Performance Evaluation of Machine Learning Classifiers for Predicting Denial-of-Service Attack in Internet of Things

Omar Almomani[1*], Adeeb Alsaaidah[2], Ahmad Adel Abu Shareha[3], Abdullah Alzaqebah[4], Malek Almomani[5]
Information System and Network Department, The World Islamic Sciences and Education University, Amman, 11947, Jordan[1]
Department of Networks and Information Security, Al-Ahliyya Amman University, Amman 19328, Jordan[2]
Department of Data Science and Artificial Intelligence, Al-Ahliyya Amman University, Amman, Jordan[3]
Computer Science Department, Al-Ahliyya Amman University, Amman, Jordan[4]
Software Engineering Department, The World Islamic Sciences and Education University, Amman, 11947, Jordan[5]

Abstract*—Eliminating security threats on the Internet of Things (IoT) requires recognizing threat attacks. IoT and its implementations are currently the most common scientific field. When it comes to real-world implementations, IoT's attributes, on the one hand, make it simple to apply, but on the other hand, they expose it to cyber-attacks. Denial of Service (DoS) attack is a type of threat that is now widespread in the field of IoT. Its primary goal is to stop or damage service or capability on a target. Conventional Intrusion Detection Systems (IDS) are no longer sufficient for detecting these sophisticated attacks with unpredictable behaviors. Machine learning (ML)--based intrusion detection does not need a massive list of expected activities or a variety of threat signatures to create detection rules. This study aims to evaluate different ML classifiers for network intrusion detection that focus on DoS attacks in the IoT environment to determine the best ML classifier that can detect the DoS attack. The XGBoost, Decision Tree (DT), Gaussian Naive Bayes (NB), Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM) ML classifiers are used to evaluate the DoS attack. The UNSW-NB15 dataset was used for this study. The obtained accuracy rate for XGboost was 98.92%, SVM 98.62%, Gaussian NB 83.75%, LR 97.74%, RF 99.48%, and DT 99.16%. where the precision rate for XGboost, SVM, Gaussian NB, LR, RF, and DT was 98.40%, 98.29%, 77.50%, 97.14%, 99.21%, and 99.12%, respectively. The sensitivity rate for XGboost, SVM, Gaussian NB, LR, RF, and DT was 99.29%, 98.76%, 91.87%, 98.06%, 99.69%, and 99.08%, respectively. The results show that the RF classifier outperformed other classifiers in terms of Accuracy, Precision, and Sensitivity.

Keywords*—Cybersecurity; IDS; DOS attack; IoT; machine learning

I. Introduction

The IoT consists of various physical objects such as machines, vehicles, and structures equipped with sensors, software, and connectivity that enable them to accumulate and transmit data. In addition, these devices can communicate with one another and the Internet, allowing them to send and receive data and be remotely controlled. Intelligent appliances, wearable technology, industrial equipment, and thermostats with Internet connectivity are a few examples of IoT devices, With the help of the IoT, several processes can be automated, and massive amounts of data can be collected. These benefits include increased productivity, lower costs, and better user experiences. Additionally, it opens new avenues for researchers' innovation.

IoT security has become a big problem as connected devices increase [1]. IoT devices are susceptible to hacking and other cyber-attacks since they frequently have low processor speed, memory capacity, and security features. Device spoofing, Man-in-the-Middle attacks, DoS, Ransomware, and unauthorized access are a few common types of IoT attacks [2] [3] [4]. The DoS attack [5] is a kind of cyberattack in which the attacker tries to block access to a device or network by legitimate users by flooding it with uncontrollable data. This can be done by deploying a botnet, or network of infected devices, to flood the target device or network with a lot of traffic, making it unavailable. Security techniques like firewalls, IDS, and traffic filtering can be used to detect and prevent malicious traffic from defending against DoS attacks in IoT [6] [7] [8].

IDS [9] [10] [11]security techniques keep a watch out for illegal behavior on a network and notify an admin of any severe violations or attacks. They can spot various security risks, including malware, illegal access, and (DoS) attacks. IDS has two types [12, 13]. Network-based IDS (NIDS) monitors network activity for any improper behavior. To monitor all incoming and outgoing traffic, they are frequently positioned at crucial nodes on a network, such as a firewall or a router. Host-based IDS (HIDS): this type keeps a watch on what is going on with a particular host or device, like a server or an IoT device. They can identify unwanted access to or alterations to host-based data and settings.

ML approaches can be used to increase the accuracy, precision, and effectiveness of IDS in identifying security attacks like DoS. ML is divided into three approaches. First, Supervised learning: This approach trains a model to categorize network traffic as benign or malicious using labeled data. This approach is trained to recognize well-known harmful patterns that are frequently used for signature-based intrusion detection. Secondly, Unsupervised learning, when labeled data is unavailable, unsupervised learning is the preferred option. The model is trained to spot data anomalies or patterns that

differ from expected behavior. Unknown or zero-day attacks can be found using this approach. Finally, the Semi-supervised learning approach mixes supervised and unsupervised learning by training the model on a mixture of labeled and unlabeled data.

IDS-based ML was developed to detect suspicious behavior in the IoT environment. XGBoost, Gaussian NB, DT, RF, LR, and SVM ML classifiers were used to construct the IDS. The developed IDS's primary objective is to assess the efficacy of detecting DoS attacks in an IoT environment. The following is the paper's contributions:

1) An intelligent IDS with high detection accuracy, precision, sensitivity, and F-measure, capabilities to detect DoS attacks in the IoT.

2) Experiments demonstrate the operation of several ML classifiers and their impact on DoS attacks in an IoT environment.

3) Random Forest classifier shows the superiority of detecting DoS attacks in an IoT environment.

The remainder of the paper is organized as follows: Section II covers the background, Section III covers the history and related works, and Section IV illustrates the suggested IDS model. Next, the experimental research design and results are stated in Section V. Finally, Section VI concludes the paper's work and findings.

II. BACKGROUND

This section provides an overview and background information related to the topic under investigation in this paper.

