Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 28 No. 2 (2025)

Mitigating DDoS Attacks: A Machine Learning Approach for Enhanced Detection and Response

Dr.S.Aruna Deepthi[1], Dr.T. Padmapriya[2], Saravanan. R[3], Dr.B.Senthilkumaran[4], Ch Bhupati[5]

[1]Assistant Professor, ECE, Vasavi College of Engineering, Hyderabad, Telangana, jaishriresch@gmail.com

[2]Melange Publications, Puducherry, India, padmapriyaa85@ptuniv.edu.in

[3]Associate Professor, Department of Information Technology, Sri Manakula Vinayagar Engineering College, Puducherry, India, rsaravanan26@gmail.com

[4]Assistant Professor, Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology (deemed to be university Estd. u/s 3 of UGC Act, 1956), Vel Nagar, Chennai, Tamilnadu, India, skumaran.gac16@gmail.com

[5]Department of IoT, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, 522302, Andhra Pradesh, India, bhupati@kluniversity.in

Article History:

Received: *23-07-2024*

Revised: *04-09-2024*

Accepted: *14-09-2024*

Abstract

Distributed Denial of Service (DDoS) assaults continue to be among the most dangerous risks on the Internet. With the advances in equipment for spotting and mitigating these attacks, crackers have improved their skills in originating new DDoS attack types with the intent of cloning normal traffic behavior therefore becoming silently powerful. The so-called low-rate DoS assaults are a portion of these effective DDoS attack types that aim to archive limited network traffic. This paper proposes a machine learning algorithm for the mitigation of DDoS outbreaks in the application layer. Our scheme seeks to increase the exactness and efficacy of DDoS attack detection by employing the robustness of machine learning procedures which include neural networks and support vector machines, in combination with superior feature engineering and real-time monitoring. Our outcomes show that, of the four Machine Learning algorithms, Maximum Learning Performance (MLP) results in the best sorting marks. Particularly MLP leads to an F1-score of 98.04% for legitimate traffic, 99.30% for attack traffic on emulated movement, and an F1-score of 99.87% for target traffic and 99.95% for legitimate transportation on real traffic. When it concerned the procedure of distinguishing emulated traffic using FL, MLP, and EC, we were capable of gaining an F1-score of 98.80% for malware traffic and 99.60% for valid movement; but, when it related to real traffic, we were managed to obtain an F1-score of 100% for the assault traffic and 100% for normal traffic.

Keywords: DDoS attack; machine learning algorithm; Enhanced detection and response; Network security

1. Introduction

Attackers employ application-layer DDoS attacks, or App-DDoS, to take down a specific server by submitting counterfeit packets in its domain. Neither the target server nor the IDS (Intrusion

Detection System) may separate between the attackers' packets and those of legitimate users as the nasty sachets reflect the movements of genuine employers and hold the authenticated IP statements. Despite the compassion the main purpose of DDoS attempts is to ensure that the websites that are targeted respond so slowly that they evolve into unusable or shut down permanently. They overwhelm the object server's bottleneck funds for the purpose of carrying out this. For application-layer DDoS attacks, TCP/IP stacks, CPU cycles, RAM, I/O bandwidth, and disk/database bandwidth are the supposed bottleneck resources. Through the process of a series of legal requests, the offender overrules the bottleneck resources in an App-DDoS attack. Every bot framework that desires to engage in the invasion must initially establish a TCP connection with its targeted server, which calls for an authenticated IP address, to kick off the attack. Artificial intelligence procedures, specifically machine learning gets closer are being studied by academics and industry participants for the purpose of recognizing App-DDoS motion when confronting a substantial amount of categorized trials (the trials which include labels YES or NO), machine learning techniques together with logistic regression, Naïve Bayesian, random forest, KNN, and SVM are likely to successfully classify binary data (the data belonging to class YES or class NO). How it functions is that the arrangement types were nominated by human professionals. Through an arrangement of nonlinear processing layers, even feature selection in deep-learning (DL) methodologies including CNN and RNN can be executed by a machine without human supervision. The labeled samples are subsequently employed to customize and train an ML model utilizing the attributes that were selected. The class of future information is anticipated via a previously acquired machine learning model. Thus, ML methods have the potential to be a helpful instrument for identifying App-DDoS traffic. Some new machine learning approaches that recognize App-DDoS traffic are examined here. A few of the new machine learning procedures for discovering App-DDoS traffic are also examined in this paper [1]. The primary dangers to safety to web security are present, during which an enormous amount of zombie instruments downpour the web server thru packets. One of the primary dangers related to the confidence of the internet is the distributed denial of service (DDoS) attack, through which an immense number of virtual machines harass the web server with tremendous quantities of data. The primary goal of the biggest hazards to web security is the immense packet drowning of the web server through numerous zombie equipment. The attack was launched in a couple of stages: an solicitation film DDoS attack began shortly thereafter the transport layer DDoS attack had expired to acquire access to the network host. Figure 1.1 shows the software defined network architecture for clear details. Given that there are lots of methods to recognize DDoS attacks at the transport layer, the attack was halted in only a few hours. However, it took a week to eliminate it because there aren't many application layer DDoS assault detection tools available [2].

