scientific reports

Check for updates

OPEN

Lightweight machine learning framework for efficient DDoS attack detection in IoT networks

Mamoona Nawaz[1], Shireen Tahira[1], Dilawar Shah[2], Shujaat Ali[2] & Muhammad Tahir[3]✉

The rapid proliferation of Internet of Things (IoT) devices has introduced significant security challenges, with Distributed Denial of Service (DDoS) attacks posing a critical threat to network integrity. Traditional detection methods often rely on computationally intensive models, rendering them unsuitable for resource-constrained IoT environments. To address this limitation, this study proposes a lightweight and scalable machine learning-based DDoS detection framework specifically designed for IoT networks. Utilizing the NSL-KDD dataset, the framework employs an Extra Trees Classifier (ETC) for feature selection, reducing dimensionality while retaining critical attributes. Reduced features were selected to enhance performance and reduce processing cost. Three supervised learning models, Random Forest, Logistic Regression, and Naïve Bayes, were implemented and evaluated based on their detection accuracy, precision, recall, and F1-score. Experimental results demonstrate that the Random Forest model achieves exceptional accuracy (99.88%), precision (99.93%), recall (99.81%), and F1-score (99.87%), outperforming both Logistic Regression (91.61% accuracy) and Naïve Bayes (87.62% accuracy). Furthermore, the proposed framework significantly reduces computational overhead compared to deep learning-based approaches, making it highly suitable for IoT deployments. This research advances IoT security by providing a scalable, efficient, and accurate solution for detecting DDoS attacks, thereby bridging the gap between high-performance requirements and resource limitations in real-world IoT applications.

Keywords Machine learning, Artificial intelligence, Internet of things, DDoS attacks, Networking, Cybersecurity

Abbreviations

DDoS	Distributed denial of service
NB	Naive bayes
LR	Logistic regression
RF	Random forest
ML	Machine learning
LSTM	Long short term memory
IoT	Internet of things
IDS	Intrusion detection systems
GPU	Graphical processing unit
ICMP	Internet control message protocol
TP	True positives
TN	True negatives
FP	False positives
FN	False negatives
ETC	Extra trees classifier
PCA	Principal component analysis
AUC	Area under the curve
TPR	True positive rate
TNR	True negative rate
FPR	False positive rate

[1]Department of Computer Science, International Islamic University, Islamabad, Pakistan. [2]Department of Computer Science, Bacha Khan University, Charsadda, Pakistan. [3]Department of Computer Science, Kardan University, Kabul, Afghanistan. ✉email: m.tahir@kardan.edu.af

FNR	False negative rate
ROC	Receiver operating characteristic

The Internet of Things (IoT) has fundamentally transformed how modern technology operates by enabling seamless interaction between physical objects and digital systems. The term IoT encompasses a wide array of physical objects and systems connected via the Internet or other communication and information transfer methods, typically without human involvement[1]. A significant need has arisen for IoT devices in both the private and business sectors. This includes a variety of smart home devices, thermostats, refrigerators, and security systems, as well as industrial, medical, and transport systems[2]. The IoT transformation brings numerous benefits, including increased efficiency, enhanced automation, and improved ease of use in healthcare, agriculture, manufacturing, and urban development. However, with the rise of IoT devices, the risks associated with their widespread use have significantly increased. These devices are highly vulnerable to attacks due to limitations like insufficient processing capabilities, weak security protocols, and inadequate system configurations[3]. There are several cyber threats to which IoT devices are vulnerable, and among these, DDoS attacks are considered particularly dangerous. A DDoS attack occurs when multiple compromised applications overwhelm a target computer or network with excessive traffic, preventing legitimate users from accessing the site. Many IoT devices carry out critical functions, such as health monitoring, industrial control, and smart city systems, and even a minor attack on these devices can result in catastrophic consequences, including service failure, property damage, or even loss of life[4].

Existing DDoS detection methods are increasingly inadequate for IoT environments. The primary limitation of signature-based techniques, which depend on identifying previously known attacks, is that these approaches struggle when faced with new, more advanced attack models[5]. Hybrid ensemble learning approaches combine multiple classifiers to improve anomaly detection accuracy and robustness in resource-constrained industrial sensor networks and SCADA systems[6]. Furthermore, while anomaly-based methods excel at detecting unusual behaviors, they often suffer from higher FPR and require significant resources. Due to limitations in processing power, memory, and bandwidth in IoT devices, many conventional approaches to detecting DDoS attacks are too computationally intensive for these devices and therefore cannot be executed in real time. Therefore, there is a pressing need for lightweight, accurate, and real-time DDoS detection mechanisms tailored for IoT systems[7]. Integrating explainable artificial intelligence (XAI) techniques into botnet detection frameworks enhances model interpretability, trustworthiness, and early detection of emerging attack patterns in IoT environments[8]. The attention mechanism enhances intrusion detection models by enabling selective focus on critical features, improving detection accuracy, especially in imbalanced network traffic scenarios[9]. Scaled dot-product attention mechanisms enable interpretable feature weighting in privacy-preserving intrusion detection systems, addressing both transparency and data sensitivity challenges in CPS-IIoT networks[10]. This study addresses this gap by proposing an efficient and lightweight ML framework for DDoS detection in IoT networks. The proposed method employs an ML approach to filter out DDoS attacks while maintaining high accuracy, supported by adaptive facilities. In this introduction, the context, problem, objectives, relevance, and scope of the study are described to establish the research need and set expectations for advancing knowledge in the field.

DDoS attacks

The IoT has transformed how humans interact with technology by enhancing the integration and information exchange among devices. Predictions indicate that IoT devices will increase to 35 billion by 2025; these insights highlight the potential risks associated with these devices[11]. Many IoT devices are characterized by inadequate security measures, making them all highly susceptible to various types of cyber threats. One of the most significant risks identified is DDoS attacks. With the growing adoption of the IoT in various industries, the safety of these devices becomes crucial. IoT devices offer insights into the performance and safety of healthcare facilities, smart cities, and industrial control systems. They can lead to severe consequences, including financial losses, disruptions to critical services, and threats to human lives[12]. Combining numerous networked devices and inadequate security measures creates an ideal environment for malicious actors. These devices often lack robust security mechanisms, making them vulnerable and putting essential infrastructure at risk of threats such as DDoS attacks. Figure 1. Illustration of a DDoS attack scenario. The diagram depicts a regular 'User' and multiple compromised 'Botnet' devices communicating through a central 'Network' to a target 'Server' labeled as 'DDoS Attack Target'. Arrows labeled 'Response' indicate the direction of network traffic flow. The botnet devices collectively send a large volume of requests via the network towards the target server, aiming to overwhelm its resources and cause a service disruption. This figure illustrates the mechanism by which numerous compromised devices, such as IoT devices integrated into a botnet, can be leveraged to execute a large-scale DDoS attack against a specific target.

In 2017, as expected, an IoT botnet known as IoT Reaper or IoTroop exploited vulnerabilities in IoT devices, thereby creating a functional botnet[13]. This malicious botnet primarily targeted household devices such as routers, IP cameras, and Network-Attached Storage systems. The IoT Reaper botnet represents a new level of both volume and complexity in DDoS attacks aimed at IoT devices[14]. This contrasts with previous types of botnets that relied on a single vulnerability in IoT devices, allowing IoT Reaper to create a vast network of compromised devices. These hackers demonstrated that IoT networks lack sufficient security measures, which should prompt more reliable safety protocols in the future. Regarding the purpose of the botnet, its function is still unclear; however, it has the potential to launch large-scale DDoS attacks, which can physically disrupt critical components. Fortunately, security experts have intervened and curtailed its spread by taking control of its server infrastructure. In addition to the healthcare sector, the control systems of smart grids managing the energy infrastructure of entire nations are also at risk of DDoS attacks. A successful strike could incapacitate a power plant's control systems, resulting in blackouts or even causing destruction[15]. Transportation, likewise, is

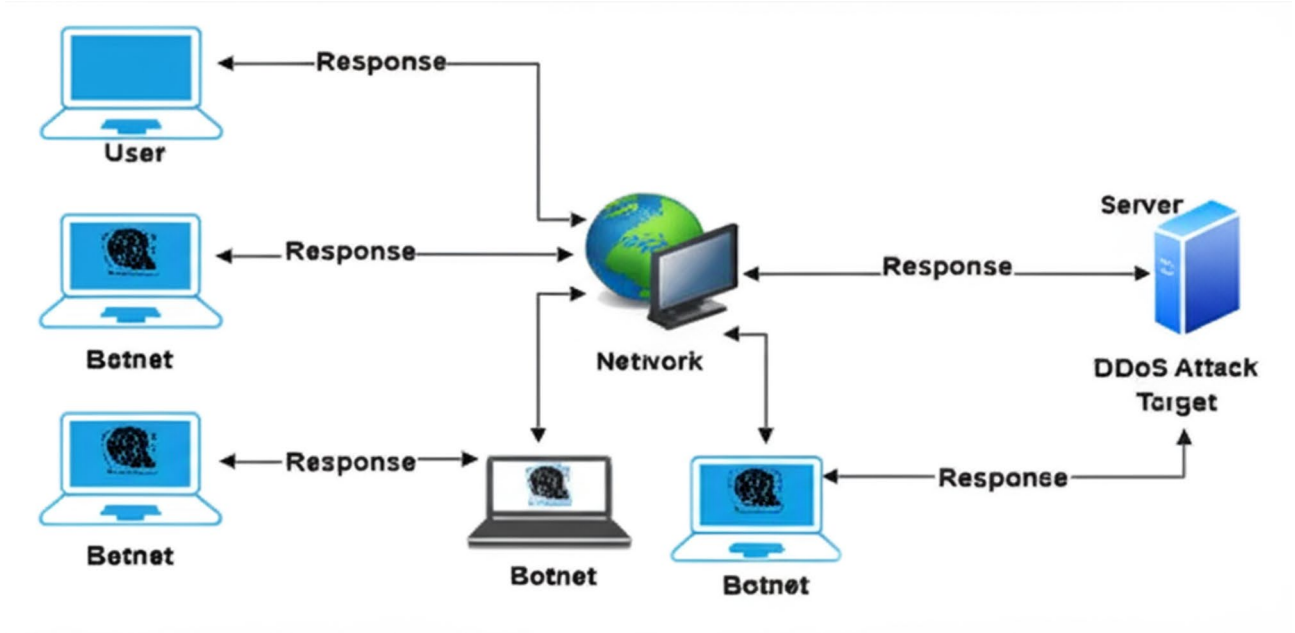


Fig. 1. DDoS attack flow in IoT Network.

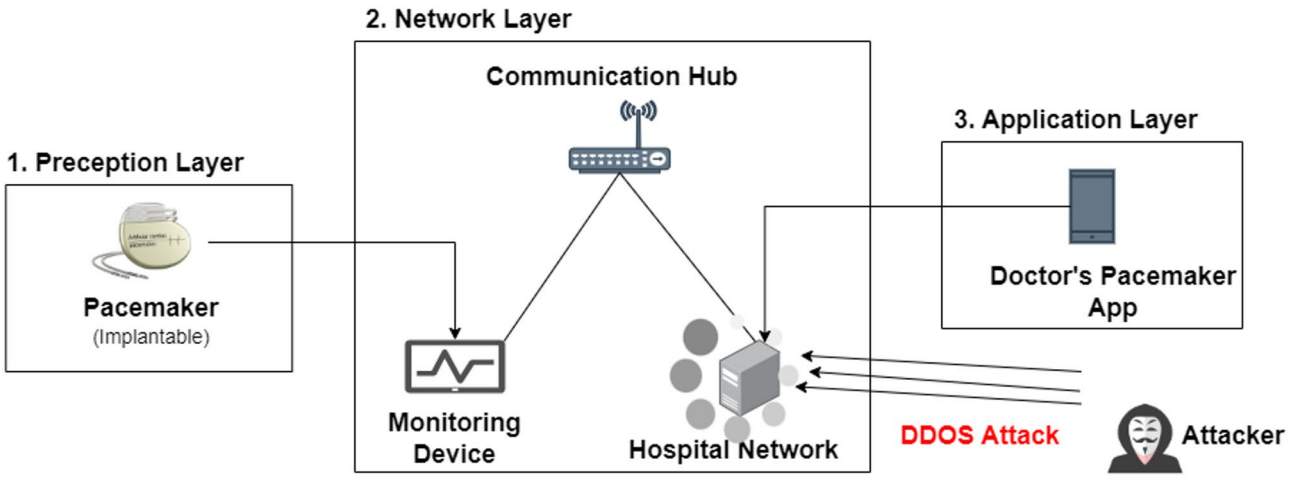


Fig. 2. DDoS Attack on Healthcare device.

not immune to such threats, as connected cars and traffic management systems can be compromised, leading to chaos in city amenities that can result in either an accident or a traffic jam. Figure 2 shows a Conceptual overview of a DDoS attack targeting a pacemaker communication system within a healthcare environment. The diagram illustrates the three primary layers involved: 1) the Perception Layer, containing the implantable pacemaker; 2) the Network Layer, comprising the Monitoring Device, Communication Hub, and Hospital Network, which facilitates data transfer; and 3) the Application Layer, where the Doctor's Pacemaker App resides. The figure highlights how an external 'Attacker' can launch a DDoS attack against the 'Hospital Network', thereby disrupting critical communication pathways between pacemakers and hospital infrastructure. Such an attack has significant implications for healthcare operations, as well as the integrity and availability of crucial patient information.

However, it can easily be demonstrated how a DDoS attack endangers IoT security in our everyday lives, particularly in healthcare, such as with pacemakers. These devices sync data with a hospital's tracking platform to monitor the patient's heartbeat. An attack on the hospital's network can inundate the servers with requests, disrupting effective communication between the pacemaker and the tracking system. This disruption can lead to a serious lack of necessary medical interventions, putting the patient's well-being at risk[16]. In the worst-case scenario, such an attack could leave a person in critical condition, necessitating immediate medical attention.

The key contributions of this paper are as follows: (1) We propose a lightweight DDoS detection framework utilizing RF, LR, and NB classifiers, specifically tailored for IoT environments. (2) An Extra Trees-based feature

selection technique is applied to enhance detection efficiency while minimizing computational load. (3) The proposed framework is validated using the NSL-KDD dataset, demonstrating superior performance compared to state-of-the-art models. (4) A comparative analysis is presented to highlight improvements in detection accuracy, precision, recall, and computational time.