A. IoT

IoT technology was developed by Kevin Ashton in 1999 [14]. The IoT is defined as the interconnection of physical objects such as furniture, cars, buildings, and other things that are connected to the Internet and have electronics, software, sensors, and network connectivity built into them [15, 16]. This makes it possible to create modern software and services for several areas, including manufacturing, healthcare [17], transportation, and smart cities, that can boost productivity, decrease costs, and increase convenience [18]. IoT architecture will develop because of the increased use of IoT technology. Fig. 1 depicts the evolution of IoT architecture.

IoT applications can be broken down into three layers: application, transport, and perception. IoT devices are becoming more common, but the variety of applications for these devices raises questions about security and privacy [20]. Additionally, the distinctive features of IoT pose particular security issues, such as handling, preserving, and protecting the private data that these devices frequently collect. Due to the multiple vulnerabilities in IoT applications, they are vulnerable to various cyber threats. Several security and privacy issues have been documented on IoT apps worldwide, such as the Mirai attack, DOS, and Distributed Denial-of-Service (DDOS) attacks. Fig. 2 shows the types of attacks in IoT.

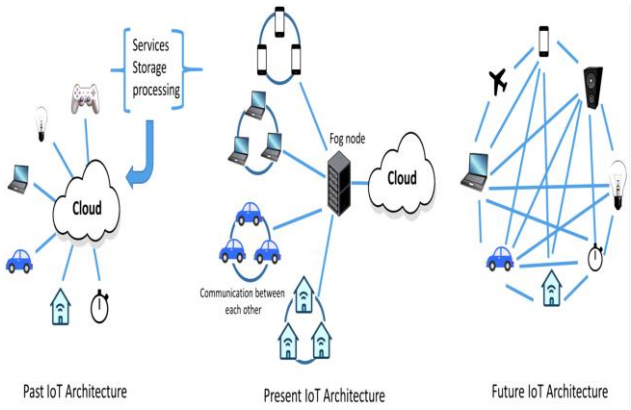

Fig. 1. Evolution of IoT architecture [19].

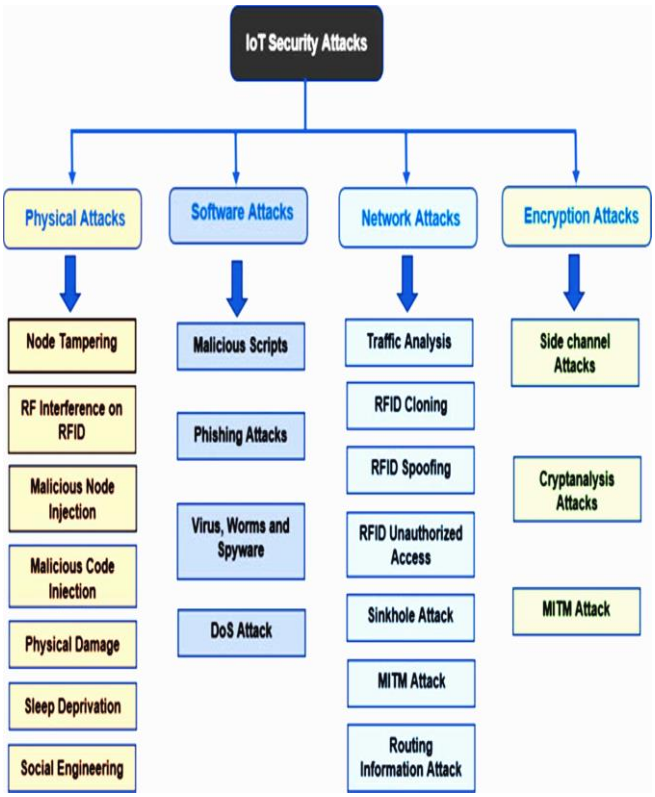

Fig. 2. IoT attack types [21].

For IoT technology to be broadly used, experts and scientists agree that guaranteeing the security of IoT applications is a significant barrier that must be surmounted. Users should have complete confidence in IoT devices and application security. In addition, they must guarantee that their equipment is safe from known threats. as they become increasingly integrated into daily routines. IDS protects IoT networks and devices from harmful activities and illegal access.

B. IDS

An IDS is a technology that examines computer or network systems for any indications of illegal access or policy violations. This can be accomplished by using a host-based or network-based method, and it can be achieved using various

tools, including hardware, software, or a combination of them. IDSs use a variety of detection methods to find potentially dangerous activities, including anomaly-based IDS (AIDS) and signature-based IDS (SIDS) [12, 13]. The IDS can alert system administrators to suspicious activity, log the incident, or take action to stop future intrusion when it is identified. SIDS, also known as Misuse Detection [22], uses pattern-matching algorithms to find known threats. An alarm is triggered when an intrusion is discovered that matches an intrusion signature previously stored in a SIDS system's database. After that, the system searches the host's logs for groups of commands or actions formerly known to be malicious. While SIDS systems typically have excellent detection accuracies for known intrusions, they may have difficulty detecting zero-day attacks since the database does not yet contain the signature of the new threat. To solve the problem of detecting zero-day attacks, AIDS is used.

Because it can surpass SIDS' limitations, AIDS has attracted much interest from researchers. AIDS creates a computer system behavior model based on machine learning, statistical, or knowledge-based techniques. An anomaly, which is viewed as an intrusion, is any significant divergence from the model. This set of methods is predicated on malicious conduct deviating from user behavior. AIDS can identify unknown or zero-day attacks because It is independent of signature databases. Such as SIDS to detect abnormal behavior.

IDS divided the data into two groups depending on the input source: Host IDS(HIDS) and Networks IDS (NIDS). HIDS analyzes data from the host system, including sources such as operating system logs, firewall logs, and database logs. IDS can recognize insider attacks that don't use network traffic, where the NIDS analyzes information from sources like packet capture to keep track of network activity. It can be utilized to monitor several computers linked to a network and detect early signs of external malicious activity before it spreads to other systems. To identify abnormal activity, ML algorithms, including XGBoost, TD, RF, Gaussian NB, LR, and SVM, have been used in the AIDS domain. The following section explains the ML algorithms used in this paper.

