MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

Enhancing Cybersecurity: Machine Learning Approaches for Predicting DDoS Attack

Farhan Sadik Ferdous[1], Tapu Biswas [1], and Akinul Islam Jony*[1]

[1] *Department of Computer Science, American International University-Bangladesh. Dhaka, Bangladesh.*

KEYWORDS

Attack
Cyber Security
DDoS
Machine Learning
CIC-DDoS2019 Dataset

ARTICLE HISTORY

Received 3 April 2024
Received in revised form 14 June 2024
Accepted 22 June 2024
Available online 4 July 2024

ABSTRACT

Dealing with network security has always been challenging, particularly with regard to the detection and prevention of Distributed Denial of Service (DDoS) attacks. Attacks like DDoS bring threats to the network by violating its availability to the probable people who are in need of using that particular server. It is a type of cyber-attack where a network is flooded with a huge amount of traffic, overwhelming the system, and making it unavailable. This type of attack focuses on making the service unavailable to rightful users, without breaching the security perimeter. In a DDoS attack, a master computer hacks a network of vulnerable computers to send a huge quantity of packets to a server from already captured zombie computers. Researchers have suggested various Machine learning (ML) algorithms to detect such attacks. To study and analyse DDoS attacks, researchers have used the CIC-DDoS2019 dataset. To find out how often a DDoS attack happens to a server along with the possible pattern of the attack and type of the attack. This dataset is utilized to train and evaluate ML models for detecting DDoS attacks. In this paper, the primary objective is to propose a decent version of DDoS dataset for investigation and evaluate the performance of various state-of-the-art classifiers, such as Gaussian Naïve Byes (GNB), Bernoulli Naïve Byes (BNB), Random Forest (RF), ID3 Decision Tree (ID3 DT), Logistic Regression (LR), K-Nearest Neighbors (KNN), AdaBoost, CART, and Bagging Classifier ML algorithms to detect DDoS attacks accurately. Along with that, the experimenter showed that DDoS attacks can be identified even more accurately if the attacks are stored in a binary way rather than categorized into 13 different types of attacks in the dataset.

1. INTRODUCTION

The Internet is crucial in the current era as it acts as a global information source for all users, making its availability imperative. The internet is vast, providing access to information, services, and resources for all sectors. In today's digital era, information security must be given top priority as everything is connected to the internet. To protect sensitive data and personal information from cyber threats and malicious attacks, it is important to take the necessary steps to ensure robust and reliable security protocols are in place. As its demand grows, security issues arise. There are various types of attacks targeting the internet that need to be recognized, classified, and protected against [1]. Among them, one of the most common attacks is DDoS in today's cyber world [2]. Although the Morris Worm is the first recorded instance of a DDoS attack, the attack was an unintentional consequence of the worm [2]. In August 2016, a botnet of over 24,000 computers in 30 countries launched a DDoS attack [3]. That's why this paper is motivated to showcase the prevention of DDoS attack incidents from happening in the future.

DDoS attack aims to exhaust computing resources, preventing normal work from proceeding. Unlike Denial of Service (DoS) attacks, which don't attempt to destroy or corrupt data, DDoS attacks are modified to include multiple sources attacking the targeted system simultaneously [4]. DDoS attacks require a controlling master computer, a target, and intermediary computers used to generate the attack. Firstly, in a DDoS attack, the master computer hacks a network of computers ("zombies") and uses them to run DoS programs, making the attack difficult to detect and defend against [5]. DoS and DDoS attacks try to render resources of a network and make them unavailable to its targeted users [6].

*Corresponding author:
E-mail address: Akinul Islam Jony < akinul@aiub.edu >.
https://doi.org/10.56532/mjsat.v4i3.306
2785-8901/

This type of attack can cause significant financial and resource damage to a company. DDoS attacks have gradually become common in the present time where so much of the whole infrastructure and services rely on the internet.

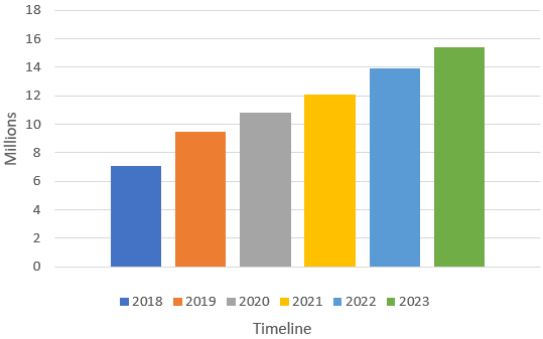

Fig. 1. DDoS Attack from 2018-2023 [7]

ML is a sector of computer science that makes machines predict from data byself without being explicitly programmed [8]. This is achieved through the use of algorithms that iteratively adjust and improve their performance as they process more data. By building a model that accurately represents a selected dataset, ML can effectively solve problems [9]. The study used nine different supervised ML models, including GNB, BNB, RF, ID3 DT, LR, KNN, AdaBoost, CART, and Bagging classifier. Every one of these algorithms considers various features of the dataset and gives correct classifications of the data. The primary focus of this study is to present a decent DDOS dataset (modified version of CIC-DDoS2019 Dataset) and find the best ensemble framework of a supervised model that can help identify and prevent DDoS scenarios.