Related work

The rising danger of cyberattacks on IoT systems has highlighted the limitations of intrusion detection systems, especially when challenged with advanced threats in complex environments [17]. To address these threats, the authors propose a new distributed framework based on deep learning attack detection techniques. This framework relies on deep learning models that primarily focus on feedforward neural networks and LSTMs to examine network traffic behaviors for detecting illegitimate performance. The proposed solution distributes detection workloads across various network nodes, resulting in improved scalability, more efficient operation, and greater network resilience. The research demonstrates how the framework accurately detects different types of attacks through performance evaluation with real-world testing. Overall, the research significantly enhances IoT system security by presenting an advanced strategy to detect cyber threats more effectively. Table 1 provides a summary of the related work.

The limitations of IDS in identifying various cyberattack types have prompted exploration into more advanced techniques. Deep learning shows significant promise for enhancing these systems, which is why a new model for multi-attack classification is proposed[18]. This model used convolutional neural networks (CNNs) and recurrent neural networks (RNNs) with long short-term memory (LSTM) units to analyze network traffic data for identifying multiple attack types. The model is trained on an extensive dataset, with evaluation based on performance metrics such as accuracy, precision, recall, and F1-score. The approach aims to improve both the operational effectiveness and predictive accuracy of IDS, thereby strengthening cyber defense mechanisms for digital networks.

Traditional techniques struggle to detect and prevent advanced DDoS attacks, prompting the exploration of deep-learning methods[19]. The authors propose utilizing deep learning approaches to enhance DDoS attack detection, given their ability to better handle complex attack patterns. This research develops a deep learning model through the implementation of neural networks, including recurrent neural networks (RNNs) with long short-term memory (LSTM) units and convolutional neural networks (CNNs), to effectively analyze network traffic data and distinguish between normal and DDoS attack traffic patterns. The model is trained on relevant data and evaluated using performance metrics such as accuracy, precision, recall, and F1-score. By applying deep learning techniques, this research aims to strengthen protection mechanisms against DDoS attacks, ultimately improving the resilience of network infrastructures.

References	Datasets	Objective	Methodology	Limitations
17	NSL KDD	To develop a robust and efficient deep learning-based distributed attack detection framework for IoT networks	Feedforward neural networks and recurrent neural networks (RNNs)	Deploying and managing a distributed system across an extensive, diverse IoT network can be complex
18	NSL KDD	To develop a deep learning-based model for classifying multiple cyber-attacks, aiming to improve the accuracy and effectiveness of intrusion detection systems	Long-Short-Term Memory Recurrent Neural Network (LSTM-RNN)	Complex Model High False Alarm Rate
19	NSL KDD	To investigate the use of a deep learning approach for detecting DDoS attacks, potentially improving the accuracy and effectiveness of DDoS detection methods	Deep Contractive Autoencoder (DCAE)	The availability and quality of the training data may limit the model's performance, potentially leading to inaccurate or biased detection results
20	NSL KDD	To enhance the speed of detection while upholding a commendable level of accuracy	SVM, logistic regression, KNN	The use of GPU technology results in decreased training and prediction time
21	NSL KDD	To categorize and predict various types of DDoSattacks through the application of machine learning	Random forest, XGBoost	Improved accuracy may be achieved using an enhanced suggested model
22	NSL KDD	To develop an optimized ensemble framework using big data analytics to effectively detect DDoS attacks targeting (IoT)	Convolutional Neural Network (CNN) embedded with a Gated Recurrent Unit (GRU)	The model's complexity results in high computational time
23	NSL KDD	To develop a precise and effective DDoS attack detection system for IoT networks with a hybrid Sample Selected RNN-ELM model	Recurrent Neural Networks (RNNs) and Extreme Learning Machines (ELMs)	The model's efficiency may hinge on the quality and variety of the training data and the particular attributes of the IoT network environment
24	NSL KDD	To provide a resilient and privacy-conscious DDoS attack detection solution for diverse IoT contexts by integrating federated learning with explainable a50rtificial intelligence approaches	Explainable Artificial Intelligence (XAI) with Federated Deep Neural Networks (FDNNs)	The efficiency of this methodology may be affected by issues including communication latency, variability in device capabilities, and the intricacy of incorporating XAI algorithms into the federated learning framework
25	NSL KDD	To develop a DDoS attack detection system using fine-tuned Multi-Layer Perceptrons	fine-tuned Multi-Layer Perceptron models	They are computationally intensive and slow
26	NSL KDD	To develop a system by integrating machine learning techniques with an SDN controller framework	Support Vector Machines, Decision Trees	The system may need to be continuously updated and adapted to address new and evolving DDoS attack techniques effectively
27	NSL KDD	To develop an intrusion detection system for IoT networks by integrating PCA for feature reduction with a CNN for accurate and efficient attack detection	Convolutional Neural Networks (CNN)	Computational overhead from CNN training could challenge resource-constrained IoT environments

Table 1. Summary of related work.

GPU-accelerated ML technology is explored in the context of improving botnet attack detection capabilities[20]. Detection methods often struggle with speed and efficiency, leading to the development of a new system that utilizes the parallel processing power of GPUs. The research investigates the performance of various machine learning algorithms, such as support vector machines (SVMs) and deep learning models, when implemented on GPU hardware. Effective feature engineering and thorough evaluations using real-world datasets are central to the methodology. By incorporating GPU acceleration, the proposed system enhances detection accuracy and speed, offering significant improvements in botnet activity identification.

Recognizing and addressing DDoS attacks continues to be a significant challenge, especially as conventional methods fall short against increasingly complex and evolving threats[21]. The study establishes a framework that creates models for detecting DDoS attacks by experimenting with various algorithms, including support vector machines (SVMs), XGBoost, Random Forests, and deep learning networks, utilizing network traffic data. These models are trained on suitable datasets and assessed using metrics like precision, recall, accuracy, and F1-score. Similarly, the goal of this research[22] is to harness machine learning to enhance DDoS protection systems, thereby increasing the operational resilience of networks against such attacks.

DDoS attack detection in the rapidly growing IoT networks presents a significant challenge due to the limitations of existing methods in resource-constrained environments[23]. It proposes a hybrid detection solution that combines recurrent neural networks (RNNs) and extreme learning machines (ELMs). RNNs are utilized for their ability to model temporal dependencies, while ELMs contribute fast training and strong generalization capabilities. The model is further optimized through a data point selection mechanism, which identifies the most informative data points for training. The resulting hybrid system offers a more accurate and efficient approach to detecting DDoS attacks, thus improving the security of IoT networks.

Explainable AI (XAI) techniques are integrated into a novel framework proposed[24]. Centralized methods have shown to be limited due to privacy concerns and resource constraints, leading the authors to integrate federated learning with explainable artificial intelligence (XAI). Federated learning enables collaborative model development across IoT devices while maintaining data privacy. Incorporating XAI improves the transparency of the model's decision-making processes, fostering both trust and understanding. This combined approach effectively addresses DDoS attack detection in diverse IoT environments, preserving privacy while providing explainable results for enhanced decision-making.

Fine-tuning Multi-Layer Perceptron (MLP) neural networks was explored in this paper[25] as a method to enhance their detection capabilities for Distributed Denial-of-Service (DDoS) attacks. The authors focus on adapting pre-trained MLP models for specific use in DDoS attack detection, addressing the shortcomings of traditional methods that often fail to identify such attacks accurately. Network traffic data is preprocessed before evaluating the MLP model's performance using accuracy, precision, and recall metrics. This research aims to improve DDoS detection systems by leveraging pre-trained models tailored to strengthen security technologies against these cyber vulnerabilities. Critical aspect of training these models is the loss function, and for binary classification problems like DDoS detection, the binary cross-entropy loss is commonly used. It is defined as:

$$L(y, \widehat{y} = -(y.log\,(\widehat{y}) + (1-y)\,.\log(1-\widehat{y}))$$

"ML-DDoSnet" is an intrusion detection system developed to combat Denial-of-Service (DDoS) attacks in IoT environments,[26]. IoT networks are vulnerable to various attacks, and the researchers apply machine learning technology to analyze network traffic and detect malicious patterns indicative of DDoS attacks. The system explores several machine learning algorithms, including support vector machines (SVMs) and decision trees, to classify network traffic as normal or malicious. The NSL-KDD dataset is used to test and validate the performance of ML-DDoSnet through training algorithms within the framework. The development and evaluation of ML-DDoSnet aim to enhance security mechanisms across IoT networks, improving their ability to defend against DDoS attacks.

Motivation

IoT networks' increasing complexity and scale necessitate effective IDS to guard against advancing threats. Conventional IDS frequently struggle with new attacks, which drives the adoption of ML to enhance IoT threat detection[28]. Lightweight ML models maintain an effective balance between accuracy and speed, which is essential for IoT devices with limited resources. ML-based solutions can adapt to evolving security threats, providing a more flexible defense than traditional methods. Research on ML-driven DDoS detection enhances the development of intelligent security solutions, strengthening the IoT environment against cyberattacks.

Dataset

NSL-KDD[29] is used for our research. This dataset has been derived from the KDD Cup 1999 dataset and has undergone specific modifications to address the impact of duplicate records on the outputs of IDS. The dataset includes 41 features spread across 125,973 training records and 22,544 testing records. These features are categorized into four groups: basic features, content features, time-based features, and host-based features, providing a comprehensive understanding of network behaviour. The NSL-KDD dataset is utilized as it resolves the redundancy issues present in the original KDD'99 dataset, offering a more balanced and dependable benchmark for intrusion detection. It helps evaluate ML models effectively by providing a range of attack patterns and realistic network traffic scenarios.

NSL-KDD was chosen due to its well-structured format, availability of labeled data, and widespread use as a benchmark in intrusion detection research. Its balanced distribution and reduced redundancy make it particularly suitable for training and evaluating ML models. Additionally, its comprehensive feature set enables robust analysis of both normal and attack traffic patterns, making it an ideal starting point for developing and

comparing classification models. However, we acknowledge that NSL-KDD is not inherently IoT-specific, and its traffic patterns may not fully reflect the heterogeneity or lightweight protocols typically found in modern IoT ecosystems. This introduces potential limitations in generalizing the model's effectiveness to real-world IoT scenarios.

Proposed technique

The proposed framework adopts a structured approach, as shown in Fig. 3. It starts with the NSL-KDD dataset. Preprocessing techniques address missing values, duplication, and normalization to guarantee data consistency. Feature selection is conducted using an ETC to pinpoint the most relevant attributes, improving detection efficiency while minimizing computational overhead. The refined dataset is subsequently used to train three ML models: Random Forest (RF), Logistic Regression (LR), and Naïve Bayes (NB), for DDoS attack classification. Model performance is assessed using accuracy, precision, recall, and F1-score. To support real-time detection in dynamic and resource-constrained IoT environments, the framework is designed with lightweight classifiers and reduced input dimensions. Feature selection significantly reduces computation time by eliminating redundant and low-importance attributes, enabling faster model inference with minimal memory usage.

Additionally, the selected models, particularly NB and LR, offer low training and prediction complexity, making them suitable for deployment on edge devices with limited processing power. The modular pipeline also ensures quick preprocessing and classification cycles, enhancing the framework's ability to respond to evolving attack patterns in real time. Model performance is assessed using accuracy, precision, recall, and F1-score. This comprehensive method enhances DDoS detection in IoT environments by combining effective preprocessing and lightweight classifiers.

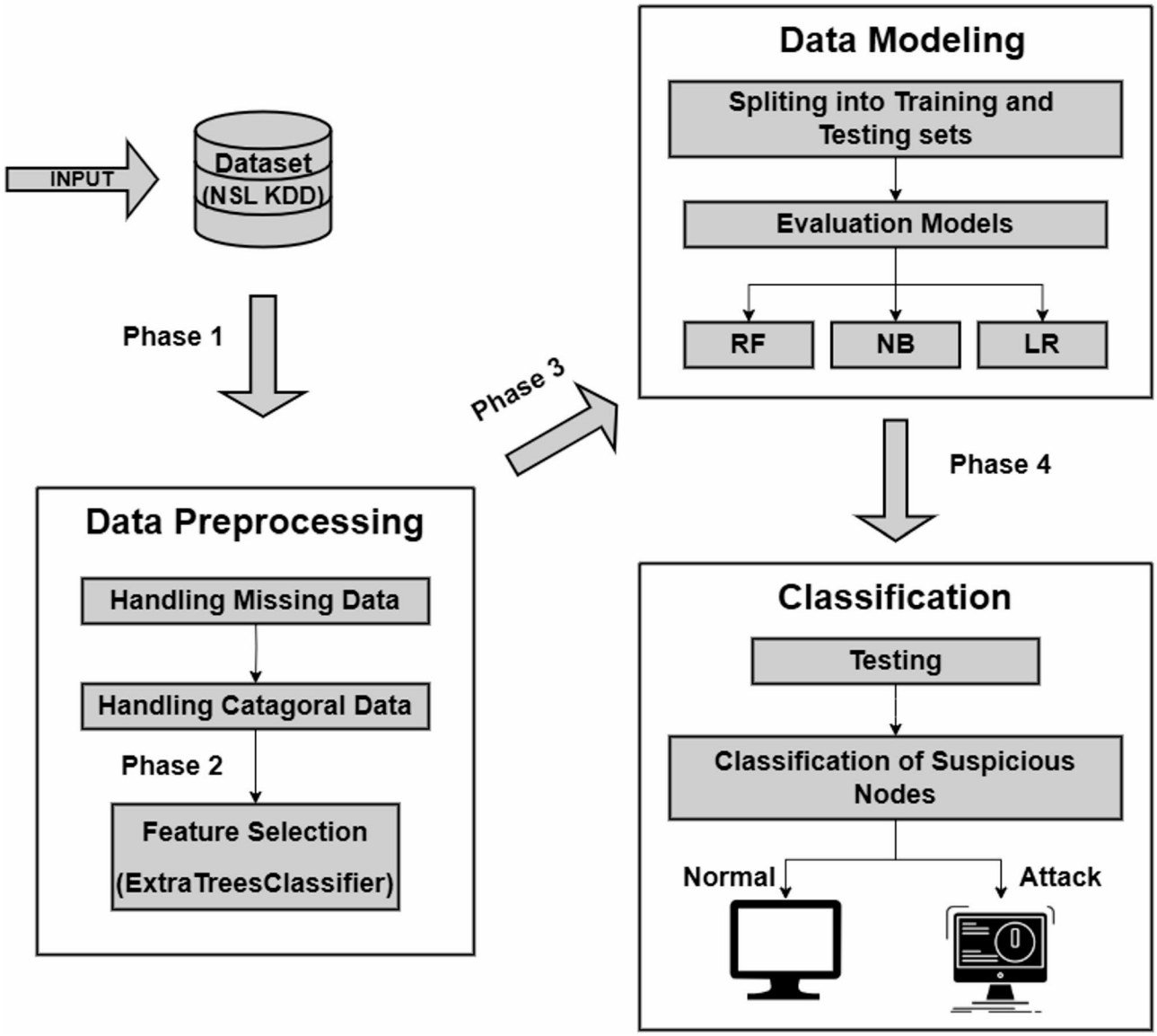


Fig. 3. Architectural Diagram of proposed technique.