C. ML Classifiers

ML is a technique for instructing computers to learn from data without explicit programming. It is a subfield of artificial intelligence that enables systems to improve automatically over time. Supervised, unsupervised, and reinforcement learning are the three main types of machine learning [23, 24]. Supervised learning solves issues when the only available data consists of labeled instances. Unsupervised learning is used to find the pattern of unlabeled data. With reinforcement learning, a computer agent learns to perform a task by repeatedly attempting it and modifying its behavior in response to the feedback it receives. The model that is trained to categorize input data into classes or categories is known as a classifier. There are numerous classifier varieties, each with unique advantages and disadvantages. The data's properties and the problem's nature determine which classifier should be used. For example, while some classifiers perform better when given high-dimensional data, others perform better with fewer features. Furthermore, although some classifiers are more susceptible to noise or outliers in the data, others are more resistant. This study selects the following Supervised learning classifiers to detect the DoS attack in IoT Environments because they have been extensively utilized in previous research on the detection of DoS attacks.

- XGBoost

XGBoost stands for Extreme Gradient Boosting, extending a version of gradient boosting [25]. It is solid and compelling, handling big datasets and producing reliable predictions. The XGBoost operates by creating a model made up of several decision trees. It begins by using the input data to train a straightforward decision tree and iteratively adds new decision trees to the model while fixing the flaws in the earlier trees. A more potent and precise model is produced due to this procedure, which is referred to as boosting. Additionally, the XGBoost provides many features, including support for missing values, management of categorical variables, and handling of imbalanced data. Additionally, it has built-in regularization to avoid overfitting and supports parallel processing to quicken the training process.

- DT

DT is a supervised ML classifier that can be used to solve classification issues [26]. It builds a tree-like model of choices and their potential effects, with each internal node standing in for a feature or attribute and each leaf node for a class label. The method recursively separates the data according to the feature values starting at the root node until it reaches a leaf node. This leaf node's class label is then used as the anticipated class for the incoming data. Decision trees are straightforward to use, easy to grasp, and capable of handling category and numerical data. But if the tree is too deep or complicated, it could be prone to overfitting.

- RF

An RF is a kind of ensemble learning classifier for classification and regression that builds several DTs during training. It produces the class that represents the mean of the classes (classification) or mean prediction (regression) of the individual trees [26]. The model performs better overall and has less overfitting when numerous decision trees are created, as opposed to depending just on one. The name's "random" component alludes to the randomly selected subsets of data that were used to train each DT.

- LR

LR attempts to calculate the chance of a particular outcome given a specific input variable. This outcome is generally binary, meaning it is composed of two possible values, such as true or false, yes or no, and so on. Multinomial logistic regression can be used to address regression with more than two possible outcomes. Logistic regression is instrumental in discovering which group a novel sample is most like. Furthermore, it is beneficial in cyber security since most security issues are categorization problems, such as recognizing attacks.

- SVM

SVM, a supervised learning technique, can be applied to classification and regression problems [27]. The fundamental

goal of SVM is to identify the optimal boundary (or hyperplane) for classifying the data into multiple groups. A boundary's margin, or the distance between it and the nearest data points from each class, should be maximized to be considered the optimal boundary. Support vectors refer to these nearby data points. By determining which side of the border additional data points fall on once the boundary has been identified, it is simple to classify them. By translating the data into a higher dimension where it may be linearly separable, SVM can also be utilized for data that cannot be separated linearly.

D. DoS Attack

A DoS is a type of cyberattack [28] in which the attacker tries to prevent the targeted users from using a computer resource by flooding it with a lot of traffic or requests. For example, a program, network, or website may be the target of a DoS attack to deny access to legitimate users. DoS attacks come in a variety of forms, including:

- Flooding attacks, such as network flood attacks that saturate networks with large packets, overload a targeted resource with traffic.
- Amplification attacks, increase the amount of traffic directed at a targeted resource by making use of a flaw in it, such as a Domain Name System (DNS) amplification attack that makes use of a DNS server to increase traffic to a targeted resource.
- Application-layer attacks, such as a Hypertext Transfer Protocol (HTTP) request flood attack that floods a targeted website with HTTP requests, target certain flaws in an application.

Using firewalls, IDS, machine learning, and other security measures can assist in detecting and preventing DoS attacks. The following section will review research that researchers did to develop IDS that could detect DoS attacks in the IoT using machine learning.

III. RELATED WORKS

This section investigates previous studies that are related to the works of this paper. Much research has been done using ML to detect DoS attacks in IoT environments. Here are a few of them that will be presented.

In a study by Shreekhand and Deepak [29], in this study, they proposed and implemented ML and neural network models based on Multilayer Perceptron (MLP) and RF for the detection of DoS attacks. The proposed model successfully identified application-layer DoS attacks. The findings indicate that, compared to the MLP algorithm, which offers an accuracy of 98.87%, the RF algorithm provides a better accuracy of 99.95%. The proposed model was examined with the CICIDS2017 dataset.

A study by Yasin. et al. [30], in this study, various ML classifiers have been examined for identifying various DDoS attacks. k-Nearest Neighbors (KNN), LR, MLP, naïve Bayes, SVM, RF, deep autoencoder, CatBoost, Stacking, and XGBoost classifiers are among the ML classifiers mentioned above. The most accurate classifiers are stacking random forest and catBoost. In addition, the proposed model has been examined with the Labris and Digiturk datasets.

A study by Naeem et al. [31], in this study, a model for detecting DoS attacks during Message Queuing Telemetry Transport (MQTT) attacks in the IoT is proposed. Averaged one-dependence estimators (AODE), C4.5, and MLP ML classifiers were used to validate the proposed model. The results show that the AODE classifier achieved the best classification accuracy in identifying the DOS attack.