2. LITERATURE REVIEW

This part covers a brief analysis of all the DDoS attacks and the present research in this sector of work.

DDoS attacks are intended [10] to make the server ineffectual to give general services, resulting in a network failure. It is frustrating for users who rely on these services, but thankfully, some measures can be taken to prevent and mitigate these attacks. Direct DDoS and Reflection-based DDoS are two common types of DDoS attacks [11, 12]. A DDoS is an attack that originated to make networks and systems' resources unavailable for legitimate users [13].

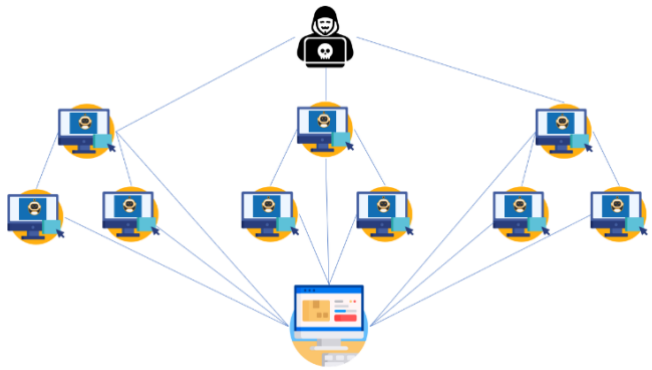

Fig. 2. DDoS Attack

Some of the examples of DDoS attacks are DNS, NTP, SSDP, LDAP, SNMP, TFTP, MSSQL, Portmap, SYN, NetBIOS, UDP, and UDP-Lag. These attacks get activated in many network layers. The hacker can either use a single or many computers as a bot to activate an attack into the network layer. More than one attack also happened on a server at a time. When the attacker uses multiple computers, these computers are called bots or zombies.

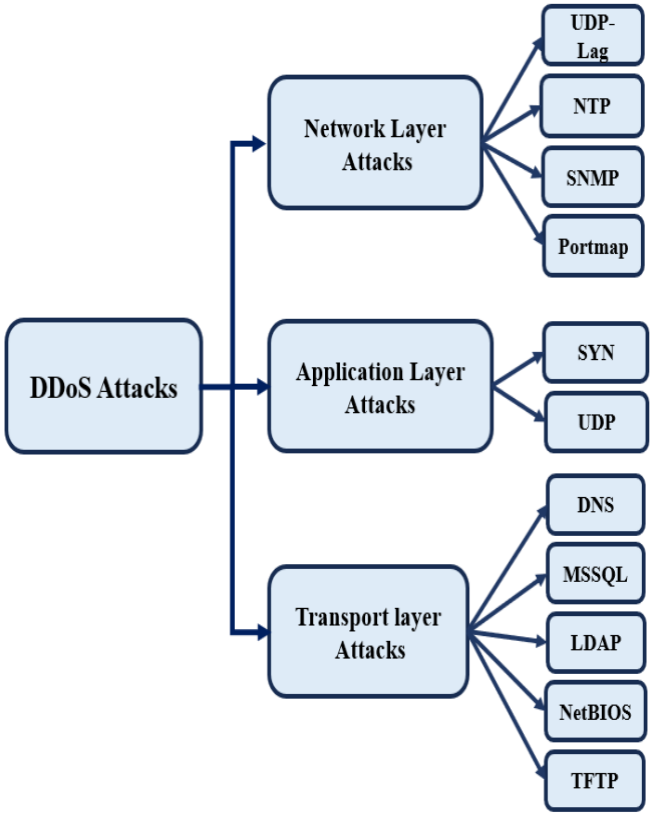

Fig 3. DDoS Attack on Different Network Layer

Hariharan. M et al. [14] describe this paper as aiming to identify DDoS attacks by utilizing the C5.0 ML algorithm. The objective is to assess the results obtained with other classifiers such as NB and C4.5 DT. The study will help determine the effectiveness of the C5.0 algorithm in identifying DDoS attacks and how it compares to other commonly used classifiers.

K. Narasimha Mallikarjunan et al. [15] contributed to creating a real-time dataset by using a Naive Bayes technique for identifying and evaluating its accuracy with the remaining methods like RF and J48.

Iman Sharafaldin et al. [16] analyzed the existing dataset completely and proposed a classification for DDoS attacks. Then present the CICDDoS2019 dataset including 11 DDoS attacks, which remedies all current shortcomings. Recommend an identification and classification way of detection based on a set of features from network flow with a prediction percentage of 78% in ID3, 77% in RF, 41% NB, and 25% LR.

Sagar Dhanraj Pande et al [17] demonstrated the classification of normal and attack samples was performed using the RF algorithm. The accuracy of the classification was 99.76%, which indicates that the algorithm was highly effective in distinguishing between the two types of samples.

Kimmi Kumari and M. Mrunalini [18] used the CAIDA 2007 dataset for expletory research. The implementation of the Naïve Byes (NB) and LR algorithm was done using the Weka data mining platform. The results of this study were carefully analyzed and compared in order to draw valid conclusions with a percentage of 99% in both algorithms.