Our methodology for DDoS attack detection is systematically structured into four distinct phases, as comprehensively depicted in Fig. 4. Phase 1, Data Preprocessing, focuses on preparing the raw IoT network dataset, including handling missing and categorical data. This is followed by Phase 2, Feature Selection, where the most impactful features are identified using techniques such as the ExtraTreesClassifier. In Phase 3, Data Modeling, the preprocessed data is split into training and testing sets, and various machine learning models, specifically Random Forest, Naive Bayes, and Logistic Regression, are trained and evaluated. The final Phase 4, Classification, involves deploying these trained models to test new data and classify network traffic as either normal or a DDoS attack.

Data preprocessing

The successful training of ML models require data preprocessing as a crucial first step to prepare the dataset. The data cleaning process is combined with data transformation to prepare the NSL-KDD raw data for analysis. Missing data preprocessing begins with a strategy that employs imputation or deletion techniques to manage incomplete records and prevent distortion of the results. Categorical attack types and protocols are converted into numeric datasets for compatibility with ML algorithms by applying one-hot encoding or label encoding functions. Certain conditions enable the use of normalization and scaling techniques to achieve standardized features, thereby preventing any single variable from dominating learning outcomes. The data preprocessing stages conclude by creating a standardized dataset prepared for model training, which includes variables formatted appropriately for both feature selection and classification processes.

The NSL-KDD dataset is known to exhibit class imbalance, particularly between normal and attack classes, which can bias classifiers toward majority classes and degrade detection performance for minority attack instances. To address this, we employed stratified sampling during the train-test split to maintain the original class distribution in both subsets. Additionally, we evaluated model performance using not only accuracy but also class-sensitive metrics such as precision, recall, and F1-score, which provide a more balanced view of model effectiveness across both majority and minority classes. Although no explicit resampling techniques like SMOTE or under sampling were applied, the strong recall and F1-scores, particularly for Random Forest, indicate the model's robustness even in the presence of class imbalance. Future work may explore synthetic resampling or ensemble-based methods for correcting imbalance to further improve detection rates for rare attack types.

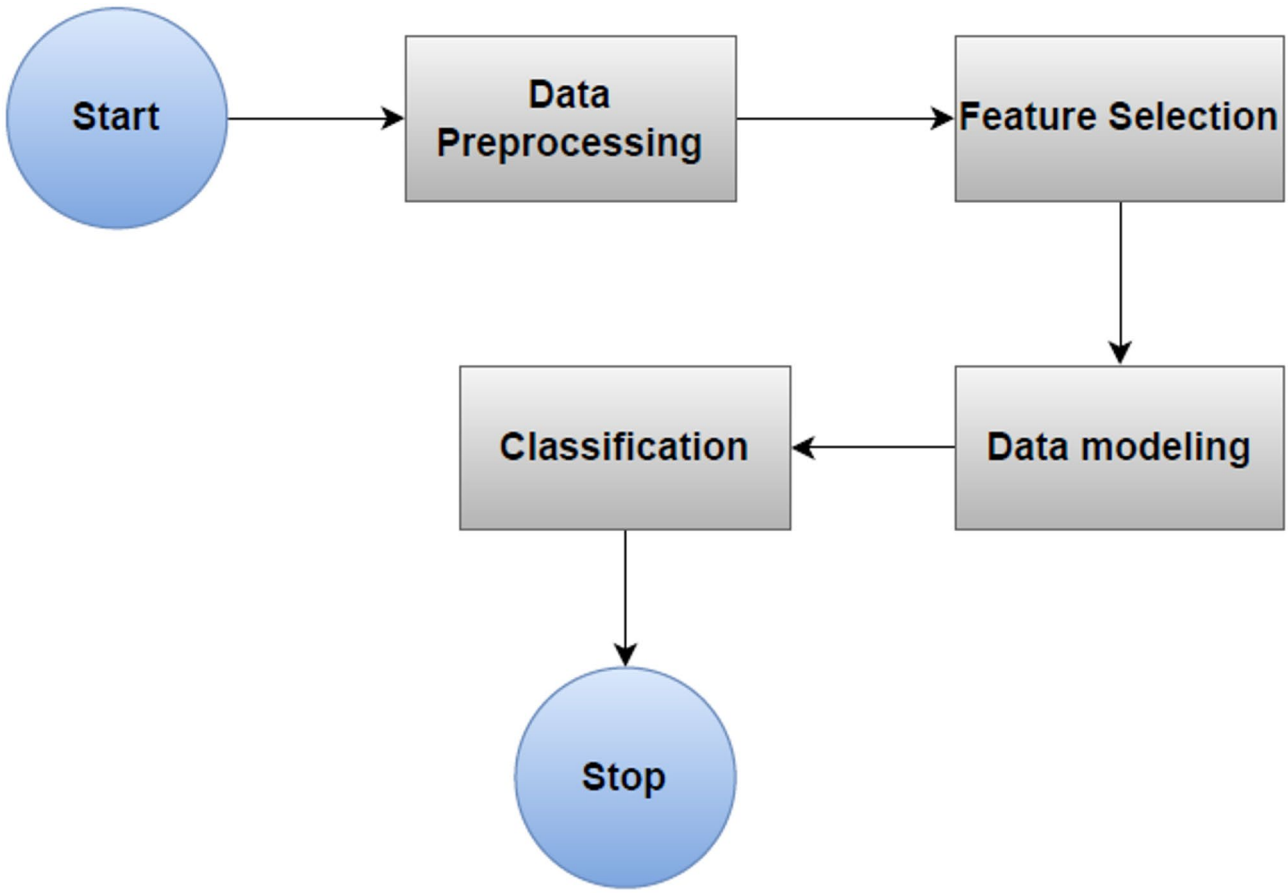


Fig. 4. Workflow for DDoS Attack Detection in IoT Networks.

Feature selection

By selecting the most relevant features from high-dimensional data, the ETC is an effective feature selection technique. Rather than reducing dimensionality through transformation, Extra Trees identifies the most informative features based on their importance scores, as shown in Fig. 5. In this study, 18 features were selected using the ExtraTreesClassifier from scikit-learn, which ranked the importance of all features and retained the top-performing ones for DDoS detection. The model was trained on the dataset, and feature importance scores were calculated using the feature_importances attribute. The least contributing features were removed to enhance

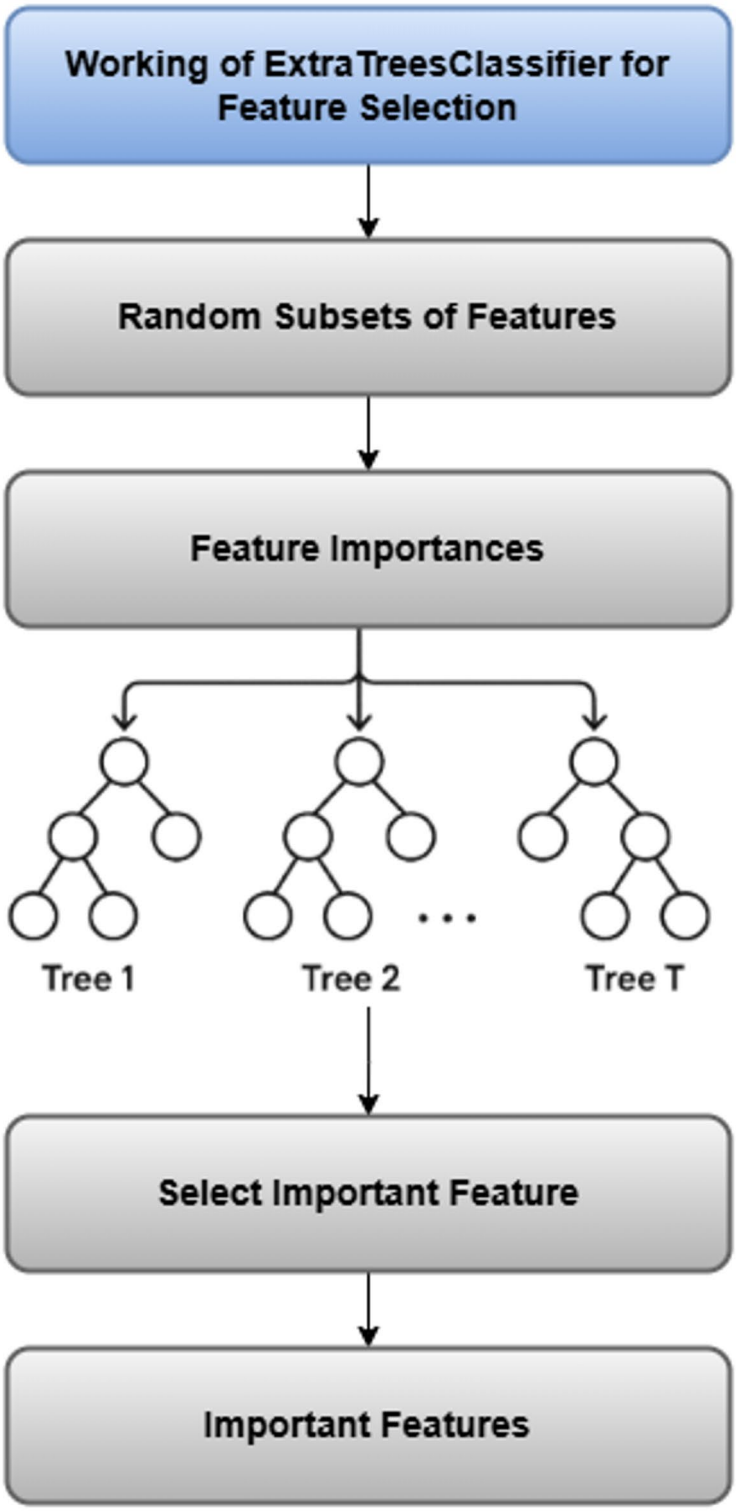


Fig. 5. Working of ExtraTreesClassifier for feature selection.

classification performance while minimizing computational overhead. By concentrating on the most relevant features, the Extra Trees-based selection improves detection efficiency for DDoS attacks in IoT networks while preserving essential data characteristics. This approach optimizes model performance, reduces processing costs, and advances IoT security by enabling more effective threat identification.

The ExtraTreesClassifier evaluates datasets by constructing multiple decision trees from various data subsets to evaluate feature importance based on successful node splits within each tree structure. The algorithm determines the amount of impurity reduction that occurs during the decision phase for each selected feature. A greater reduction in impurity indicates stronger feature importance, as these characteristics enhance the decision tree's efficiency in segregating analytical datasets. Figure 6 shows the most relevant features that the ExtraTreeClassifier selects. However, several limitations should be considered when interpreting these results. First, the use of the NSL-KDD dataset, while common in intrusion detection research, may introduce dataset bias due to its synthetic nature and class imbalance, potentially limiting the model's ability to generalize to more recent or real-world traffic. Second, the model's performance is contingent on assumptions made during preprocessing and feature selection, such as treating all retained features as equally reliable across contexts and discarding others that might hold value under different network conditions. Third, the reliance on importance scores calculated in a single training context assumes feature relevance remains static, which may not hold in dynamic IoT environments. As such, the generalizability of the selected features and the trained model to other datasets or deployment scenarios should be further validated through cross-dataset evaluation or real-time testing.

Figure 6 displays a horizontal bar chart showing the top 18 features and their corresponding importance scores. The most important feature is 'same_srv_rate', followed by 'dst_host_srv_serror_rate', and 'dst_host_serror_rate', indicating their significance in the classification model.

Model selection

We have selected three supervised learning classifiers, RF, LR, and NB, for DDoS attack detection in IoT environments. RF was selected for its robustness, scalability, and ability to handle high-dimensional data, which makes it well-suited for detecting complex attack patterns. LR was chosen for its probabilistic approach, which estimates the likelihood of an attack and provides interpretable decision boundaries. NB was included due to its efficiency in handling categorical data and its ability to classify attacks based on probabilistic assumptions of feature independence. These models were selected to strike a balance between accuracy, computational efficiency, and interpretability, thereby ensuring an effective and lightweight solution for IoT security.

To ensure robust and optimized performance, each classifier was configured with specific hyperparameters. For the RF model, we set the number of estimators to 100, used the default maximum depth, a minimum of two samples required to split an internal node, and one sample for leaf nodes, with a fixed random state of 42 to ensure reproducibility. LR was configured with an L2 regularization penalty, a regularization strength parameter $C = 1.0$, the 'lbfgs' solver, and a maximum iteration count of 1000 to ensure convergence. The NB model used the GaussianNB implementation with default settings, which is well-suited for normalized continuous data and performs efficiently with minimal parameter tuning. These configurations were selected based on prior

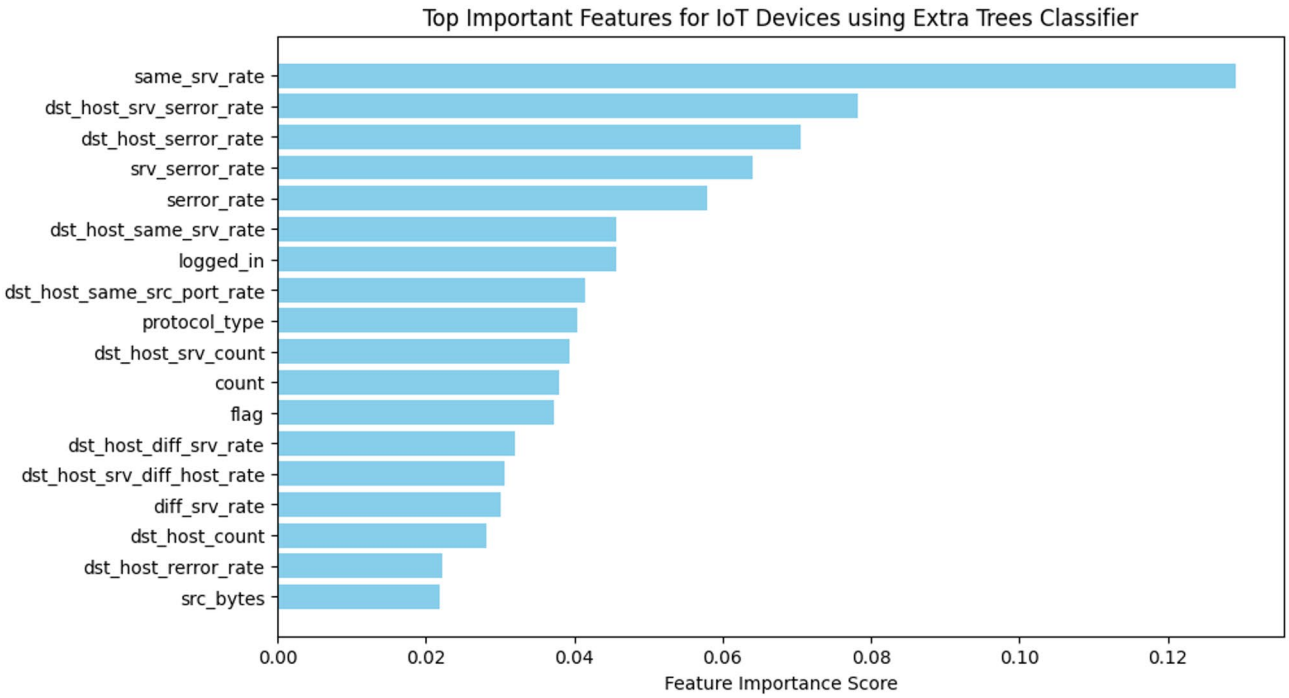


Fig. 6. Top 18 Selected features.

literature and empirical validation to strike a balance between model accuracy, efficiency, and suitability for IoT environments.