Another study by Jiyeon Kim et al. [32], in this study, propose a Convolutional Neural Network (CNN) based algorithm for identifying DoS attacks by taking into account the size of the kernel and the number of convolutional layers and then comparing the proposed model with Recurrent Neural Networks (RNN). The proposed model has been examined with KDD and CSE-CIC-IDS 2018 datasets. The obtained results show that the CNN outperformed the RNN.

A study by Rios. In this study, et al. [33] proposes a model for detecting DOS attacks in communication networks based on a Broad Learning System (BLS) that produces high results with less time training. The proposed model has been examined with CICIDS2017 and CSE-CIC-IDS 2018 datasets. The results show that Non-incremental BLS frequently produced the highest accuracy and F-Score, although BLS with incremental learning typically required less training time.

In another study by Verma and Ranga [34], in this study, an extensive investigation into anomaly-based IDS for protecting IoT from DoS attacks is conducted for seven ML classifiers. These classifiers include the RF, Adaboost, Gradient Boosted Machine, XGBoost, Extremely Randomized Trees, classification and regression trees, and MLP. The investigation model has been examined with frequently used datasets, including CIDDS-001, UNSWNB15, and NSL-KDD. The statistical analysis of performance metrics is conducted using Friedman and Nemenyi post-host tests to identify significant differences between classifiers. The results show that classification, regression trees, and the XGboost classifier have the best response time.

In a study by Muhammad Zeeshan et al. [35], in this study, protocol-based deep intrusion detection is proposed for IoT networks. The proposed protocol aims to find similar features of UNSW-NB15 and the Bot-IoT by comparing them. The outcome of the comparison was extracting 26 features and combining normal packets from UNSW-NB15 and DoS/DDoS from Bot-IoT. Furthermore, the proposed protocol was trained using the Long short-term memory networks (LSTM) deep learning technique, and The results show that classification accuracy was 96.3%.

In a study by Alaeddine Mihoub et al. [36], in this study, a model is proposed for the IoT to identify and prevent DoS/DDoS attacks. The identification part of the proposed model is based on the multi-class classifier that adopts the "Looking-Back" concept. Used ML classifiers in the model are DT, RF, KNN, MLP, LSTM, and RNN. The model was tested and evaluated using the Bot-IoT dataset. The obtained results show that the Looking-Back-enabled RF classifier achieves the best accuracy.

In another study by Alimi et al. [37], in this study, a model is proposed for detecting DoS attacks in IoT based on a redefined LSTM deep learning approach. The model was tested and evaluated using NSL-KDD and CICIDS-2017 datasets. The conducted results show that the proposed model has a detection accuracy of 99.22% for DoS attacks on the CICIDS-2017. In comparison, the NSL-KDD dataset attained 98.60%.

In the study by Kimmi Kumari and M. Mrunalini [38], in this study, mathematical and ML model models have been proposed for DoS attack detection using Logistic Regression and Naive Bayes. The proposed models have been tested and evaluated using the CAIDA dataset 2007. The obtained results show the mathematical model is 99.75% accurate, while the ML model is 100% accurate.

IV. Proposed Model

Detailed step-by-step instructions of the proposed model and an explanation are provided in this section. Fig. 3 shows the proposed model flowchart.

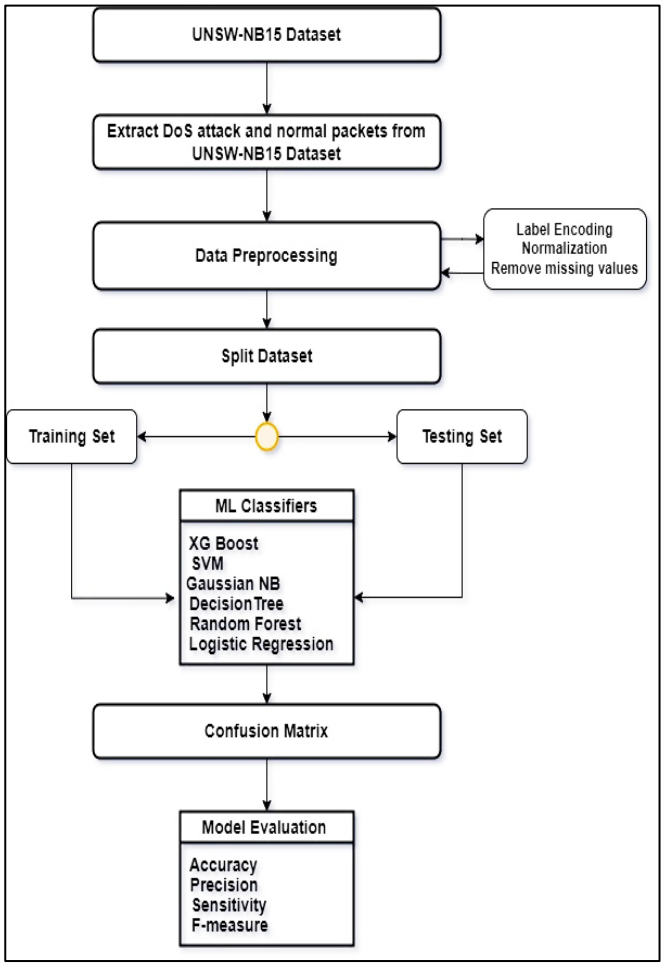

Fig. 3. Proposed model flowchart.

A. UNSW-NB15 Dataset

The University of New South Wales in Sydney, Australia, developed the network intrusion detection dataset known as UNSW-NB15 [39]. It is made to aid intrusion detection research and serve as a realistic testbed for evaluating IDS. It is frequently used for research and development of IDS because it comprises many actual network attacks as well as normal traffic. As a result, the dataset is commonly used in academic research projects and publications by the cybersecurity research community. The UNSW-NB15 dataset includes nine different kinds of network attacks, including worm attacks, backdoor attacks, DoS attacks, exploits attacks, fuzzers attacks, generic attacks, reconnaissance attacks, shellcode attacks, and generic attacks. Fig. 4 shows the attack distribution of the UNSW-NB15 dataset.