Marwane Zekri et al. [19] implemented a highly effective system for detecting DDoS that utilizes advanced algorithms such as Naive Bayesian, C4.5, and K-Means to reduce the DDoS threat. This system also incorporates signature detection techniques, which enable it to generate a decision tree (DT) for automatic and accurate identification of signature attacks for DDoS flooding attacks.

Raniyah Wazirali and Rami Ahmad contributed [20] to the evaluation of the use of ML approaches in WSN node traffic and how they impact the overall WSN network lifetime. The author thoroughly analyzes the performance metrics of different ML classifications such as KNN, LR, SVM (Support Vector Machine), DT, NB, Gboost, LSTM (Long Short-Term Memory) (e.g., [24]), and MLP, (Multi-Layer Perceptron) on a WSN-dataset of different sizes. To accurately assess the effectiveness of these algorithms, the author used various performance metrics such as Accuracy, F1-score, FPR (False Positive Ratio), FNR (False Negative Ratio), and training execution time.

Salim Salmi and Lahcen Oughdir demonstrated [21] that several deep learning (DL) algorithms like DNN, CNN, RNN, CNN+RNN based IDS (were trained on WSN-DS dataset) for identifying 4 types of DoS attacks (Blackhole, Grayhole, Flooding, and Scheduling) that affect WSNs.

Mohamed Idhammad et al. [10] present an innovative approach for detecting DDoS attacks online using network Entropy estimation, Co-clustering, Information Gain Ratio, and Extra-Trees technique. They used 3 (three) public datasets such as NSL-KDD, UNB ISCX 12, and UNSW-NB15, and achieved a predictive accuracy of 98.23%, 99.88%, and 93.71% respectively. Moreover, the false positive rates were minimal, with values of 0.33%, 0.35%, and 0.46%, respectively.

Rami J. Alzahrani and Ahmed Alzahrani [22] worked on implementing six ML algorithms that include NB, KNN, DT, SVM, RF, and LR by using WEKA tools to identify DDoS attacks from the same CICDDoS2019 dataset. The most satisfying accuracy result in the introduced determination was obtained by DT and RF methods with 99% accuracy.

The previous research mentioned datasets were unable to execute and capture all the DDoS attacks presented in the dataset CIC-DDoS2019, which doesn't contain enough data to match the credibility of the dataset. The created dataset in this paper was used to obtain accuracy through various ML algorithms, but only random data from every attack was chosen to create an idealized dataset. This research study highlighted the difference in accuracy when the target variables are modified. Also, this paper initiated to highlight the accuracy, recall, precision, and F1 score by various ML techniques using the CIC-DDoS2019 dataset.

3. METHODS AND MATERIALS

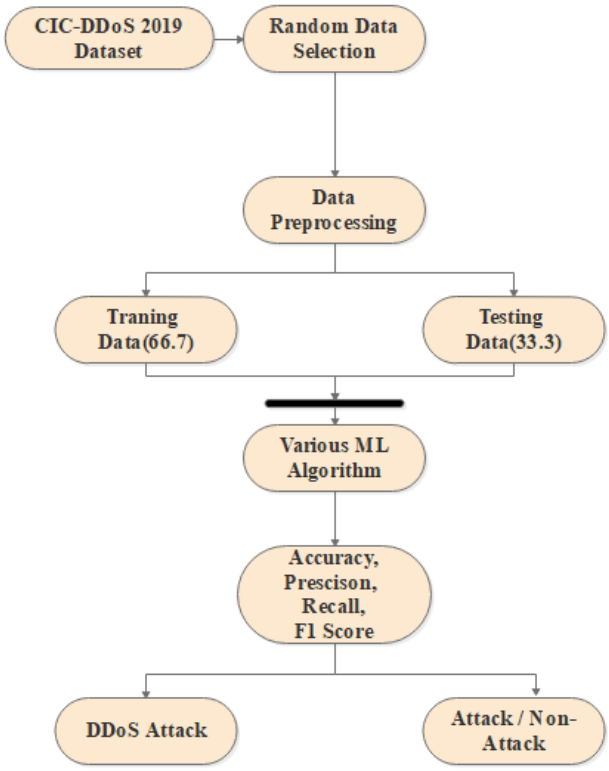

Fig. 4. Working Procedure

3.1 Dataset

The dataset is about DDoS attacks [23]. DDoS attacks are a significant threat to network security. The attacks overwhelm the target network with malicious traffic, making it inaccessible to users. The dataset used in this study was obtained from the Canadian Institute for Cybersecurity [23]. The dataset has 87 feature columns and 1 target column, totaling 88 columns. There are 14 different categorical values in the target column, including NTP, DNS, LDAP, MSSQL, NetBIOS, SNMP, SSDP, UDP, UDP-Lag, WebDDoS, SYN, TFTP, Portmap, and BENIGN (represents non-attack). Canadian Institute for Cybersecurity has created two datasets – one is for training and the other one is for testing [16]. The training dataset was created for the day of January 12th, starting at 10:30 AM and ending at 5:15 PM. This dataset includes 12 types of DDoS attacks, such as NTP, DNS, LDAP, MSSQL, NetBIOS, SNMP, SSDP, UDP, UDP-Lag, WebDDoS, SYN and TFTP. The testing dataset was created for the day of March 11th, starting at 9:40 AM and ending at 5:35 PM. This dataset includes 7 types of attacks, such as Portmap, NetBIOS, LDAP, MSSQL, UDP, UDP-Lag, and SYN.