To ensure reproducibility and transparency, all experiments were conducted using Python 3.10 and the Scikit-learn library (version 1.2.2) in a Jupyter Notebook environment. The computational setup included a system with an Intel Core i5-1165G7 CPU @ 2.80 GHz, 16 GB RAM, and a Windows 11 operating system. No GPU acceleration was used, reflecting a resource-constrained environment typical of many IoT edge devices. Default parameters were used for the classifiers unless specified otherwise. The NSL-KDD dataset was preprocessed and split into training and test sets using an 80:20 ratio. All models were trained and evaluated using tenfold cross-validation to ensure consistent performance estimates.

Classification

Our proposed methodology, as comprehensively outlined in Fig. 7, systematically processes IoT network data for DDoS attack detection. A critical phase within this methodology involves the deployment of trained classification models to perform binary classification, distinguishing between normal network traffic and DDoS attacks. This deployment stage is preceded by robust data preprocessing, feature selection, and the rigorous training and evaluation of selected classifiers. The evaluation process identifies the most effective models for predicting class labels of unknown network traffic instances. This classification depends on features extracted and refined in earlier stages, where data cleaning, preprocessing, and feature selection improved relevance assessment. Using RF, NB, and LR, the testing dataset is classified to detect potential DDoS attack signals targeting IoT devices. Model predictions are validated against ground truth labels to differentiate between normal and malicious traffic.

Performance evaluation emphasizes accuracy, precision, recall, and F1-score, offering insights into detection effectiveness. Precision indicates the accuracy of attack identification, recall gauges the model's ability to identify attack instances, and the F1-score harmonizes both metrics. Moreover, FP (normal traffic incorrectly classified as attacks) and FN (attacks wrongly classified as normal) are examined to evaluate model reliability. This phase guarantees that the chosen models provide a fast, accurate, and efficient solution for real-world IoT environments that require robust DDoS attack detection. Pseudocode 1 presents the workflow of the proposed intrusion detection system using the NSL-KDD dataset. If the input data is raw, it undergoes preprocessing, including encoding categorical features and normalizing numerical ones. The data is then split into training and test sets. An ExtraTreesClassifier identifies the most important features, which are retained for model training. Three classifiers, RF, NB, and LR, are trained and evaluated. For each, if prediction is successful, performance metrics such as accuracy, precision, recall, and F1-score are calculated; otherwise, an error is logged.

Evaluation metrics

A set of evaluation metrics measures both the performance and effectiveness of the proposed method throughout this study. These metrics, commonly used in ML research, enable a quantitative assessment of classification model success. The formulas for these performance assessment metrics are derived from standard methodologies. Our experimental results demonstrate an improvement over the baseline, highlighting the effectiveness of our proposed technique for detecting DDoS attacks on IoT devices.

To ensure robustness and generalization of the results, we employed tenfold cross-validation during model training and evaluation. This approach partitions the dataset into 10 subsets, iteratively training the model on 9 folds while validating on the remaining one. This process minimizes bias due to data partitioning and ensures a more reliable estimate of model performance. We report the average value and standard deviation for each metric across the 10 folds, providing insight into the consistency of the model's performance.

In the equations presented below, various parameters are defined: TP, TN, FP, FN, L, and M. Specifically, TP stands for true positives (correctly predicted normal class), TN stands for true negatives (correctly predicted attack class), FP signifies false positives (incorrectly predicted normal class), and FN indicates false negatives (incorrectly predicted attack class). L and M are the actual and predicted class labels, respectively. These metrics are expressed by the following formulas:

Detection Accuracy Accuracy measures the percentage of correctly identified instances, encompassing both normal and attack instances, within a dataset. This metric is defined as the ratio of correctly classified instances to the total number of instances, as shown in Eq. (1). High accuracy signifies that the model is performing effectively, closely aligning its predictions with actual observations. It is essential to consider accuracy in cases where the dataset may be imbalanced to prevent an overrepresentation of FP or FN.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision Precision calculates the model's ability to accurately identify positive instances, particularly in differentiating attacks from irrelevant data. As shown in Eq. (2), precision is the ratio of True Positive Rate (TPR) to the sum of True Positive Rate (TPR) and False Positive Rate (FPR). A higher precision indicates that the model is accurate in its positive predictions and reduces the occurrence of FP. The presence of FP significantly affects the diagnostic accuracy of DDoS attack detection, potentially resulting in unnecessary alerts or actions.

$$Precision = \frac{TPR}{TPR + FPR} \tag{2}$$

Recall Recall evaluates the efficacy of the model in accurately identifying all genuine attacks, with a priority on minimizing the number of missed TP. According to Eq. (3), recall is defined as the ratio of the True Positive Rate (TPR) to the total of the True Positive Rate (TPR) and the False Negative Rate (FNR). Recall underscores

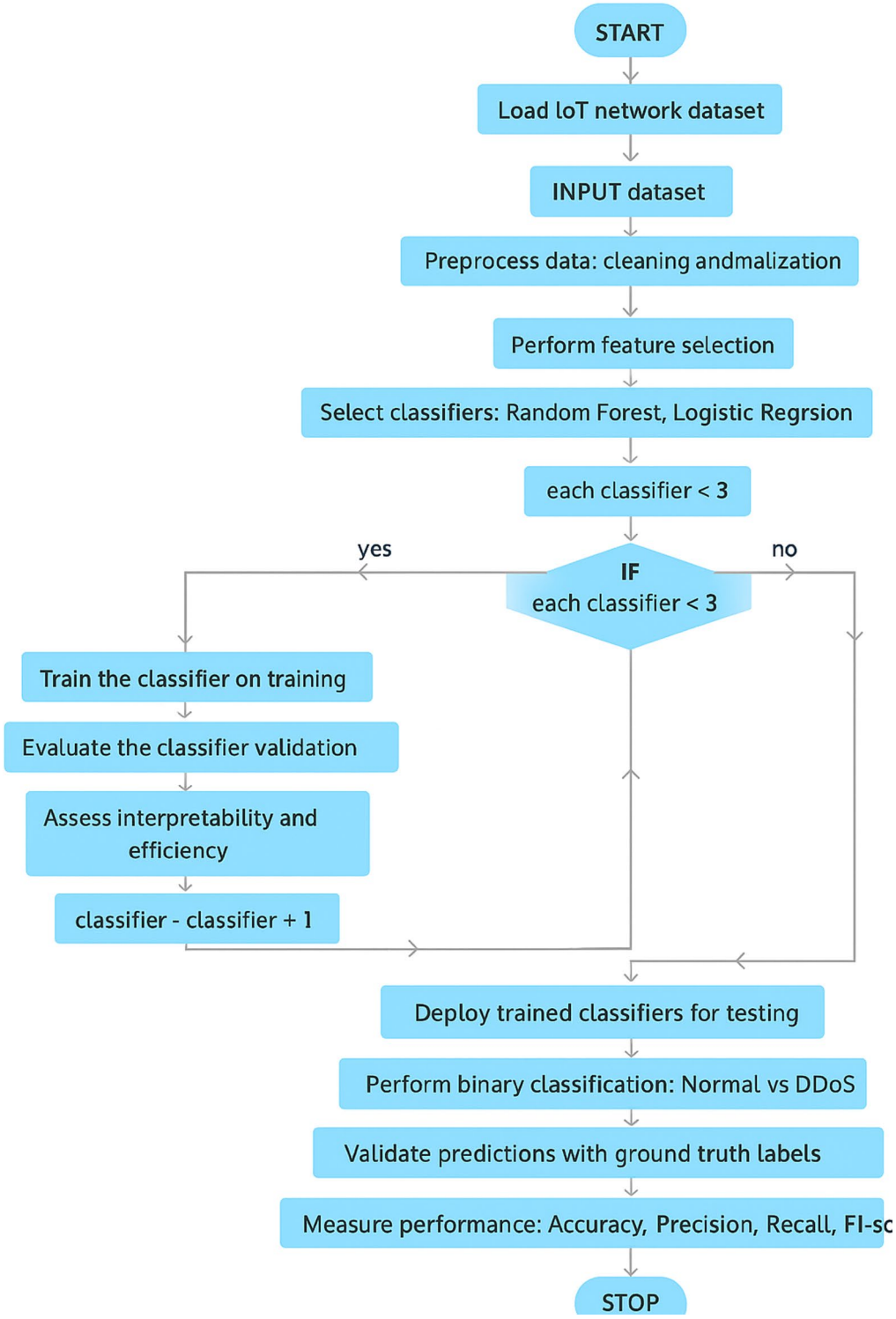


Fig. 7. Methodology pipeline of detection of DDoS attacks.

the model's proficiency in detecting attacks with precision, specifically emphasizing the reduction of FN. While precision quantifies the accurate positive predictions, recall signifies the fraction of positive instances that have been successfully identified. A diminished recall may suggest instances of undetected attacks, potentially undermining the overall effectiveness of the attack detection system.

$$Recall = \frac{TPR}{TPR + FNR} \tag{3}$$

F1-Score The F1-Score is a metric that combines both precision and recall into a single value, emphasizing their symmetry in Eq. (4). It is the harmonic mean of precision and recall, providing a balanced measure of a model's performance. The F1-Score takes into account both FP and FN, offering a more comprehensive evaluation than relying solely on precision or recall. It is computed using the formula below:

$$F1Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

Each of these metrics is reported as an average over the 10 validation folds, and standard deviations are included to reflect the variability in performance. This provides a more comprehensive evaluation of the model's stability under different data splits.

Computational Time Computational time refers to the total time required for a ML model to process data and produce results. It is an important metric for evaluating the efficiency and speed of the model, especially in real-time applications. The average computational time is calculated as the total processing time, representing the overall time taken to complete all necessary computations for a given task.

$$\text{Average Computational Time} = TotalProcessingTime$$

Results and discussion

To evaluate the effectiveness of our approach, we utilize Detection Accuracy, Precision, Recall, F1-score, and Computational Time as performance metrics. The assessed metrics correspond with those found in the existing research, thereby promoting consistent performance comparisons among various models and techniques.

Pseudocode 1 outlines the overall workflow for the proposed intrusion detection system using the NSL-KDD dataset. The system begins by checking whether the input data is in raw format. If it is, preprocessing steps are performed, such as encoding categorical features and normalizing numerical features, to prepare the data for model training. After preprocessing, the dataset is divided into a training set and a test set. An ETC is applied to the training set to rank and select the most important features, which are then utilized to reduce dimensionality in both sets. Following feature selection, three classifiers, RF, NB, and LR, are trained on the training set. Each model is subsequently used to make predictions on the test set. If the predictions are successful, performance metrics, including accuracy, precision, recall, and F1-score, are computed. If the prediction process fails, an error is logged accordingly. This structured workflow ensures data consistency, model comparability, and robust performance evaluation across different classification algorithms.

```
System receives NSL-KDD dataset
IF data == "RAW" then
  Preprocess data:
    - Encode categorical features
    - Normalize numeric features
ENDIF
Split data into TrainingSet and TestSet
Use ExtraTreesClassifier on TrainingSet
Select top important features
Apply feature selection on both sets
FOR model IN [RandomForest, NaiveBayes, LogisticRegression]
  Train model on TrainingSet
  Predict on TestSet
  IF prediction == "successful" then
    Evaluate: Accuracy, Precision, Recall, F1
  ELSE
    Log Error: "Model Prediction Failed"
  ENDIF
ENDFOR
```

Pseudocode 1: Attack detection

By integrating cross-validation and reporting standard deviation, our evaluation approach improves the reliability and credibility of the performance results, particularly when deploying in dynamic, real-world IoT environments. Several classifiers support the proposed technique. Specifically, this method employs three classifiers, RF, NB, and LR, which have been meticulously selected for their efficacy in distinguishing normal traffic patterns from those indicative of DDoS attack data. The verification of each algorithm was conducted through a comprehensive performance analysis encompassing a range of metrics, including accuracy, TPR, FPR, TNR, FNR, as well as computational time.

The comparative performance analysis evaluates these algorithms across seven distinct evaluation metrics. The objective of this research is to ascertain which classifier yields the highest accuracy in the context of DDoS attack detection. Table 2 presents the performance attributes, providing readers with a concise overview of the assessment results. This analytical framework helps identify the most suitable classifier for protecting IoT networks against DDoS attacks.

Table 3 Performance of classification models using tenfold cross-validation. Results are reported as mean ± standard deviation for accuracy, precision, recall, and F1-score. To assess the robustness and consistency

Evaluation Metrics	Random Forest (%)	Logistic Regression (%)	Naive Bayes (%)
Detection Accuracy	99.88	91.61	87.62
Precision	99.93	92.53	83.57
Recall	99.81	91.61	89.30
F1 Score	99.87	90.89	87.40

Table 2. Performance measures of selected lightweight Machine learning models.

Evaluation Metrics	Random Forest (%)	Logistic Regression (%)	Naive Bayes (%)
Detection Accuracy	99.88 ± 0.05	91.61 ± 0.31	87.62 ± 0.45
Precision	99.93 ± 0.04	92.53 ± 0.29	83.57 ± 0.50
Recall	99.81 ± 0.06	91.61 ± 0.32	89.30 ± 0.38
F1 Score	99.87 ± 0.05	90.89 ± 0.34	87.40 ± 0.41

Table 3. Performance measures with cross validation.