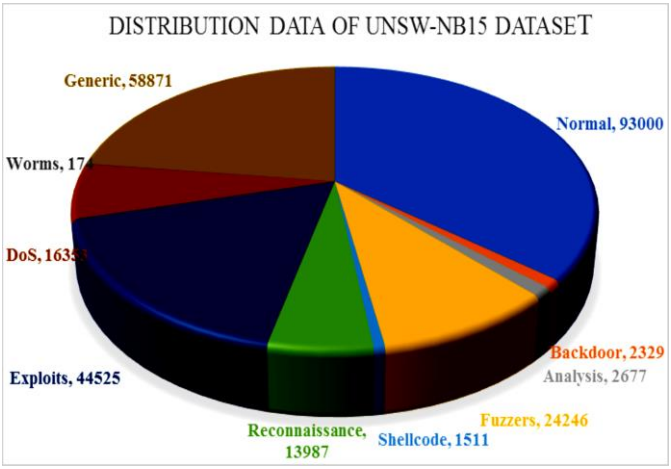

Fig. 4. Attack in the UNSW-NB15.

This paper focuses on the DoS attack; therefore, the DoS attack and normal traffic have been extracted from the UNSW-NB15 dataset.

B. Dataset Preprocessing

The following preprocessing procedures must be done on the UNSW-NB15 dataset before it can be used with the proposed models.

- Label Encoding: In this stage, the category variables are transformed into numerical form so that algorithms can interpret them. This encoding process is accomplished by giving each category in the dataset an individual number. This is helpful since many ML algorithms perform better with numerical data than category data. Reduced data dimensions and enhanced model performance are two benefits of label encoding. The Label Encoding is done using The Scikit-Learn Library's preprocessing module in Python.
- Normalization: Normalization is a technique used in dataset processing by transforming and scaling data to fit inside a predetermined range. The purpose of doing this is often to lessen the influence of outliers and guarantee that the data falls within a close range. The Min-Max normalizing method is a popular approach where the data is often between 0 and 1. The following formula can be used to normalize a value x using the Min-Max normalization:

$$\text{Min/Max normalization} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

- Remove the missing: In a dataset, removing the missing value is a part of the cleaning process that eliminates errors, inconsistencies, and unwanted information. This process ensures that the data is accurate, consistent, and prepared for analysis easier. There are several ways to remove missing values (NaN values) from a dataset using Python. The dropna() method from the Pandas library and the fillna() method to replace a given value for missing values were used.

C. Dataset Splits

Dataset splitting is breaking up a large dataset into more manageable chunks for uses like model testing and training. A dataset is frequently divided into Training sets: this is the primary dataset used to train a model. It is utilized to discover the underlying patterns in the information and contains a sizable amount of data. Testing set: The test set is used to evaluate how well the trained model has worked in practice. It contains data, the model hasn't seen before and estimates how well it performs on actual data.

The split ratios for training and test sets vary depending on the size and complexity of the dataset, but a typical split ratio is 70–30. (Training - Test). Therefore, ensuring the split is accurate and the various subgroups do not overlap is crucial. Python's most used data split method is train_test_split from the scikit-learn library.

D. ML Classifiers

An ML classifier is used to categorize incoming data as either abnormal or normal. The XGBoost, DT, RF, NB, LR, and SVM classifiers are discussed in detail in Section II(C). Because these are the most well-known classifiers used in the literature for IDS, these classifiers were chosen.

E. Confusion Matrix and Model Evaluations

In machine learning, a confusion matrix is a table that assesses how well a classification model performs. It lists the model's accurate and inaccurate predictions based on a test data sample. True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) are the confusion matrix's four main components. These components are used to produce some metrics, including F1-score, recall/sensitivity, accuracy, and precision, which provide a comprehensive picture of the model's performance.

$$\text{Accuracy}=(TP+TN)/(TP + TN +FP +FN \quad (2)$$

$$\text{Precision}= TP /(TP + FP) \quad (3)$$

$$\text{Recall}= TP /(TP + FN) \quad (4)$$

$$\text{F-Measure} =(2 \times \text{Precision} \times \text{Recall})/(\text{Precision}+\text{Recall}) \quad (5)$$

V. Experimental Design and Results

This section presents the Experiment Design and Results

A. Experimental Design

The model was evaluated using Windows 7 and an i7 processor running at 3.40 GHz with 6.0 GB of RAM. The experiments were conducted using the open-source Anaconda (spider) for the UNSW-NB15 datasets under both normal and DoS attacks. Scikit-learn tools in Python were used to implement the model. This model uses the classifiers DT, RF, NB, LR, and SVM. The obtained confusion matrix of prediction for several classifiers is shown in Fig. 5.