3.2 Dataset Preprocessing

To propose the modified version of the dataset, both the training and testing datasets were used. Here data were randomly selected 50,000 data for each type of attack (NTP, DNS, LDAP, MSSQL, NetBIOS, SNMP, SSDP, UDP, UDP-Lag, SYN, and TFTP) from the training dataset, including all BENIGN data. Additionally, 50,000 Portmap attack data were also randomly collected, including all benign data from the testing dataset. WebDDoS data was removed from the dataset since there was the minimum account of data available for

WebDDoS. After merging all datasets, a new dataset was created that contained 661597 records. The new dataset has 87 feature columns, 1 target column with 13 target variables.

After removing the object datatype column during data cleaning, the dataset was left with a total of 82 columns. After replacing the infinity values with NaN, the corresponding rows are removed from the dataset. The final dataset contains 638455 records. The target variable in the final dataset was renamed from its original names (DrDoS_DNS, DrDoS_LDAP, etc.) to more concise names (DNS, LDAP, etc.).

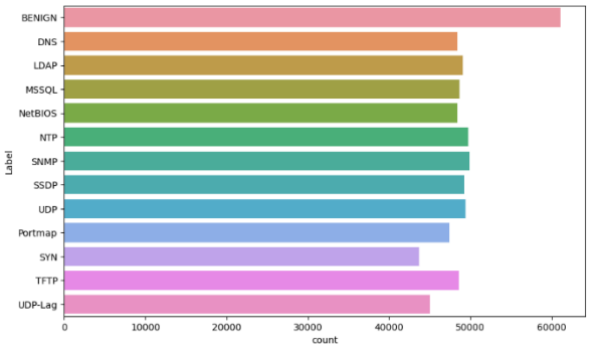

Fig. 5. Target Matrix Visualization (13 types of attacks)

After that, a new target column was created named "Class". Here The attacks and non-attacks are identified as "ATTACK" and "BENIGN".

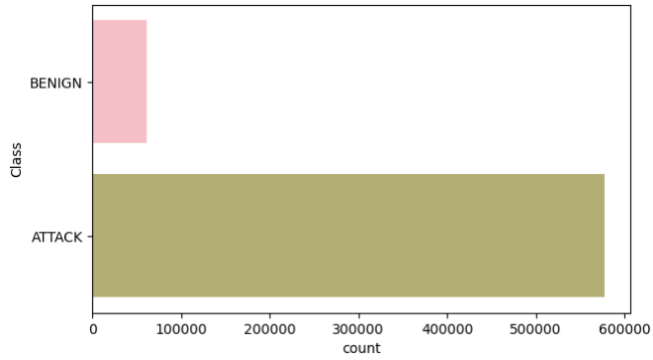

Fig. 6. Target Matrix Visualization (Attack/Non-attack)

3.3 Machine Learning Algorithms

3.3.1 Naïve Byes (NB)

NB is a powerful tool [15] in machine learning that allows to classify data based on assigned class labels. It is based on the Bayes theorem [25] and uses conditional probabilities to make predictions. The main assumption of Naive Bayes [22] is that all attributes are independent of each other. When this assumption is met, Naive Bayes classifiers can outperform other models like LR while using less training data.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (1)$$

3.3.2 Random Forest (RF)

RF is a well-known [17] ML technique that is commonly used for classification which was established by Leo Breiman. RF combines [16] the concepts of DT and ensemble learning. It uses a forest of several DTs, each using randomly selected data attributes as its input. To prevent the DT [25] from being identical, in an RF, a subset of characteristics is randomly selected for each node. The rest of the parameters are then used for the DT within the forest.

$$P(Y = 1|X) = \frac{1}{1 + e^{(-(b0+b1\times X1+b2\times X2+\cdots+bn\times Xn))}} \quad (2)$$

3.3.3 Decision Tree (DT)

DT will be immensely effectual [15] in terms of modeling and extraction from vast amounts of data by generating a set of decision rules [25]. The leaves of the tree present classes, while every child node and its branches present a composition of attributes that result in classification. To classify an object, the process starts with the root node and then moves down the corresponding branches until reaching the leaf node.

3.3.4 Logistic Regression (LR)

LR [16] is a classification method that can be used to solve multiclass problems. LR is similar to other regression analyses in that it can be used to represent data and establish the relationship between features (attributes) and target variables (classes).

3.3.5 K Nearest Neighbor (KNN)

Another supervised ML technique is KNN which classifies a new object based on its nearest neighbors, with the value of k being a positive integer that is defined beforehand. KNN has proven to be extremely effective [19], particularly in classification problems across various domains. Essentially, KNN will locate all points that are less distant from the unknown data, and then select those that are closest to it. As a result, it is sometimes referred to as a distance-based algorithm. It can be time-consuming to train KNN classifiers when values are not readily available, and KNNs can also be quite expensive in terms of storage time due to the vast data involved.