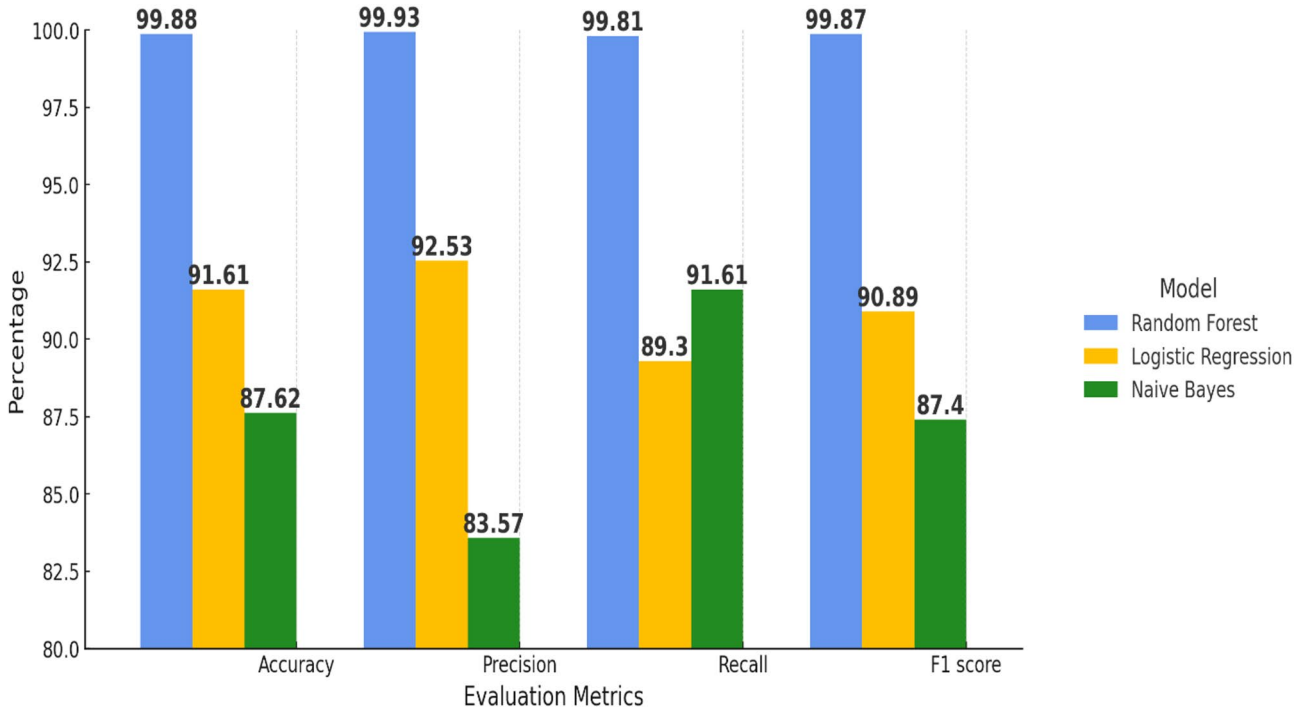


Fig. 8. Performance measure chart of selected Machine learning models.

of the proposed models, a tenfold cross-validation approach was applied. The results demonstrate that the RF classifier achieved the highest performance, with an average detection accuracy of 99.88% ± 0.05, precision of 99.93% ± 0.04, recall of 99.81% ± 0.06, and F1-score of 99.87% ± 0.05, indicating highly stable and accurate DDoS detection. LR followed with a respectable accuracy of 91.61% ± 0.31, while NB attained 87.62% ± 0.45. The low standard deviations across all metrics indicate consistent model behavior across different data splits, affirming the framework's generalizability for detecting DDoS attacks in dynamic and resource-constrained IoT environments.

The findings indicate that the RF algorithm surpasses both LR and NB in all four key evaluation metrics: Accuracy, Precision, Recall, and F1-score, as shown in Fig. 8. Notably, it achieves the highest classification accuracy of 99.88%, establishing RF as the most effective model for identifying normal attack combinations. The RF attains maximum precision with a score of 99.93%, reflecting an optimal balance between true attack detection and minimal false alarm rates. According to the Recall measurement, RF's capability to detect actual attacks is recorded at 91.81%. Furthermore, RF exhibits the leading F1-score of 99.87%, which underscores its proficiency in maintaining a favourable relationship between precision and recall rates. In contrast, LR yields satisfactory results, with an Accuracy of 91.61% and an F1-score of 90.89%; however, its Precision of 92.53% and Recall of 89.3% fall short when compared to those of the RF model. The F1-score of 87.4% for NB highlights its insufficient performance metrics across all evaluation criteria in the detection of DDoS attacks relative to the other classifiers analyzed. Therefore, RF emerges as the most viable model for detecting DDoS attacks within IoT networks.

The comparative performance evaluation of the proposed RF, LR, and NB classifiers against the state-of-the-art CNN-GRU_SMA architecture demonstrates notable advancements, as shown in Fig. 9 and Table 4. Among the models tested, RF consistently outperforms the others across all evaluation metrics, achieving the highest accuracy of 99.88%, precision of 99.93%, recall of 99.81%, and F1-score of 99.87%. These results underscore the model's exceptional ability to accurately identify and classify DDoS traffic.

Although the CNN-GRU_SMA model exhibits competitive performance, achieving a precision of 98.18%, recall of 99.44%, and F1-score of 98.45%, it still trails slightly behind RF. Logistic Regression and NB, while demonstrating acceptable outcomes, report significantly lower metrics, particularly in precision and recall,

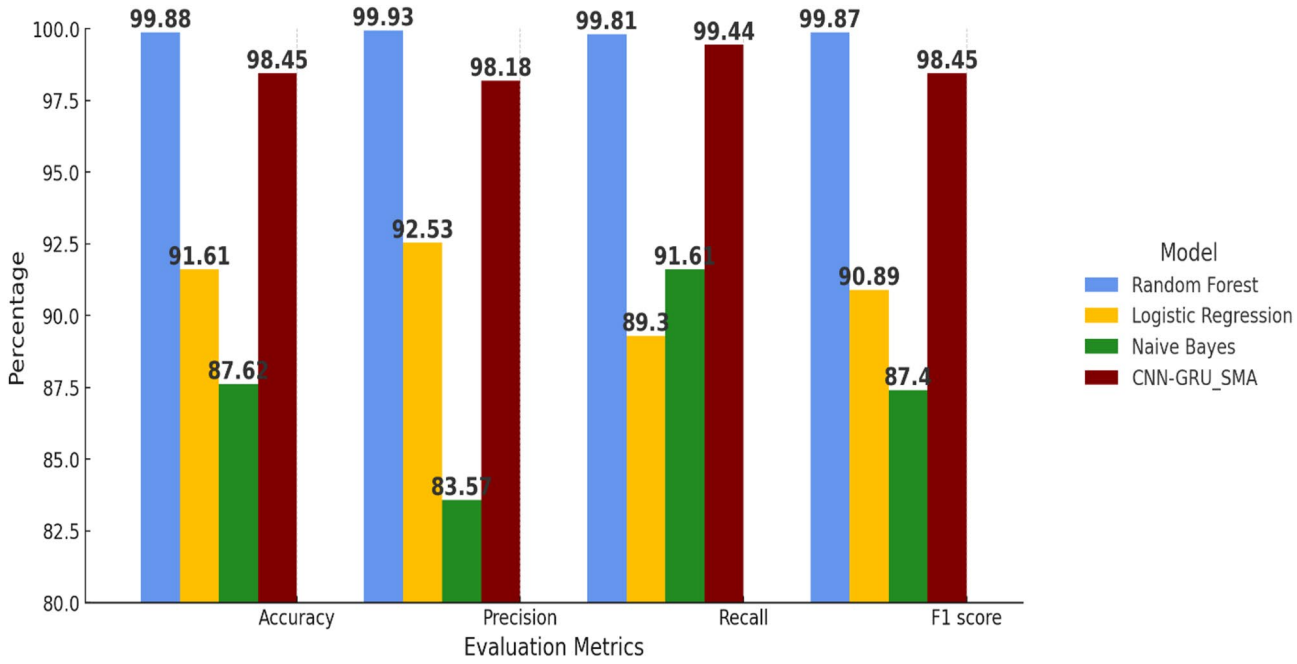


Fig. 9. Performance measure with base paper.

Models	Detection Accuracy (%)	Precision (%)	Recall (%)
Random Forest	99.88	99.93	99.81
Logistic Regression	91.61	92.53	91.61
Naïve Bayes	87.62	83.57	89.30
FFNN	98.67	98.30	99.22
LSTM	96.44	95.74	97.66
RNN	98.68	98.07	99.60
DCAE	92.45	92.46	92.45
CNN-GRU-SMA	98.45	98.18	99.44

Table 4. Comparative Evaluation of the Proposed Framework Against Existing Techniques.

with NB showing the weakest performance across all evaluation criteria. These findings affirm that while deep learning-based models like CNN-GRU_SMA offer a viable approach for DDoS detection, the RF classifier strikes an optimal balance between detection accuracy and computational simplicity, making it particularly well-suited for deployment in resource-constrained IoT environments.

Computational time tests show that CNN-GRU_SMA requires the longest runtime, despite its impressive classification accuracy. As shown in Fig. 10, the RF model achieves the best classification accuracy, but it operates with a computational time of 32.0 s, which is significantly faster than CNN-GRU_SMA's 70.3 s. The longer processing time for CNN-GRU_SMA is due to its complex model structure and advanced architectural design. Although CNN-GRU_SMA provides excellent classification accuracy, RF delivers comparable precision with a substantial speed advantage. The fastest model in the comparison is NB, completing execution in just 8.0 s, demonstrating its adaptability for less complex applications or lower-data volume scenarios. LR, with a runtime of 92.0 s, is the most time-consuming approach among the four models. The analysis of processing times highlights the need for a balance between computational performance and model accuracy, particularly in time-sensitive DDoS detection systems. While CNN-GRU_SMA requires longer processing times, it offers strong accuracy, but RF stands out with both high accuracy and faster processing times. NB excels in speed, making it suitable for simpler applications.

To validate the effectiveness of the proposed lightweight ML framework, its performance was compared against several well-known deep learning-based models commonly used for DDoS detection in IoT networks, including FFNN, LSTM, RNN, DCAE, and CNN-GRU-SMA. As shown in Table 4, the proposed Random Forest (RF) model achieved superior results, with a detection accuracy of 99.88%, precision of 99.93%, and recall of 99.81%. Among the deep learning models, CNN-GRU-SMA showed competitive performance, particularly in recall (99.44%), yet still fell short of the RF model overall.

The superior performance of the RF model can be attributed to several factors. First, its ensemble nature allows it to effectively handle noisy or irrelevant features, which is critical in intrusion detection tasks involving

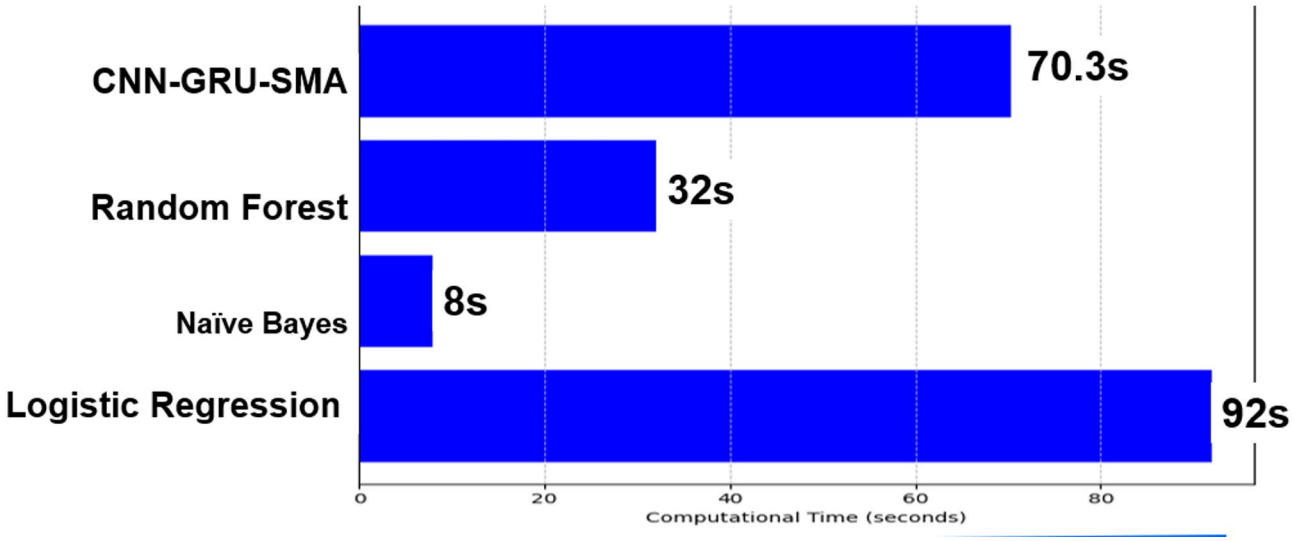


Fig. 10. Computational time comparison.

high-dimensional datasets like NSL-KDD. Additionally, the integration of Extra Trees-based feature selection enhances the model's focus on the most informative attributes, further improving its predictive precision. In contrast, deep learning models, while powerful, often require large datasets to generalize well and can be sensitive to parameter tuning and overfitting in limited or imbalanced data scenarios. They also entail higher computational costs due to complex architectures and longer training times.

In contrast, the proposed RF-based framework maintains high detection performance while remaining computationally lightweight, making it well-suited for real-time deployment in resource-constrained IoT environments. These comparative findings emphasize not only the robustness of the proposed approach but also its practical value in achieving a balance between detection accuracy, interpretability, and deployment feasibility.

Each classifier was trained using the designated training set and subsequently evaluated on the corresponding testing set to ascertain their accuracy, precision, recall, and F1 score in detecting DDoS attacks. A confusion matrix serves as a valuable tool in assessing the effectiveness of a machine learning-based detection system for identifying DDoS attacks, particularly in IoT devices. Figures 11, 12, and 13 present the confusion matrices for the selected lightweight ML models.

The implementation of RF, complemented by enhanced feature selection methodologies such as Error-Tree Correction and Principal Component Analysis (PCA), yielded superior outcomes compared to the initial models delineated in the foundational paper, shown in Fig. 14. Specifically, the RF model leveraging ETC achieved remarkable performance metrics with an Accuracy of 99.88%, Precision of 99.93%, and Recall of 99.81%, culminating in an F1-score of 99.87%. Additionally, when PCA was employed for feature extraction, the RF model sustained commendable performance, achieving an Accuracy of 99.87%, Precision of 99.79%, Recall of 99.94%, and an F1-score of 99.86%. The findings of this research clearly found that the RF methodology, in conjunction with ETC-based feature extraction, offers enhanced accuracy in the detection of DDoS attacks when compared with the detection systems detailed in the foundational study.