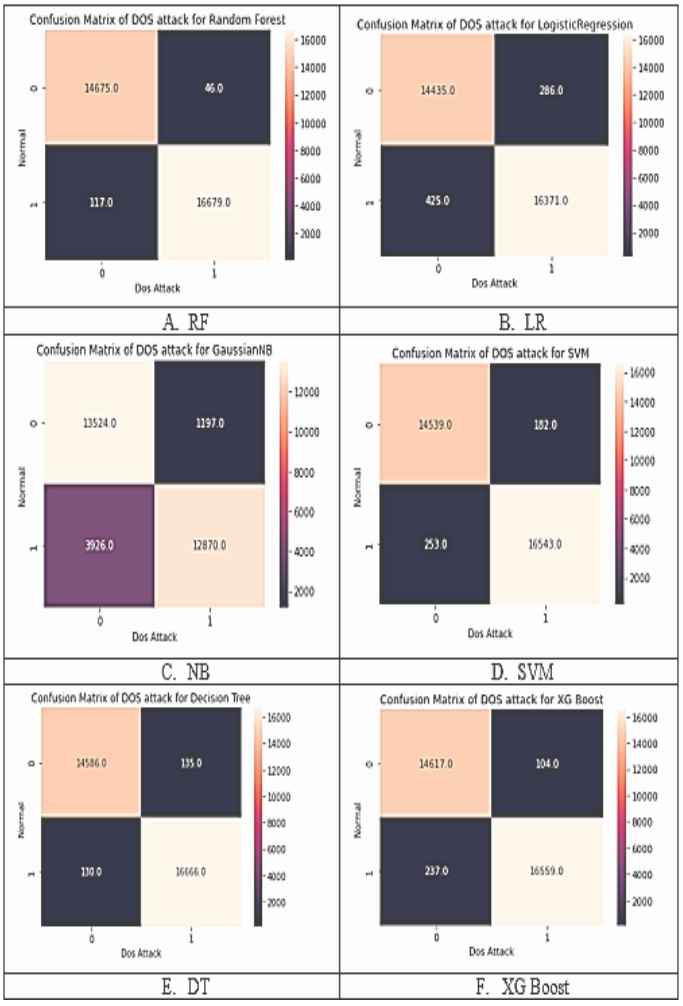

Fig. 5. Confusion matrix of each classifier.

Table I presents the outcomes of the model's assessment metrics by employing the confusion matrix, which was attained, as depicted in Fig. 7.

TABLE I. Performance Metrics Results

	Accuracy	Precision	Sensitivity	F-measure	TP	FN	FP	TN
XG Boost	**98.9**	**98.4**	**99.2**	**98.8**	**99.2**	**0.7**	**1.4**	**98.5**
SVM	**98.6**	**98.2**	**98.7**	**98.5**	**98.7**	**1.2**	**1.5**	**98.4**
NB	**83.7**	**77.5**	**91.8**	**84.0**	**91.8**	**8.1**	**23.3**	**76.6**
LR	**97.7**	**97.1**	**98.0**	**97.6**	**98.0**	**1.9**	**2.5**	**97.4**
RF	**99.4**	**99.2**	**99.6**	**99.4**	**99.6**	**0.3**	**0.7**	**99.3**
DT	**99.1**	**99.1**	**99.0**	**99.1**	**99.0**	**0.9**	**0.7**	**99.2**

B. Finding

The obtained results of this study are analyzed in this section. As seen in Fig. 6, the accuracy of the various ML classifiers is demonstrated. Accuracy is the percentage of accurate predictions made by a classifier in comparison to the actual value of the label. The accuracy rate for XGboost was 98.92%, SVM 98.62%, NB 83.75%, LR 97.74%, RF 99.48%, and DT 99.16%. According to the data collected, the RF classifier outperformed the other classifiers in terms of accuracy due to the following, the RF comprises Multiple decision trees, therefore, it has less classification error. A measure of precision is a percentage that shows what proportion of the objects the classifier recognized are accurate forecasts. For example, the precision for XGboost, SVM, NB, LR, RF, and DT in Fig. 7 is 98.40%, 98.29%, 77.50%, 97.14%, 99.21%, and 99.12%, respectively. The outcomes show that the RF classifier was more precise than the other classifiers.

Sensitivity, also known as recall or true positive rate, is a commonly used metric in machine learning to evaluate binary classification models. It indicates the number of true positive cases the classifier correctly categorized as positive. Sensitivity values for XGBoost, SVM, NB, LR, RF, and DT as in Fig. 8, 99.29%, 98.76%, 91.87%, 98.06%, 99.69%, and 99.08%, respectively. The obtained Sensitivity values demonstrate the superiority of RF classifiers over other classifiers. A classifier's performance can be assessed using the F-measure since it considers both the precision and sensitivity values. This metric is beneficial when there is an uneven distribution of positive and negative classifications. F-measure values for XGBoost, SVM, NB, LR, RF, and DT as in Fig 9, 98.85%, 98.53%, 84.08%, 97.60%, 99.45%, and 99.10%, respectively. The results show that RF classifiers outperform other classifiers regarding F-measure values.

From the analysis mentioned above, the following conclusions can be drawn:

1) RF is the best classifier to detect the DoS attack in an IoT environment compared to XGBoost, DT, NB, LR, and SVM.

2) Gaussian NB is the worst classifier to detect the DoS attack in an IoT environment.

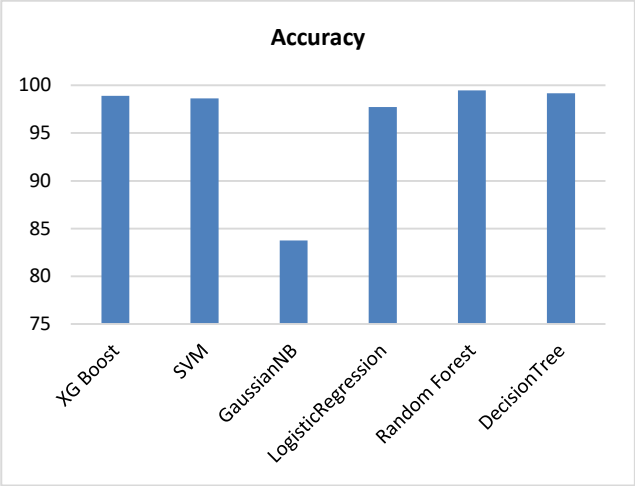

Fig. 6. Accuracy of ML classifiers.

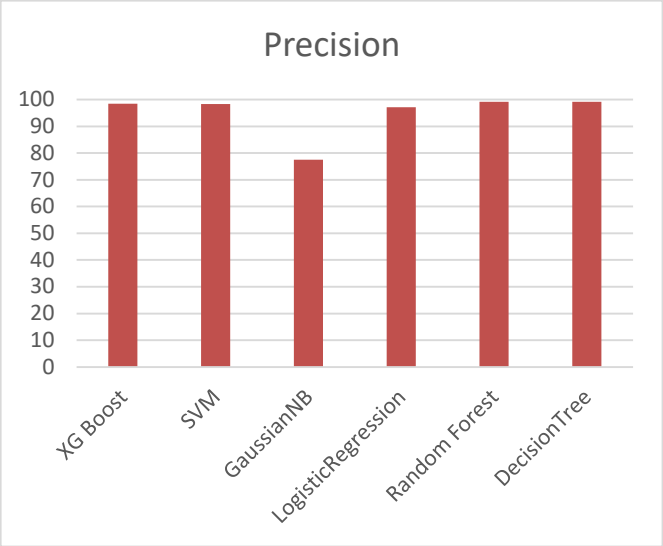

Fig. 7. Precision of ML classifiers.