3.3.6 Adaboost

The AdaBoost algorithm [26] is a powerful technique proposed by Freund and Schapire for creating a strong classifier by linking multiple weak classifiers together in a series. The key to its success is the ability to identify and focus on difficult-to-classify samples, assigning higher weights to misclassified points as the process continues. As a result, subsequent weak classifiers are better equipped to handle these challenging samples. By combining the predictions of all the weak classifiers final prediction is made, with each classifier's contribution weighted based on its individual performance. Overall, the AdaBoost algorithm is an effective tool for improving classification accuracy and producing robust classifiers.

3.3.7 Cart

CART measures [27] The system analyses the amount of impurity in the data provided and creates a binary tree structure. Every internal node of the tree generates a decision based on a specific attribute, resulting in exactly two possible

classes being outputted. CART calculates the Gini index for each attribute and selects the one with the lowest value. By selecting the attribute with the lowest Gini index recursively, the tree is built. It's an interesting approach that seems to be effective in certain scenarios.

3.3.8 Bagging

One of the most well-known and frequently used ensemble techniques for data classification and prediction is Bootstrap Aggregation [27]. By joining the outcomes of every classifier to create a final output class, an improved combined classifier is created. Studies have shown that this bagging approach can show better results than individual classifiers built from the real training data, as it can remove the inconsistency of single inducers. To combine the results of each classifier, a voting approach is used. With the measure same probability each classifier selects each instance, unlike in boosting where it depends on its weight.

3.4 Evaluation Metrics

Four performance metrics precision, recall, accuracy, and F1-measure are used to assess the dataset prediction.

Accuracy is the proportion of accurately categorized DDoS attack instances or benign instances out of all instances in the dataset. It identifies the ratio of accurate predictions corresponding to every sample [21, 24].

$$Accuracy = \frac{TP}{TP + FN + TN + FP} \quad (3)$$

Precision is the proportion of identifying correctly detected attacks to every packet categorized as attacks. [16, 21].

$$Precison = \frac{TP}{TP + FP} \quad (4)$$

Recall is the ratio of the capacity to accurately classify an attack upon all generated flows [16, 21].

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

F-Measure is an integration mean of the precision and recall into a single measure [16, 21, 24].

$$F - Measure = \frac{TN + TP}{2TP + FN + FP} \quad (6)$$

4. RESULTS AND FINDINGS

The experiments are conducted using the CIC-DDoS2019 dataset [23]. This research paper experimented with nine different machine learning models to predict their accuracy, precision, recall, and F1 score. The models that have been used in this paper include GNB, BNB, RF, ID3 DT, LR, KNN, AdaBoost, CART, and the Bagging classifier.

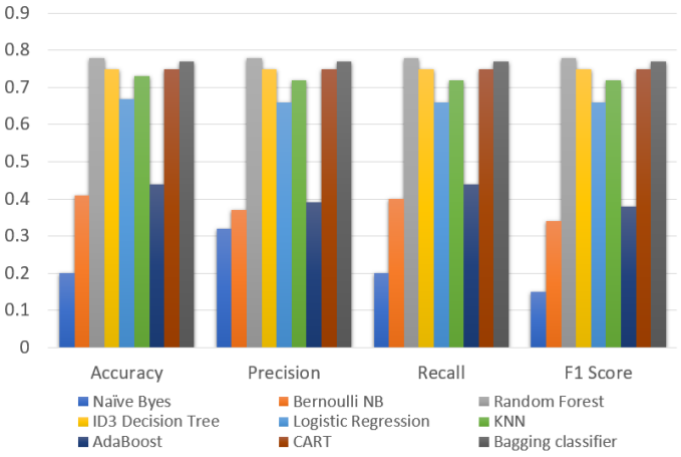

Fig. 7. DDoS Attack Detection Result (13 types of attacks)

In Fig. 7, the accuracy, precision, recall, and F1 scores of nine ML classifiers are shown. The table indicates that the NB has the lowest accuracy, precision, recall, and F1 score which is 20%, 32%, 20%, and 15%. While the RF classifier performs exceptionally well with high accuracy, precision, recall, and F1 score which are 78%, 78%, 78%, and 78%. On the other hand, the performances of ID3 DT, KNN, CART, and Bagging classifiers are very similar to the RF classifier whose accuracy is 73% to 75%. Along with the NB classifier BNB, LR, and AdaBoost were unable to predict a decent amount of attack from the dataset.