The comparative analysis of LR employing two feature selection methodologies, namely ETC.

and PCA, elucidates the significant influence of these techniques on the model's overall performance. The findings indicate that ETC consistently outperforms PCA across all evaluative metrics: Accuracy, Precision, Recall, and F1-score. Specifically, ETC facilitates LR in achieving an impressive Accuracy of 91.61%, accompanied by 92.53% Precision, 89.3% Recall, and an F1-score of 90.89%. Conversely, the PCA technique results in a comparatively lower performance output when applied to the sample dataset. The resultant metrics comprising Accuracy at 90.09%, Precision at 90.41%, Recall at 88.61%, and an F1-score at 89.51% reflect this disparity.

In Fig. 15, the results substantiate that ETC surpasses the efficiency of PCA, optimizing both Precision and Accuracy metrics for LR models, despite PCA exhibiting an acceptable level of effectiveness in model construction. The selected feature selection strategy, namely ETC, demonstrates superior attributes that empower LR to achieve heightened detection capabilities concerning DDoS attacks on IoT devices. Comprehensive analysis indicates that both ETC and PCA contribute significantly to performance enhancements for LR models, thereby underlining the importance of feature selection methodologies in model optimization.

The comparative analysis of the NB classifier reveals several significant performance discrepancies, as demonstrated in Fig. 16. The ETC consistently yields superior results across all evaluated metrics: Accuracy, Precision, Recall, and F1-score. Specifically, when utilizing ETC, NB attains an Accuracy of 87.62%, Precision of 83.57%, Recall of 91.61%, and an F1-score of 87.4%. Conversely, the application of PCA for feature selection results in marginally lower performance metrics, with Accuracy recorded at 87.14%, Precision at 80.11%, Recall at 91.29%, and an F1-score of 85.34%.

The ETC methodology consistently outperforms PCA in terms of Precision and F1-score, while PCA exhibits a slight advantage in Recall. However, the reduction in Precision associated with PCA underscores the inherent

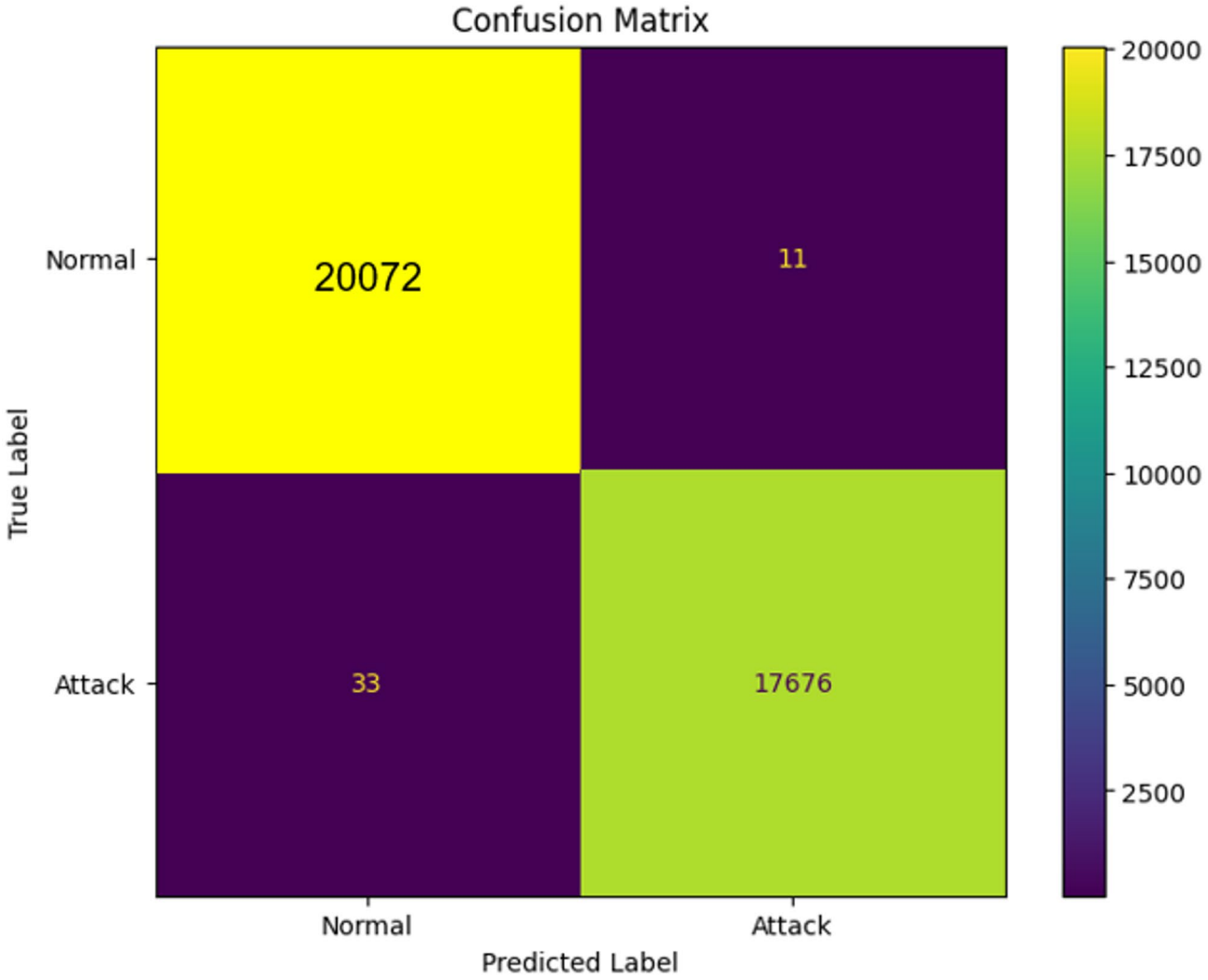


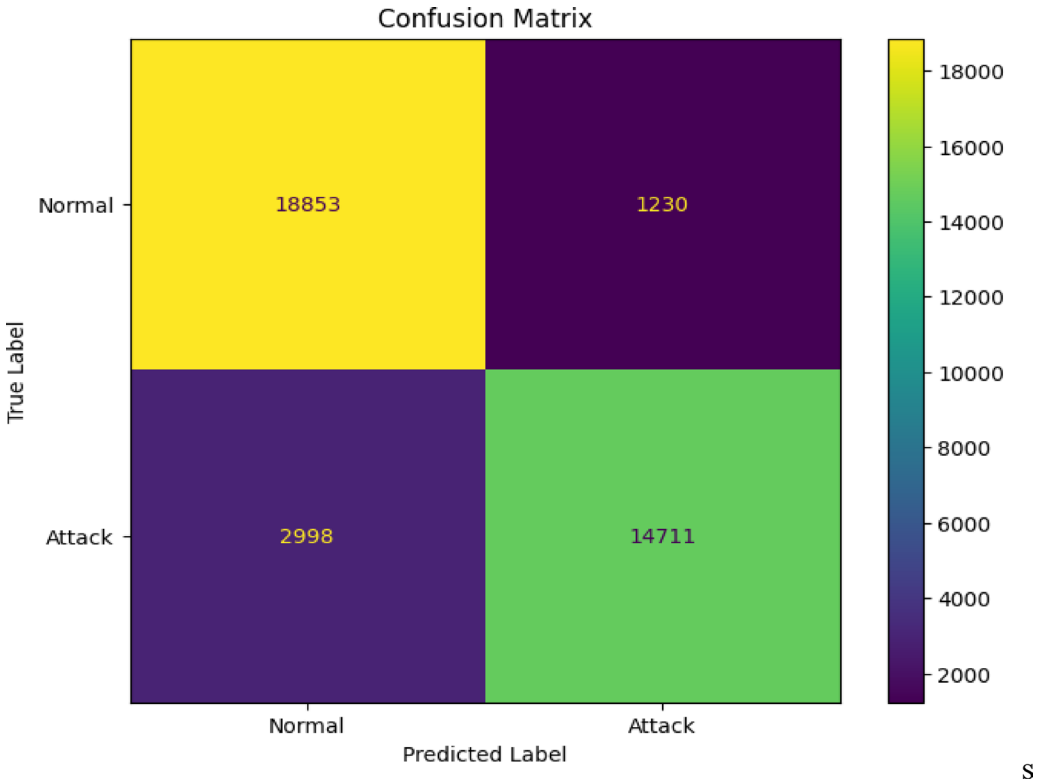
Fig. 11. Confusion matrix of Random Forest.

trade-offs involved in employing various feature selection techniques. When juxtaposed with the findings documented in the foundational paper, NB utilizing the ETC validates a pronounced enhancement in Precision and F1-score, thus establishing it as a more dependable model for the detection of DDoS attacks, although it is still less proficient than more sophisticated models such as RF. These results indicate that the ETC constitutes a more effective feature selection strategy for NB, significantly augmenting its capacity to identify attacks without detracting from its accuracy.

An analysis of the performance of the RF model, both with and without the implementation of the ETC for feature selection purposes, shown in Fig. 17. The findings reveal a substantial improvement in the model's performance upon the application of feature selection through ETC. Specifically, the accuracy exhibits a marked enhancement from 91.39% in the absence of feature selection to 99.88% with feature selection implemented. Precision demonstrates a significant increase from 87.03 to 99.93% following the inclusion of ETC. In a similar vein, recall escalates from 91.34 to 99.81%, and the F1 score also reflects an advancement from 87.65 to 99.87%. These results indicate that the utilization of ETC for feature selection markedly elevates the performance of the RF model across all assessed metrics, thereby emphasizing the effectiveness of ETC in enhancing model accuracy, precision, recall, and F1 score.

The performance comparison of the LR model, both with and without the ETC for feature selection, evaluated across four key metrics: Accuracy, Precision, Recall, and F1 Score. With feature selection, the model exhibits a significant rise in Accuracy, increasing from 74.53% without feature selection to 91.61% with it. Precision shows in Fig. 18 a notable increase as well, climbing from 62.65 to 92.53%. Recall improves from 74.43 to 91.61%, and the F1 Score experiences a considerable enhancement from 66.5 to 90.89%. These findings indicate that utilizing ETC for feature selection greatly enhances the performance of the LR model, boosting all essential metrics.

NB model's performance with and without feature selection via the ETC, evaluated across four metrics: Accuracy, Precision, Recall, and F1 Score, shown in Fig. 19. Implementing feature selection leads to a marked enhancement in Accuracy, which increases from 35.89% (without feature selection) to 87.62% (with feature selection). Precision also grows significantly, from 51.12 to 83.57% post feature selection. Recall sees a rise



S

Fig. 12. Confusion matrix of Logistic Regression.

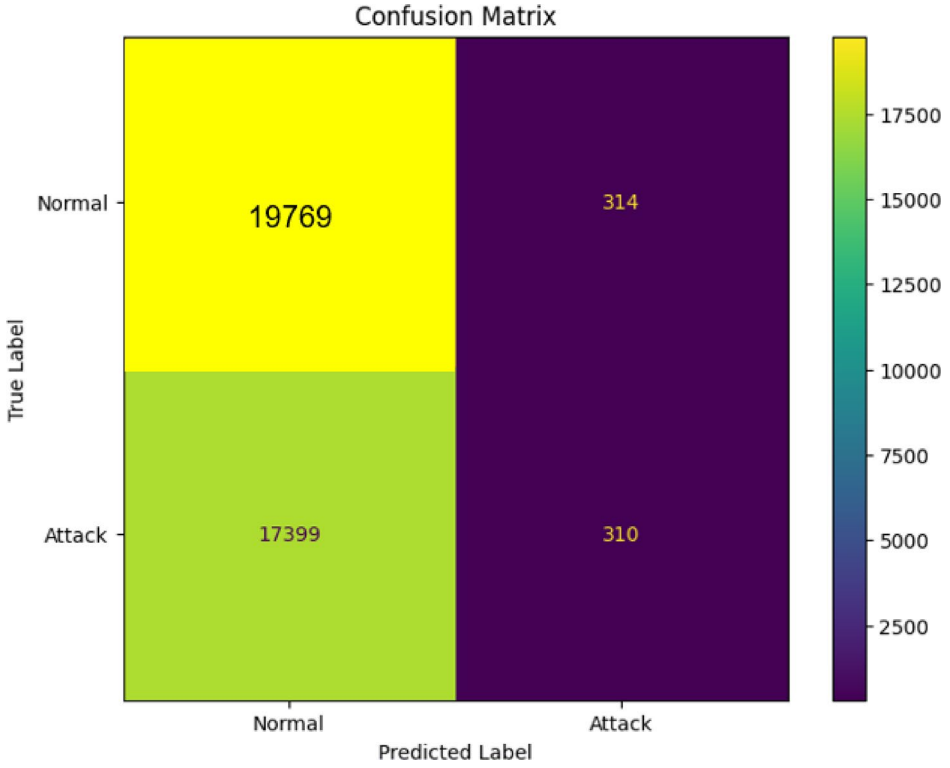


Fig. 13. Confusion matrix of Naïve bayes.

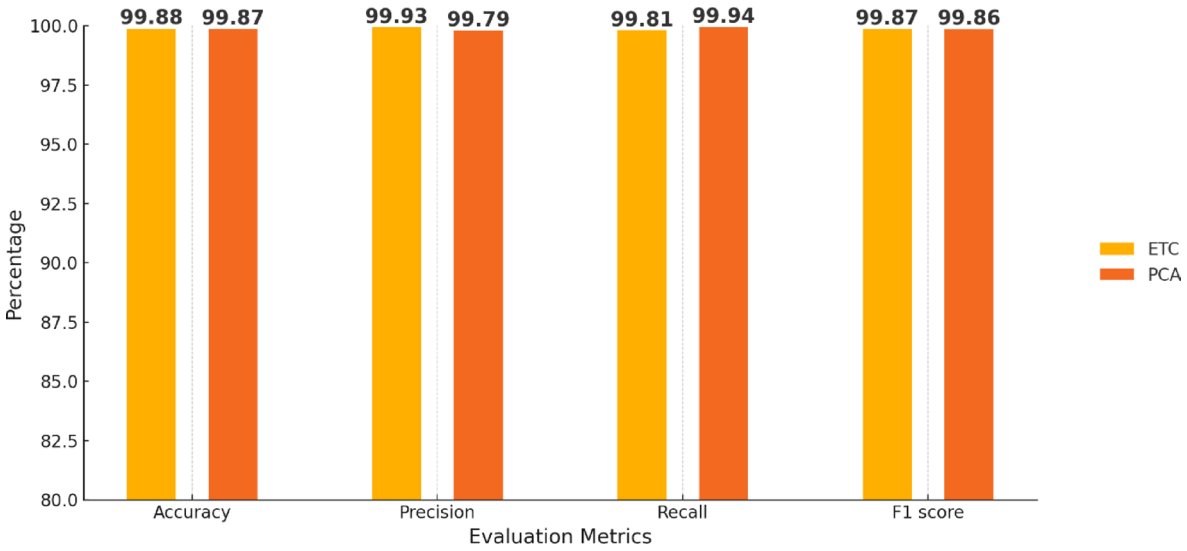


Fig. 14. Performance measure of Random Forest.