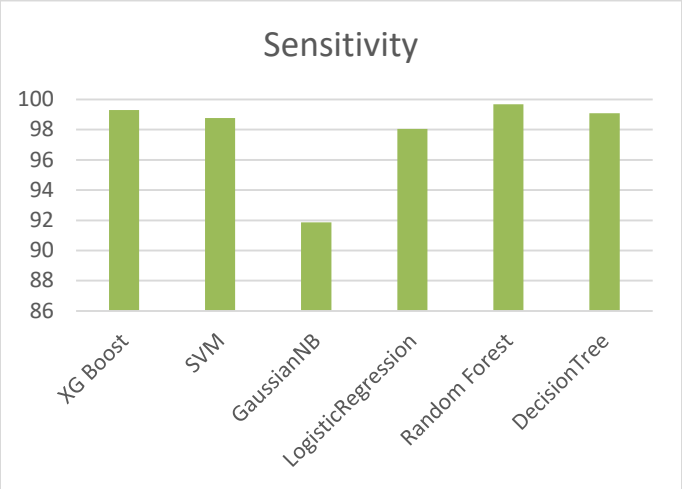

Fig. 8. Sensitivity of ML classifiers.

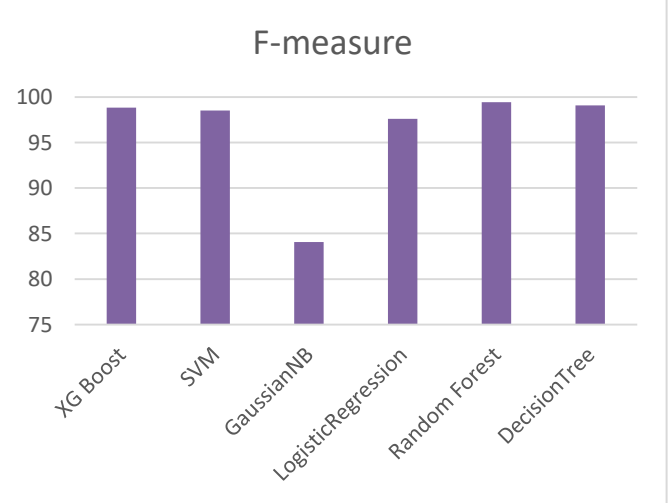

Fig. 9. F-Measure of ML classifiers.

VI. Conclusion and Future Works

Security measures must be included in IoT environments to prevent and combat DoS attacks. Therefore, this paper introduced an IDS model using XGBoost, DT, RF, NB, LR, and SVM to detect DoS attacks in the IoT environment on the UNSW-NB15 datasets. The outcomes of the model were examined using the confusion matrix. Accuracy, precision, sensitivity, and F-Measure were used to evaluate the model's performance. The obtained experimental data proves that the RF is the most efficient classifier among other examined classifiers to detect DoS attacks in IoT environments. Its accuracy was 99.48%, precision 99.21%, sensitivity 99.69%, and F-Measure 99.45%. This study was limited to evaluating some of the ML classifiers with specific attacks. In the future direction of the research, the modern dataset for IDS and other types of DoS attacks, such as (DDoS) attacks, deep learning, and reinforcement learning approaches will be considered to examine the model performance.

References

[1] L. Farhan, S. T. Shukur, A. E. Alissa, M. Alrweg, U. Raza, and R. Kharel, "A survey on the challenges and opportunities of the Internet of Things (IoT)," in 2017 Eleventh International Conference on Sensing Technology (ICST), 2017, pp. 1-5: IEEE.

[2] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning," IEEE Access, vol. 10, pp. 40281-40306, 2022.

[3] O. Almomani, M. A. Almaiah, M. MADI, A. Alsaaidah, M. A. Almomani, and S. Smadi, "Reconnaissance attack detection via boosting machine learning classifiers," in AIP Conference Proceedings, 2023, vol. 2979, no. 1: AIP Publishing.

[4] A. Almomani et al., "Ensemble-Based Approach for Efficient Intrusion Detection in Network Traffic," vol. 37, no. 2, 2023.

[5] G. Carl, G. Kesidis, R. R. Brooks, and S. Rai, "Denial-of-service attack-detection techniques," IEEE Internet computing, vol. 10, no. 1, pp. 82-89, 2006.

[6] R. Vishwakarma and A. K. J. T. s. Jain, "A survey of DDoS attacking techniques and defence mechanisms in the IoT network," vol. 73, no. 1, pp. 3-25, 2020.

[7] P. Kumari, A. K. J. C. Jain, and Security, "A Comprehensive Study of DDoS Attacks over IoT Network and Their Countermeasures," p. 103096, 2023.

[8] X. Zhu and H. Deng, "A security situation awareness approach for iot software chain based on markov game model," 2022.

[9] O. Almomani, M. A. Almaiah, A. Alsaaidah, S. Smadi, A. H. Mohammad, and A. Althunibat, "Machine learning classifiers for network intrusion detection system: comparative study," in 2021 International Conference on Information Technology (ICIT), 2021, pp. 440-445: IEEE.

[10] A. H. Mohammad, T. Alwada'n, O. Almomani, S. Smadi, N. ElOmari, and Continua, "Bio-inspired Hybrid Feature Selection Model for Intrusion Detection," Computers, Materials, vol. 73, no. 1, pp. 133-150, 2022.

[11] A. Alzaqebah, I. Aljarah, O. Al-Kadi, and R. Damaševičius, "A modified grey wolf optimization algorithm for an intrusion detection system," Mathematics, vol. 10, no. 6, p. 999, 2022.

[12] O. Almomani, "A hybrid model using bio-inspired metaheuristic algorithms for network intrusion detection system," Comput. Mater. Contin, vol. 68, no. 1, pp. 409-429, 2021.