Table 1. Performance Metrics for ML Algorithms (13 Types of Attacks)

Prediction Model	Accuracy	Precision	Recall	F1 Score
Gaussian NB	0.20	0.32	0.20	0.15
Bernoulli NB	0.41	0.37	0.40	0.34
Random Forest	0.78	0.78	0.78	0.78
ID3 Decision Tree	0.75	0.75	0.75	0.75
Logistic Regression	0.67	0.66	0.66	0.66
KNN	0.73	0.72	0.72	0.72
AdaBoost	0.44	0.39	0.44	0.38
CART	0.75	0.75	0.75	0.75
Bagging Classifier	0.77	0.77	0.77	0.77

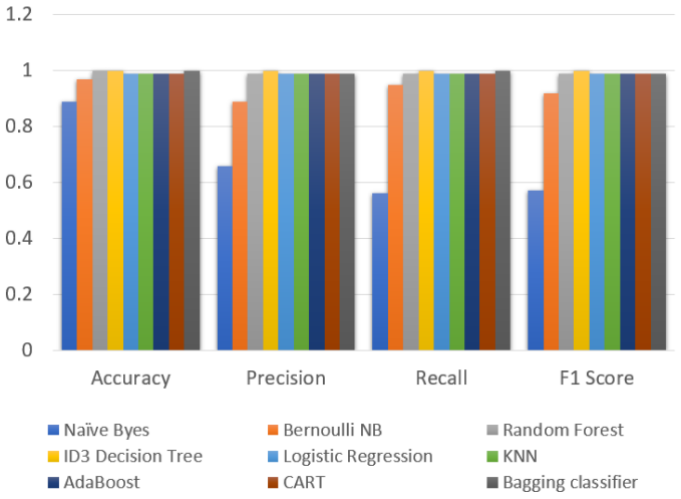

Fig. 8. DDoS Attack Detection Result (Attack/Non-attack)

In Fig. 8, the accuracy, precision, recall, and F1 scores of nine classifiers are displayed. Here the target variable is class where the attacks are categorized as "ATTACK" and non-attacks are categorized as "BENIGN". The table indicates that the Naïve Bayes technique has the lowest accuracy, precision, recall, and F1 score which is almost 90% while the RF, ID3 DT, and Bagging classifier perform exceptionally well with very high accuracy, precision, recall, and F1 score which is 100%. On the other hand, the performances of KNN and CART classifiers perform well too with an accuracy of 99%.

In Table 2, three of the classifiers identified the attacks with almost 100% accuracy with a recall of also nearly 100%. Here we can see the prediction model performs much more efficiently when it has fewer labels to identify in the target. Whenever the attacks are categorized in 13 different names the accuracy seems to be less than it is in 2 categories.

Table 2. Performance Metrics for ML Algorithms (Attack/Non-attack)

Prediction Model	Accuracy	Precision	Recall	F1 Score
Gaussian NB	0.89	0.66	0.56	0.57
Bernoulli NB	0.97	0.89	0.95	0.92
Random Forest	1.00	0.99	0.99	0.99
ID3 Decision Tree	1.00	1.00	1.00	1.00
Logistic Regression	0.99	0.99	0.99	0.99
KNN	0.99	0.99	0.99	0.99
AdaBoost	0.99	0.99	0.99	0.99
CART	0.99	0.99	0.99	0.99
Bagging Classifier	1.00	0.99	1.00	0.99

Here the prediction model performs much more efficiently when it has fewer labels to identify in the target variable. Whenever the attacks are categorized in 13 different names the accuracy seems to be less than when it is in 2 categories. So, in the end, the result showcased that if attacks are stored in the dataset as attack and non-attack it can be identified way more accurately rather than when it is categorized. Also, for identifying the attack RF, ID3 DT, KNN, AdaBoost, CART, and the Bagging are most recommended. Among them, RF and AdaBoost are the best models to predict attacks because they can overcome overfitting and improve performance through ensemble methods.

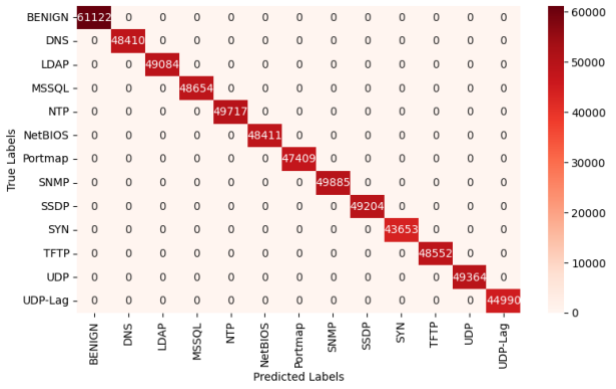

Fig. 10. Confusion Matrix (13 types of attacks)

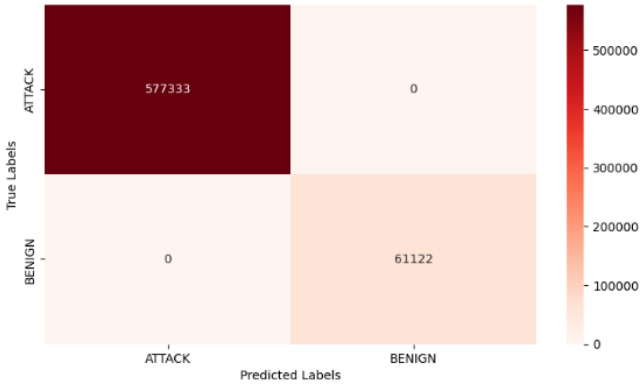

Fig. 11. Confusion Matrix (Attack/Non-attack)