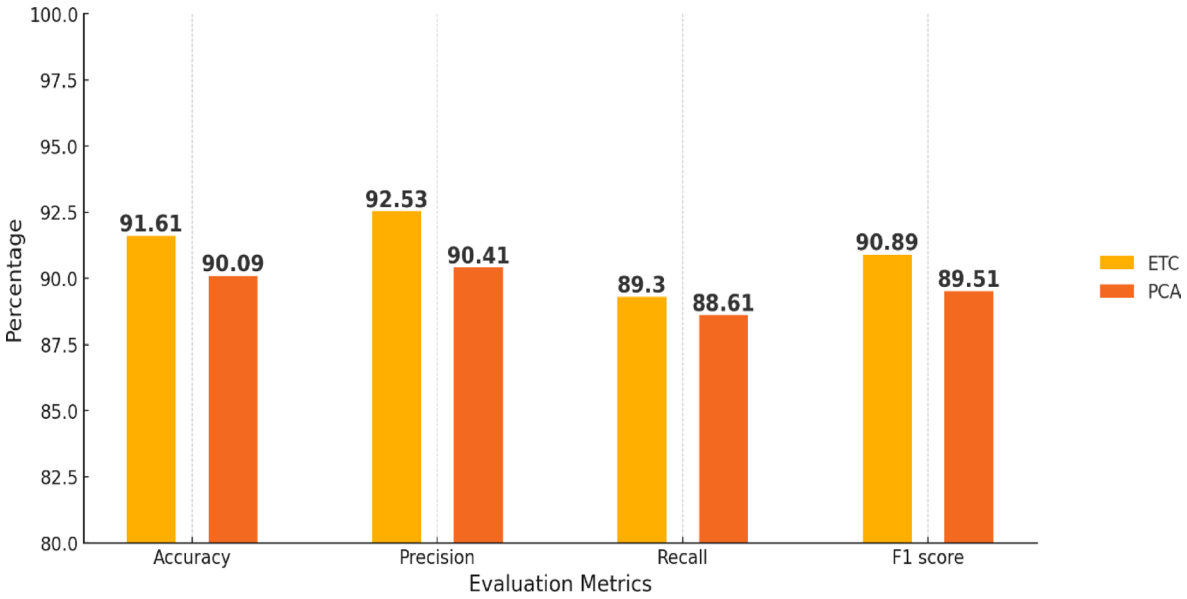


Fig. 15. Performance measure of Logistic Regression.

from 35.87 to 91.61%, and the F1 Score shows impressive growth from 26.96 to 87.4%. These findings indicate that utilizing ETC for feature selection significantly boosts the NB model's performance, notably enhancing all primary metrics, particularly Recall and F1 Score.

A visual comparison known as the Receiver Operating Characteristic (ROC) curve illustrates the detection accuracy of DDoS attacks by examining Gaussian Naive Bayes alongside LR and RF classifiers, as shown in Fig. 20. AUC values accompany each model to provide a numerical assessment of performance metrics. RF proves to be superior to both LR and Gaussian Naive Bayes in balancing the TP Rate, achieving an area under the curve (AUC) of 0.92 according to the ROC curve analysis. RF shows the strongest ability to distinguish between attack and normal traffic, indicated by its position utmost from the random guessing diagonal line. The performance metrics for LR and NB yield lower results, although LR maintains a slight advantage by attaining a better AUC than NB.

The overall results illustrate that the application of the ETC for feature selection significantly enhances the detection of DDoS attacks targeting IoT devices. This approach not only optimizes feature representation but also leads to improved performance metrics of ML classifiers. Our study utilized the NSL KDD dataset, which is crucial for analyzing patterns associated with DDoS attacks.

Among the classifiers examined, the RF model emerged as the most effective, outperforming LR and NB in terms of accuracy, precision, recall, and F1-score. RF exhibited superior overall performance, characterized

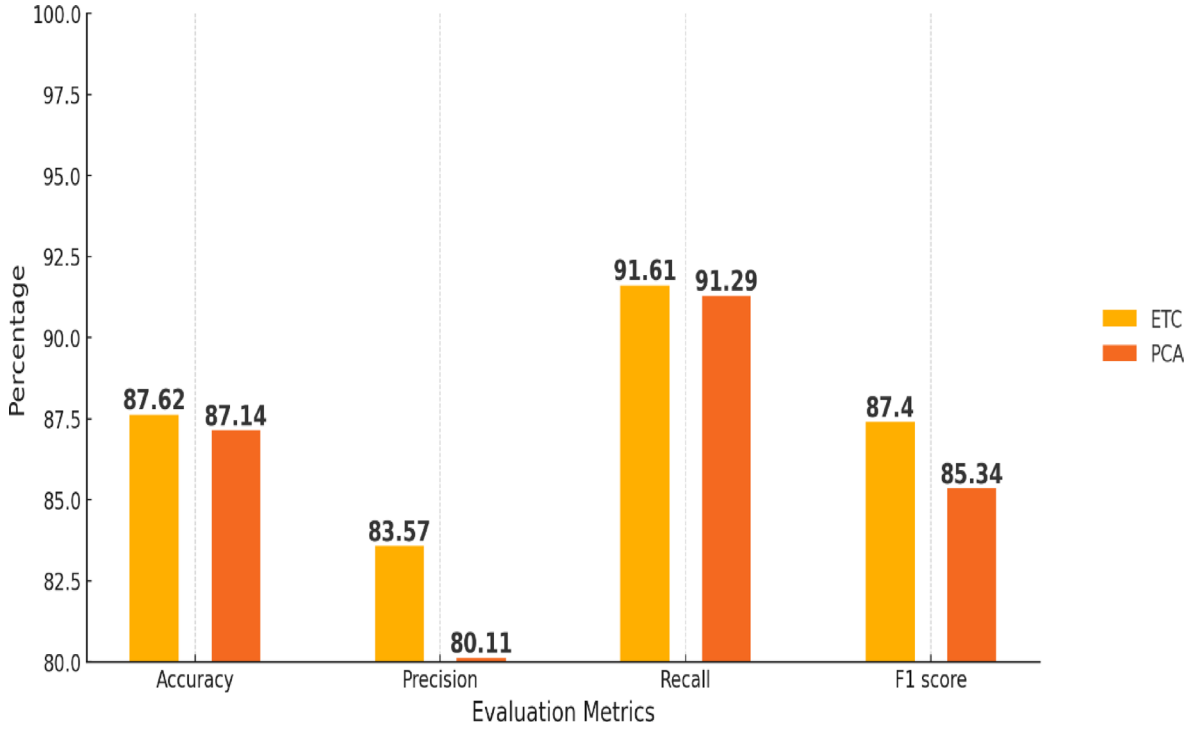


Fig. 16. Performance measure of Naïve bayes.

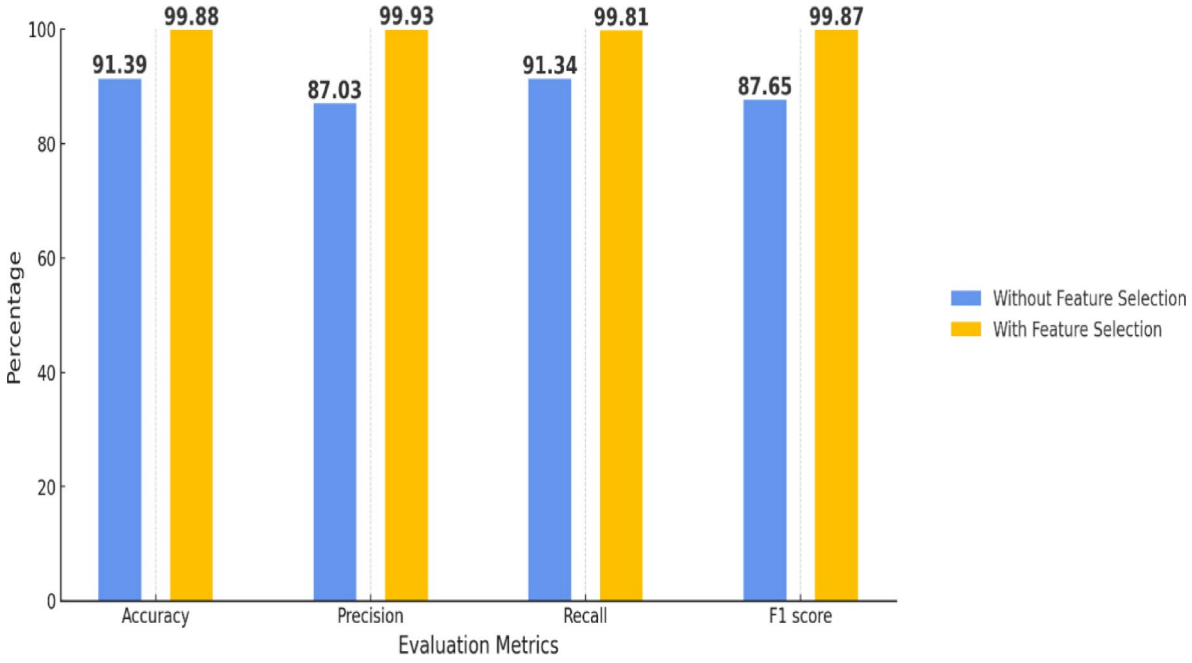


Fig. 17. Comparison of RF performance with and without feature selection.

by reduced training and prediction durations, thereby affirming its suitability for real-time DDoS detection applications. Additionally, feature selection utilizing the Extra Trees methodology demonstrated enhanced efficiency compared to outdated techniques like the Pearson correlation coefficient, resulting in a decreased computational burden while maintaining high levels of accuracy.

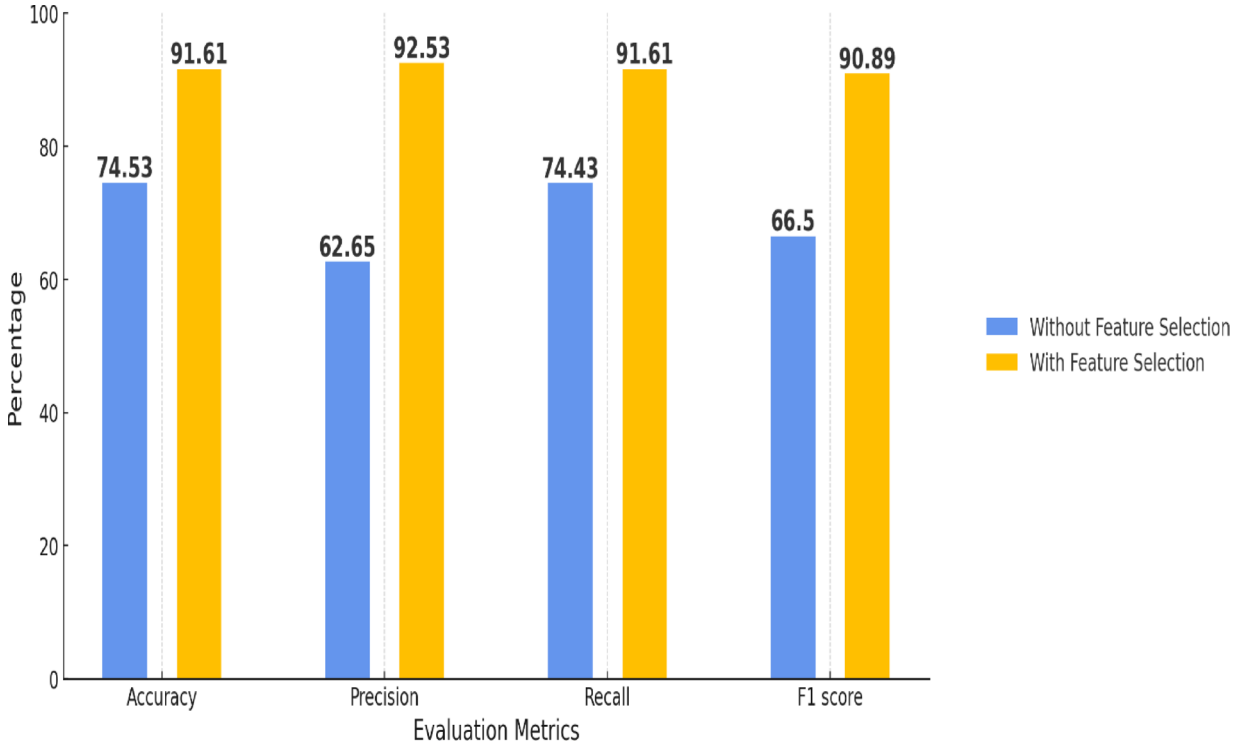


Fig. 18. Comparison of LR performance with and without feature selection.

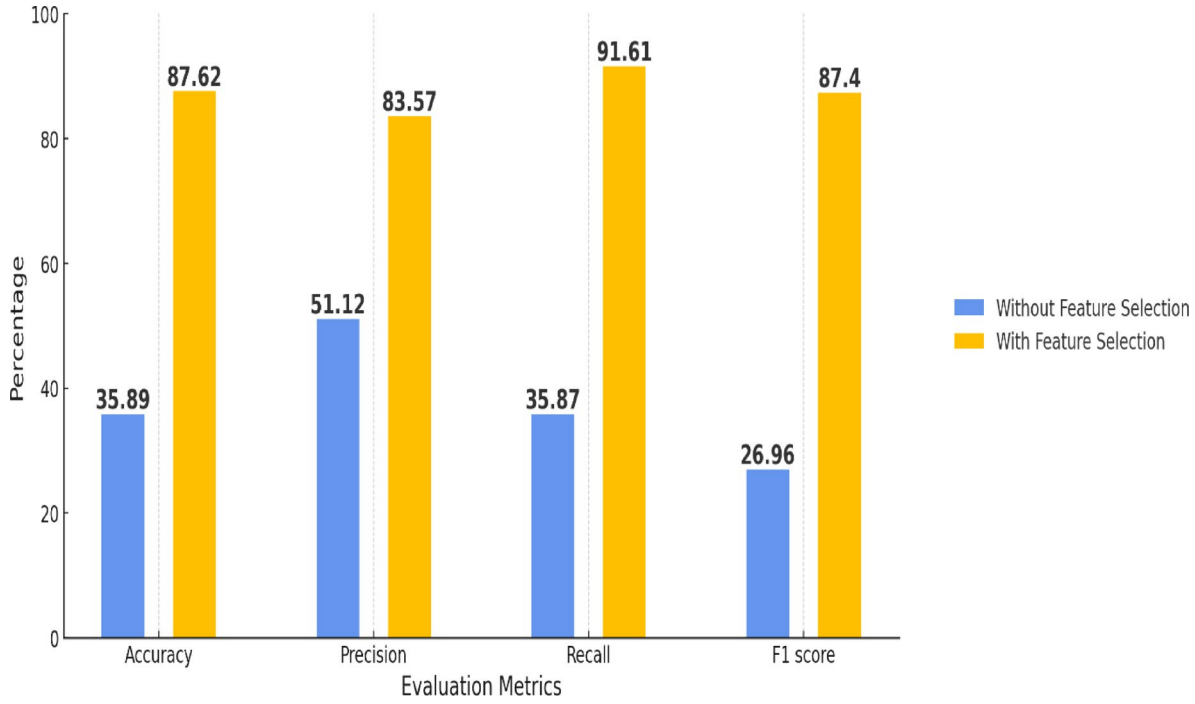


Fig. 19. Comparison of NB performance with and without feature selection.