[13] O. Almomani, "A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms," Symmetry, vol. 12, no. 6, p. 1046, 2020.

[14] P. Gokhale, O. Bhat, S. Bhat, and Technology, "Introduction to IOT," International Advanced Research Journal in Science, Engineering, vol. 5, no. 1, pp. 41-44, 2018.

[15] P. Matta, B. Pant, and Technology, "Internet of things: Genesis, challenges and applications," Journal of Engineering Science, vol. 14, no. 3, pp. 1717-1750, 2019.

[16] A. Al Zaqebah, O. Almomani, M. Almomani, A. Alsaaidah, A. A. Abu-Shareha, and A. Althunibat, "Improving Routing Decision Algorithm for RPL Networks," in 2023 International Conference on Information Technology (ICIT), 2023, pp. 544-549: IEEE.

[17] M. Almaiah, F. Hajjej, A. Ali, M. Pasha, and O. Almomani, "An AI-Enabled Hybrid Lightweight Authentication Model for Digital Healthcare Using Industrial Internet of Things Cyber-Physical Systems," Sensors, vol. 22, p. 1448, 2022.

[18] R. Masadeh, B. AlSaaidah, E. Masadeh, M. d. R. Al-Hadidi, and O. Almomani, "Elastic Hop Count Trickle Timer Algorithm in Internet of Things," Sustainability, vol. 14, no. 19, p. 12417, 2022.

[19] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on IoT security: application areas, security threats, and solution architectures," IEEE Access, vol. 7, pp. 82721-82743, 2019.

[20] R. Masadeh, O. Almomani, E. Masadeh, and R. e. Masa'deh, "Secure CoAP Application Layer Protocol for the Internet of Things Using Hermitian Curves," in The Effect of Information Technology on Business and Marketing Intelligence Systems: Springer, 2023, pp. 1869-1884.

[21] H. F. Atlam, G. B. Wills, and s. cities, "IoT security, privacy, safety and ethics," Digital twin technologies, pp. 123-149, 2020.

[22] D. Mudzingwa and R. Agrawal, "A study of methodologies used in intrusion detection and prevention systems (IDPS)," in 2012 Proceedings of IEEE Southeastcon, 2012, pp. 1-6: IEEE.

[23] A. Haldorai, A. Ramu, and M. Suriya, "Organization internet of things (IoTs): supervised, unsupervised, and reinforcement learning," in Business Intelligence for Enterprise Internet of Things: Springer, 2020, pp. 27-53.

[24] A. Sholiyi, J. A. Alzubi, O. A. Alzubi, O. Almomani, and T. O'Farrell, "Near capacity irregular turbo code," arXiv preprint arXiv:.01358, 2016.

[25] S. Smadia, O. Almomanib, A. Mohammadc, M. Alauthmand, and A. Saaidahe, "VPN Encrypted Traffic classification using XGBoost," International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 7, 2021.

[26] M. Madi, F. Jarghon, Y. Fazea, O. Almomani, A. Saaidah, and C. Sciences, "Comparative analysis of classification techniques for network fault management," Turkish Journal of Electrical Engineering, vol. 28, no. 3, pp. 1442-1457, 2020.

[27] M. A. Almaiah et al., "Performance Investigation of Principal Component Analysis for Intrusion Detection System Using Different Support Vector Machine Kernels," Electronics, vol. 11, no. 21, p. 3571, 2022.

[28] S. SMADI, M. ALAUTHMAN, O. ALMOMANI, A. SAAIDAH, and F. ALZOBI, "Application layer denial of services attack detection based on stacknet," International Journal of Advanced Trends in Computer Science and Engineering, vol. 3929, no. 3936, pp. 2278-3091, 2020.

[29] S. Wankhede and D. Kshirsagar, "DoS attack detection using machine learning and neural network," in Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5: IEEE.

[30] Y. Gormez, Z. Aydin, R. Karademir, and V. C. Gungor, "A deep learning approach with Bayesian optimization and ensemble classifiers for detecting denial of service attacks," International Journal of Communication Systems, vol. 33, no. 11, p. e4401, 2020.

[31] N. F. Syed, Z. Baig, A. Ibrahim, and C. Valli, "Denial of service attack detection through machine learning for the IoT," Journal of Information Telecommunication, vol. 4, no. 4, pp. 482-503, 2020.

[32] J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," Electronics, vol. 9, no. 6, p. 916, 2020.

[33] A. L. G. Rios, Z. Li, K. Bekshentayeva, and L. Trajković, "Detection of denial of service attacks in communication networks," in IEEE

international symposium on circuits and systems (ISCAS), 2020, pp. 1-5: IEEE.

[34] A. Verma and V. Ranga, "Machine learning based intrusion detection systems for IoT applications," Wireless Personal Communications, vol. 111, pp. 2287-2310, 2020.

[35] M. Zeeshan et al., "Protocol-based deep intrusion detection for dos and ddos attacks using unsw-nb15 and bot-iot data-sets," IEEE Access, vol. 10, pp. 2269-2283, 2021.

[36] A. Mihoub, O. B. Fredj, O. Cheikhrouhou, A. Derhab, and M. Krichen, "Denial of service attack detection and mitigation for internet of things using looking-back-enabled machine learning techniques," Computers Electrical Engineering, vol. 98, p. 107716, 2022.

[37] K. O. Adefemi Alimi, K. Ouahada, A. M. Abu-Mahfouz, S. Rimer, and O. A. Alimi, "Refined LSTM Based Intrusion Detection for Denial-of-Service Attack in Internet of Things," Journal of Sensor Actuator Networks, vol. 11, no. 3, p. 32, 2022.

[38] K. Kumari and M. Mrunalini, "Detecting Denial of Service attacks using machine learning algorithms," Journal of Big Data, vol. 9, no. 1, pp. 1-17, 2022.

[39] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in 2015 military communications and information systems conference (MilCIS), 2015, pp. 1-6: IEEE.