *et al.*, "Integrating Artificial Intelligence Agents with the Internet of Things for Enhanced Environmental Monitoring: Applications in Water Quality and Climate Data," *Electronics*, vol. 14, no. 4, p. 696, Feb. 2025, doi: 10.3390/electronics14040696.

[26] J. D. Gadze, A. A. Bamfo-Asante, J. O. Agyemang, H. Nunoo-Mensah, and K. A.-B. Opare, "An Investigation into the Application of Deep Learning in the Detection and Mitigation of DDOS Attack on SDN Controllers," *Technologies*, vol. 9, no. 1, p. 14, Feb. 2021, doi: 10.3390/technologies9010014.

[27] R. Mishra and A. Mishra, "Current research on Internet of Things (IoT) security protocols: A survey," *Comput. Secur.*, vol. 151, p. 104310, Apr. 2025, doi: 10.1016/j.cose.2024.104310.

[28] Y.-W. Chen, J.-P. Sheu, Y.-C. Kuo, and N. Van Cuong, "Design and Implementation of IoT DDoS Attacks Detection System based on Machine Learning," in *2020 European Conference on Networks and Communications (EuCNC)*, Jun. 2020, pp. 122–127, doi: 10.1109/EuCNC48522.2020.9200909.

[29] N. Islam *et al.*, "Towards Machine Learning Based Intrusion Detection in IoT Networks," *Comput. Mater. Contin.*, vol. 69, no. 2, pp. 1801–1821, 2021, doi: 10.32604/cmc.2021.018466.

[30] N. Imtiaz *et al.*, "A Deep Learning-Based Approach for the Detection of Various Internet of Things Intrusion Attacks Through Optical Networks," *Photonics*, vol. 12, no. 1, p. 35, Jan. 2025, doi: 10.3390/photonics12010035.

[31] Y. Kayode Saheed, A. Idris Abiodun, S. Misra, M. Kristiansen Holone, and R. Colomo-Palacios, "A machine learning-based intrusion detection for detecting internet of things network attacks," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 9395–9409, Dec. 2022, doi: 10.1016/j.aej.2022.02.063.

[32] T. S. Othman and S. M. Abdullah, "An Intelligent Intrusion Detection System for Internet of Things Attack Detection and Identification Using Machine Learning," *ARO-THE Sci. J. KOYA Univ.*, vol. 11, no. 1, pp. 126–137, May 2023, doi: 10.14500/aro.11124.

[33] R. Iranzad and X. Liu, "A review of random forest-based feature selection methods for data science education and applications," *Int. J. Data Sci. Anal.*, pp. 1–15, Feb. 2024, doi: 10.1007/s41060-024-00509-w.

[34] M. Altalhan, A. Algarni, and M. Turki-Hadj Alouane, "Imbalanced Data Problem in Machine Learning: A Review," *IEEE Access*, vol. 13, pp. 13686–13699, 2025, doi: 10.1109/ACCESS.2025.3531662.

[35] J. Li *et al.*, "Application of XGBoost algorithm in the optimization of pollutant concentration," *Atmos. Res.*, vol. 276, p. 106238, Oct. 2022, doi: 10.1016/j.atmosres.2022.106238.

[36] Z. Fan, J. Gou, and S. Weng, "Complementary CatBoost based on residual error for student performance prediction," *Pattern Recognit.*, vol. 161, p. 111265, May 2025, doi: 10.1016/j.patcog.2024.111265.

[37] C. K. K. Reddy *et al.*, "Twined ensemble framework for network security: integrating Random Forest, AdaBoost, and Gradient Boosting for enhanced intrusion detection," *Discov. Internet Things*, vol. 5, no. 1, p. 107, Oct. 2025, doi: 10.1007/s43926-025-00199-1.

[38] A. M. Elshewey, E. Selem, and A. H. Abed, "Improved CKD classification based on explainable artificial intelligence with extra trees and BBFS," *Sci. Rep.*, vol. 15, no. 1, p. 17861, May 2025, doi: 10.1038/s41598-025-02355-7.

[39] "The UNSW-NB15 Dataset," *UNSW Research*. [Online]. Available at: https://research.unsw.edu.au/projects/unsw-nb15-dataset.

[40] WittigenZ, "IoTNet24 Dataset for IDS," 2024. [Online]. Available at: https://www.kaggle.com/datasets/wittigenz/hydras.

[41] M. Hamsa Ahmed and M. Shokhan Al-Barzinji, "Integrated Computer Vision and Adaptive Machine Learning for Real-Time Fall Detection and Personalized Elderly Care," *Sci. Res. J. Eng. Comput. Sci.*, vol. 6, no. 1, pp. 1–5, Jan. 2026, doi: 10.47310/srjecs.2026.v06i01.003.