5. CONCLUSION

DoS and DDoS attacks are becoming [28] increasingly sophisticated, making it difficult for real (target) users to access network resources. It's crucial to develop a comprehensive defense strategy to combat these attacks. The research study presented in this paper started by understanding and analyzing the relationship between the instance of the dataset time of the CIC-DDoS2019 Dataset. It also tries to analyze the DDoS attack ways and gives a thorough analysis of how it happens. An observant fact has been identified that requests [18] tend to decrease as the rate at which requests are processed increases. This shortening was chosen carefully using a random function so that there would be no bias in choosing the data. Various ML [20] models have been introduced to identify DDoS attacks. The outcome showcases

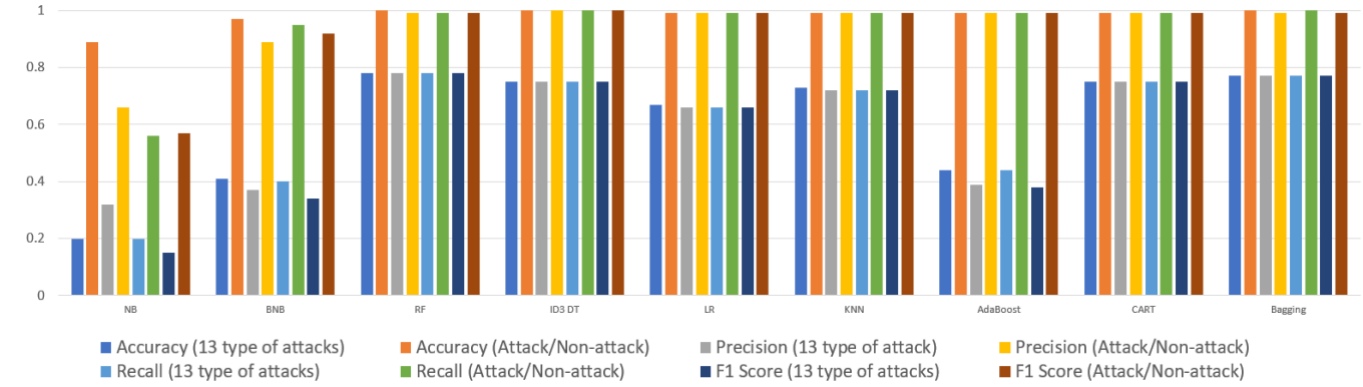

Fig. 9. Comparison Between Table 1 and Table 2

that the proposed way of work is not only feasible but also delivers superior performance as compared to various recent and relevant approaches that have been documented in the literature [29].

This model still operates more as a damage control system [14] rather than a prevention system. It seems that the detection only occurs after the damage has already been done. Also, the predictions are only detected by the ML algorithm. This model doesn't showcase a new way of detecting DDoS attacks.

In the future, a throughout analysis hopes to be conducted by using the DL algorithm. In this research paper it hoped to find improved ways to stop cyber-attacks more effectively, as well as new techniques and efficient algorithms, being developed to stop DDoS attacks. This project aims to investigate the real-time execution and verification of the method to address the problem at hand [30].

REFERENCES

[1] S. Chakraborty, P. Kumar, and B. Sinha, "A study on DDoS attacks, danger and its prevention," Int. J. Res. Anal. Rev., vol. 6, no. 2, pp. 10-15, 2019.

[2] K. H. Zaboon and A. A. Abdullah, "A Review of the Common DDoS Attack: Types and Protection Approaches Based on Artificial Intelligence," Fusion: Practice and Applications, vol. 7, no. 1, pp. 08-08, Dec. 2021.

[3] L. E. Jaramillo, "Malware detection and mitigation techniques: Lessons learned from Mirai DDOS attack," Journal of Information Systems Engineering & Management, vol. 3, no. 3, pp. 19, Jul. 16, 2018.

[4] A. I. Jony and S. A. Hamim, "Navigating the Cyber Threat Landscape: A Comprehensive Analysis of Attacks and Security in the Digital Age", Journal of Information Technology and Cyber Security, vol. 1, no. 2, pp. 53-67, 2023.

[5] I. V. Kotenko and A. V. Ulanov, "Agent-based simulation of DDoS attacks and defense mechanisms," Journal of Computing, vol. 4, no. 2, pp. 16-37, 2005.

[6] Q. Yan, F. R. Yu, Q. Gong, and J. Li, "Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges," IEEE communications surveys & tutorials, vol. 18, no. 1, pp. 602-622, Oct. 5, 2015.

[7] Cisco, "Annual Internet Report (2018–2023) White Paper," Accessed June 11, 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html.

[8] A. I. Jony and A. K. B. Arnob, "Securing the Internet of Things-Evaluating Machine Learning Algorithms for Detecting IoT Cyberattacks using CIC-IoT2023 Dataset", International Journal of Information Technology and Computer Science, 2024. (In Press).

[9] S. S. Shanto, Z. Ahmed and A. I. Jony, "Mining User Opinions: A Balanced Bangla Sentiment Analysis Dataset for E-Commerce", Malaysian Journal of Science and Advanced Technology, vol. 3, no. 4, pp.272-279, 2023.