Conclusion and future work

The detection of DDoS attacks targeting IoT devices represent a pivotal challenge in safeguarding the security and reliability of contemporary networks. This investigation aimed to implement binary classification methodologies for identifying DDoS attacks on IoT devices using RF, LR, and NB models. The findings indicated that all three models demonstrated the capability to detect DDoS traffic. Notably, the RF model demonstrated superior

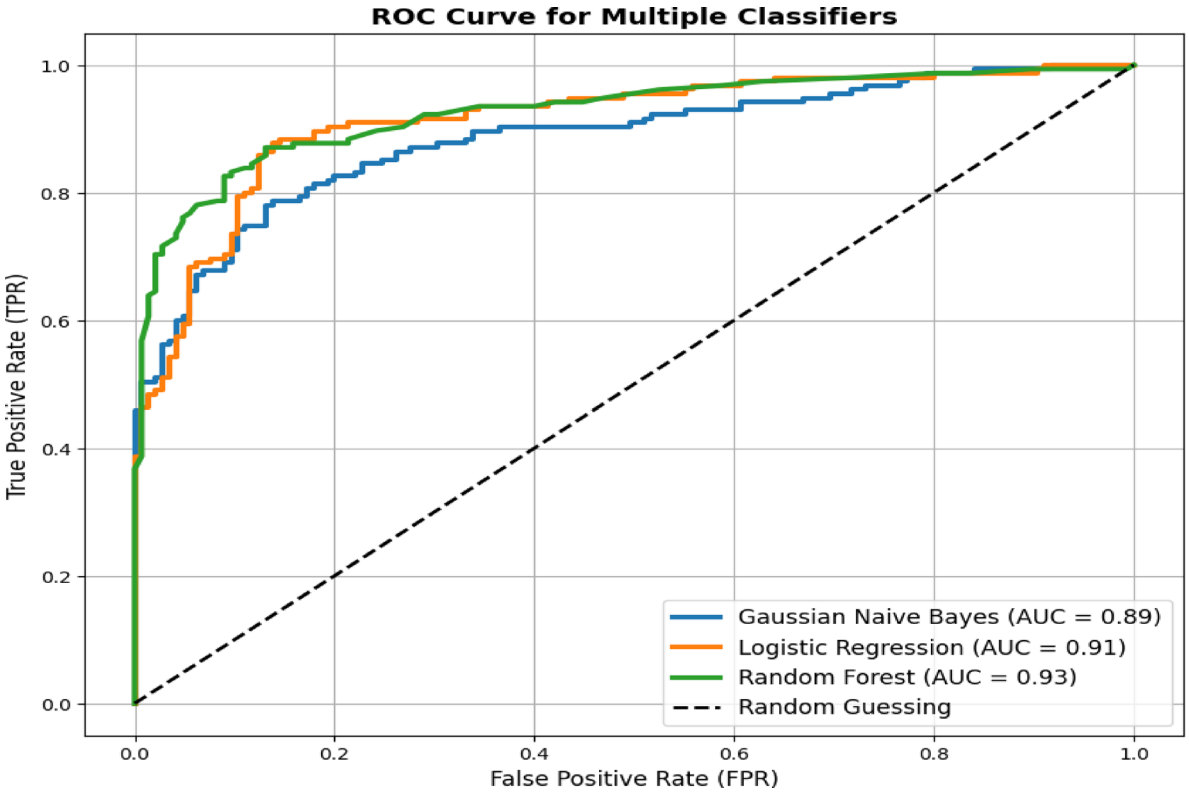


Fig. 20. Comparative performance measure of machine learning models.

performance, achieving the highest accuracy score of 99.88% and F1-scores, which surpassed the comparative efficacy of the other models. Furthermore, integrating the ETC for feature selection significantly improved the efficiency and effectiveness of the detection procedure by omitting unnecessary features while retaining critical attributes. This refined approach ensures scalability and adaptability within IoT environments, which commonly operate under limited resources. However, several limitations must be acknowledged. The reliance on the NSL-KDD dataset, while widely used for benchmarking, may not reflect the latest patterns and threats present in real-world IoT traffic, potentially limiting the generalizability of the findings. Assumptions made during preprocessing, such as uniform normalization and categorical encoding, may not optimally translate across varied IoT device configurations. Additionally, the Extra Trees-based feature selection, though efficient, is based on global importance and may not fully capture context-specific or localized feature relevance. These factors may impact the robustness of the model in live environments.

From a practical deployment perspective, real-world integration poses additional challenges, including interoperability across heterogeneous IoT platforms, real-time detection under constrained bandwidth and latency conditions, and the need for lightweight implementations on edge devices. Ensuring compatibility with existing IoT communication protocols and resource management strategies will be crucial for successful implementation. Furthermore, updating models to adapt to evolving threat landscapes remains a vital consideration for sustained effectiveness.

Ultimately, this study contributes to the enhancement of security frameworks for IoT devices in the face of emerging cybersecurity threats by applying ML models and implementing robust feature selection strategies, while also offering insight into the practical considerations necessary for real-world deployment.

Future work

Future research in this domain should prioritize several critical areas to enhance the detection of DDoS attacks targeting IoT devices. Addressing the current limitations, future studies should incorporate more recent and diverse datasets to reduce bias and better reflect the complexity of real-world IoT environments. It is imperative to broaden the scope of the study to encompass a diverse array of attack vectors and con of IoT devices, as this will be vital for developing a more robust detection framework. Additionally, assessing the scalability and real-time performance of the proposed models across various network environments is essential for determining their effectiveness under a range of conditions.

The application of similar ML techniques to other IoT-specific datasets will not only bolster the validation of existing methodologies. It may also reveal new insights that could further optimize detection accuracy. Furthermore, the creation of an intuitive front-end application capable of detecting attacks on various IoT devices would facilitate effective implementation and enhance user engagement with security protocols. Collectively,

these initiatives are poised to significantly improve DDoS detection in IoT networks, offering stronger defenses against the evolving landscape of cyber threats.

Data availability

The dataset used and analyzed during the current study is available from the NSL KDD dataset repository on Kaggle. https://www.kaggle.com/datasets/hassan06/nslkdd.

Received: 23 April 2025; Accepted: 1 July 2025
Published online: 10 July 2025

References

1. Dachyar, M., Zagloel, T. Y. M. & Saragih, L. R. Knowledge growth and development: Internet of things (IoT) research, 2006–2018. *Heliyon* **5**(8), e02264. https://doi.org/10.1016/j.heliyon.2019.e02264 (2019).
2. Hussein, A. H. Internet of things (IOT): Research challenges and future applications, *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 6, 2019.
3. Mouha, R. A. R. A. Internet of things (IoT). *J. Data Anal. Inf. Process.* **09**(02), 77–101. https://doi.org/10.4236/jdaip.2021.92006 (2021).
4. Mu, X. & Antwi-Afari, M. F. The applications of Internet of Things (IoT) in industrial management: A science mapping review. *Int. J. Prod. Res.* **62**(5), 1928–1952. https://doi.org/10.1080/00207543.2023.2290229 (2024).
5. Gelgi, M., Guan, Y., Arunachala, S., Rao, M. S. S. & Dragoni, N. Systematic literature review of IoT botnet DDOS attacks and evaluation of detection techniques. *Sensors* **24**(11), 3571. https://doi.org/10.3390/s24113571 (2024).
6. Kayode Saheed, Y., Harazeem Abdulganiyu, O. & Ait Tchakoucht, T. A novel hybrid ensemble learning for anomaly detection in industrial sensor networks and SCADA systems for smart city infrastructures. *J. King Saud Univ. Comput. Inf. Sci.* **35**(5), 101532. https://doi.org/10.1016/j.jksuci.2023.03.010 (2023).
7. Lamptey, R., Saedi, M., Stankovic, V. Machine-Learning anomaly detection for early identification of DDOS in smart home IoT devices.
8. Saheed, Y. K. & Chukwuere, J. E. XAIEnsembleTL-IoV: A new eXplainable artificial intelligence ensemble transfer learning for zero-day botnet attack detection in the internet of vehicles. *Results Eng.* **24**, 103171. https://doi.org/10.1016/j.rineng.2024.103171 (2024).
9. Abdulganiyu, O. H., Ait Tchakoucht, T., Alaoui, A. E. H. & Saheed, Y. K. Attention-driven multi-model architecture for unbalanced network traffic intrusion detection via extreme gradient boosting. *Intell. Syst. Appl.* **26**, 200519. https://doi.org/10.1016/j.iswa.2025.200519 (2025).
10. Kayode Saheed, Y. & Ebere Chukwuere, J. CPS-IIoT-P2Attention: Explainable privacy-preserving with scaled dot-product attention in cyber-physical system-industrial IoT network. *IEEE Access* **13**, 81118–81142. https://doi.org/10.1109/ACCESS.2025.3566980 (2025).
11. Mahadik, S., Pawar, P. M. & Muthalagu, R. Efficient intelligent intrusion detection system for heterogeneous internet of things (HetIoT). *J. Netw. Syst. Manag.* **31**(1), 2. https://doi.org/10.1007/s10922-022-09697-x (2023).
12. Elgazzar, K. et al. Revisiting the internet of things: New trends, opportunities and grand challenges. *Front. Internet Things* **1**, 1073780. https://doi.org/10.3389/friot.2022.1073780 (2022).
13. Hussain, F., Abbas, S. G., Husnain, M., Fayyaz, U. U., Shahzad, F., Shah, G. A. IoT DoS and DDoS attack detection using ResNet, In *2020 IEEE 23rd International Multitopic Conference (INMIC)*, IEEE, 2020, pp. 1–6.
14. Nižetić, S., Šolić, P., López-de-Ipiña González-de-Artaza, D. & Patrono, L. Internet of things (IoT): Opportunities, issues and challenges towards a smart and sustainable future. *J. Clean. Prod.* **274**, 122877. https://doi.org/10.1016/j.jclepro.2020.122877 (2020).
15. Taherdoost, H. Security and internet of things: Benefits, challenges, and future perspectives. *Electronics* **12**(8), 1901. https://doi.org/10.3390/electronics12081901 (2023).
16. Sadek, I., Codjo, J., Rehman, S. U. & Abdulrazak, B. Security and privacy in the internet of things healthcare systems: Toward a robust solution in real-life deployment. *Comput. Methods Progr Biomed. Update* **2**, 100071. https://doi.org/10.1016/j.cmpbup.2022.100071 (2022).
17. Jullian, O. et al. Deep-learning based detection for cyber-attacks in IoT networks: A distributed attack detection framework. *J. Netw. Syst. Manag.* **31**(2), 33. https://doi.org/10.1007/s10922-023-09722-7 (2023).
18. Silivery, A. K., Rao Kovvur, R. M., Solleti, R., Kumar, L. S. & Madhu, B. A model for multi-attack classification to improve intrusion detection performance using deep learning approaches. *Meas. Sens.* **30**, 100924. https://doi.org/10.1016/j.measen.2023.100924 (2023).
19. Aktar, S. & Yasin Nur, A. Towards DDoS attack detection using deep learning approach. *Comput. Secur.* **129**, 103251. https://doi.org/10.1016/j.cose.2023.103251 (2023).
20. Motylinski, M., MacDermott, Á., Iqbal, F. & Shah, B. A GPU-based machine learning approach for detection of botnet attacks. *Comput. Secur.* **123**, 102918. https://doi.org/10.1016/j.cose.2022.102918 (2022).
21. Ismail, et al. A machine learning-based classification and prediction technique for DDoS attacks. *IEEE Access* **10**, 21443–21454. https://doi.org/10.1109/ACCESS.2022.3152577 (2022).
22. Ahmad, I., Wan, Z. & Ahmad, A. A big data analytics for DDOS attack detection using optimized ensemble framework in internet of things. *Internet Things* **23**, 100825. https://doi.org/10.1016/j.iot.2023.100825 (2023).
23. Hariprasad, S. Detection of DDoS Attack in IoT networks using sample selected RNN-ELM, 2022.
24. Almadhor, A., Altalbe, A., Bouazzi, I., Hejaili, A. A. & Kryvinska, N. Strengthening network DDOS attack detection in heterogeneous IoT environment with federated XAI learning approach. *Sci. Rep.* **14**(1), 24322. https://doi.org/10.1038/s41598-024-76016-6 (2024).
25. Sanmorino, A., Marnisah, L. & Kesuma, H. D. Detection of DDoS attacks using fine-tuned multi-layer perceptron models. *Eng. Technol. Appl. Sci. Res.* **14**(5), 16444–16449. https://doi.org/10.48084/etasr.8362 (2024).
26. Revathi, M., Ramalingam, V. V., Amutha, B. A machine learning based detection and mitigation of the DDOS attack by using SDN controller framework, *Wirel. Pers. Commun.*, pp. 1–25, 2021.
27. Mazid, A., Kirmani, S. & Abid, M. Enhanced intrusion detection framework for securing IoT network using principal component analysis and CNN. *Inf. Secur. J. Glob. Perspect* https://doi.org/10.1080/19393555.2024.2408256 (2024).
28. Alve, S. R., Mahmud, M. Z., Islam, S., Chowdhury, M. A., Islam, J. Smart IoT security: Lightweight machine learning techniques for multi-class attack detection in IoT networks Feb. 06, 2025, *arXiv*: arXiv:2502.04057. https://doi.org/10.48550/arXiv.2502.04057.
29. "NSL-KDD." Accessed: Aug. 29, 2024. [Online]. Available: https://www.kaggle.com/datasets/hassan06/nslkdd

Author contributions

M.N. conceived the study, designed the methodology, implemented the experiments, conducted data analysis,

and wrote the original manuscript. S.T. supervised the research, provided critical feedback on the methodology, reviewed/edited the manuscript, and secured funding. D.S. contributed to data preprocessing and feature selection implementation. S.A. assisted with model validation and performance evaluation. M.T. provided domain expertise in IoT security, reviewed the manuscript, and helped with result interpretation. All authors discussed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

Additional information

Correspondence and requests for materials should be addressed to M.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.