[10] Z. Chao-Yang, "DOS attack analysis and study of new measures to prevent," in 2011 International Conference on Intelligence Science and Information Engineering, IEEE, Aug. 2011, pp. 426-429.

[11] M. Idhammad, K. Afdel, and M. Belouch, "Semi-supervised machine learning approach for DDoS detection," Applied Intelligence, vol. 48, pp. 3193-3208, Oct. 2018.

[12] D. Tang and X. Kuang, "Distributed denial of service attacks and defense mechanisms," in IOP Conference Series: Materials Science and Engineering, vol. 612, no. 5, p. 052046, Oct. 2019.

[13] N. Tripathi, "DoS and DDoS Attacks: Impact, Analysis and Countermeasures."

[14] M. Hariharan, H.K. Abhishek, and B.G. Prasad, "DDoS attack detection using C5.0 machine learning algorithm," IJ Wireless and Microwave Technologies, vol. 1, pp. 52-59, 2019.

[15] K. Narasimha Mallikarjunan, A. Bhuvaneshwaran, K. Sundarakantham, and S. Mercy Shalinie, "DDAM: detecting DDoS attacks using machine learning approach," in Computational Intelligence: Theories, Applications and Future Directions-Volume I: ICCI-2017, pp. 261-273, Singapore, Aug. 2018.

[16] I. Sharafaldin, A.H. Lashkari, S. Hakak, and A.A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in 2019 International Carnahan Conference on Security Technology (ICCST), pp. 1-8, Oct. 2019.

[17] S. Pande, A. Khamparia, D. Gupta, and D.N. Thanh, "DDOS detection using machine learning technique," in Recent Studies on Computational Intelligence: Doctoral Symposium on Computational Intelligence (DoSCI 2020), pp. 59-68, Springer Singapore, 2021.

[18] K. Kumari and M. Mrunalini, "Detecting Denial of Service attacks using machine learning algorithms," Journal of Big Data, vol. 9, no. 1, pp. 1-7, Dec. 2022.

[19] M. Zekri, S. El Kafhali, N. Aboutabit, and Y. Saadi, "DDoS attack detection using machine learning techniques in cloud computing environments," in 2017 3rd international conference of cloud computing technologies and applications (CloudTech), pp. 1-7, Oct. 2017.

[20] R. Wazirali and R. Ahmad, "Machine Learning Approaches to Detect DoS and Their Effect on WSNs Lifetime," Computers, Materials & Continua, vol. 70, no. 3, Mar. 2022.

[21] S. Salmi and L. Oughdir, "Performance evaluation of deep learning techniques for DoS attacks detection in wireless sensor network," Journal of Big Data, vol. 10, no. 1, pp. 1-25, Dec. 2023.

[22] R. J. Alzahrani and A. Alzahrani, "Security analysis of DDoS attacks using machine learning algorithms in networks traffic," Electronics, vol. 10, no. 23, p. 2919, Nov. 25, 2021.

[23] University of New Brunswick, "Canadian Institute for Cybersecurity DDoS Attack Dataset (2019)," [Online]. Available: https://www.unb.ca/cic/datasets/ddos-2019.html.

[24] A. I. Jony and A. K. B. Arnob, "A long short-term memory based approach for detecting cyber attacks in IoT using CIC-IoT2023 dataset", Journal of Edge Computing, vol. 3, no. 1, pp. 28-42, 2024. Available from: https://doi.org/10.55056/jec.648.

[25] X. D. Hoang and Q. C. Nguyen, "Botnet detection based on machine learning techniques using DNS query data," Future Internet, vol. 10, no. 5, p. 43, May 18, 2018.

[26] T. H. Kim, D. C. Park, D. M. Woo, T. Jeong, and S. Y. Min, "Multi-class classifier-based AdaBoost algorithm," in Intelligent Science and Intelligent Data Engineering: Second Sino-foreign-interchange Workshop, IScIDE 2011, Xi'an, China, October 23-25, 2011, Revised Selected Papers 2 2012, pp. 122-127.

[27] S. Bashir, U. Qamar, F. H. Khan, and M. Y. Javed, "An efficient rule-based classification of Diabetes using ID3, C4.5, & CART ensembles," in 2014 12th International Conference on Frontiers of Information Technology, Dec. 17, 2014, pp. 226-231.

[28] S. Sikkanan and M. Kasthuri, "Denial-of-service and botnet analysis, detection, and mitigation," in Research Anthology on Combating Denial-of-Service Attacks, 2021, pp. 20-48.

[29] F. S. Lima Filho, F. A. Silveira, A. de Medeiros Brito Junior, G. Vargas-Solar, and L. F. Silveira, "Smart detection: an online approach for DoS/DDoS attack detection using machine learning," Security and Communication Networks, vol. 2019, pp. 1-5, Oct. 13, 2019.

[30] C. Kemp, C. Calvert, T. M. Khoshgoftaar, and J. L. Leevy, "An approach to application-layer DoS detection," Journal of Big Data, vol. 10, no. 1, p. 22, Feb. 13, 2023.