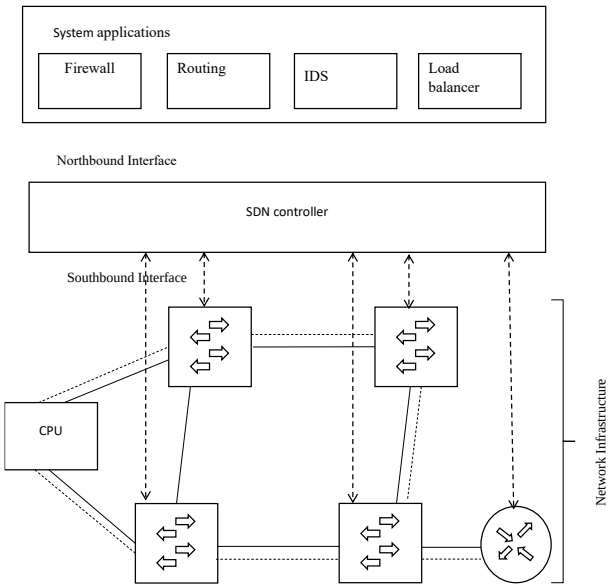

Fig. 1.1. Software defined networking architecture

The technique of DDoS attack revealing by machine learning processes is now increasingly the focus of investigation to cope with the previous problems. Researchers admire the machine learning algorithm's potential to find unreasonable information contained within massive quantities of data. The machine learning detection model has the advantage of being immediately changed by new data. Certain inadequacies persist. Longer prediction intervals are crucial for machine learning algorithms as a result of their significant computational complexity. The machine learning algorithms used to identify DDoS assaults fail to inspect the time correlation of traffic data. The current study introduces Principal Component Analysis-Recurrent Neural Network (PCA-RNN), an ML method that recognizes DDoS attacks, in reply to these obstacles. For the reason to guarantee that our algorithm can handle every possible type of violence, we initially extract every important detail, which assists with the problem of an individual app scenario. The features comprise four aspects: the web violence feature, the flow time feature, the flood feature, and the slow attack feature [3]. Likewise, high-tech government resources representing various countries have pointed out that one of the most fashionable ways that crackers get involved with official websites is by using DDoS attacks. Over time, researchers have developed a passion for DDoS attack detection. The explosive progress of Internet of Things (IoT) services takes contributed to an immense rise in DDoS outbreaks. As an outcome, network scientists and computer scientists went looking for several methods to detect DDoS attacks and put together moves to anticipate an exact attack. Although rule-based computers have been created to detect such attacks, their usefulness has been constrained by the complicated characteristics of DDoS, where numerous variables participate heavily [4].

The basic subject matter structure for the article appears as follows: Section 3 illustrates how machine learning techniques get closer and might reduce DDoS attacks in the network; Section 4 evaluates the empirical findings of this mitigation of DDoS strikes; and Section 5 delivers a conclusion.

2. Related Works

Aslam, M.et. al [5] To recognize DDoS occurrences for network communications of SDN-enabled IoT, we established an AMLSDM agenda in this paper that is centered on an adaptive machine learning classification model. A DDoS mitigation technology is a further component in the AMLSDM structures, permitting network resources to be transformed into ordinary network traffic. The AMLSDM framework's multilayered feed-forwarding design combines SVM, NB, kNN, LR, and FR predictors in the main level. The first layer's yield is introduced into the next film, EV, which enhances the first level classifiers' performance to pinpoint DDoS attacks. At the third layer, the expert adaptive machine learning model get ahead DDoS attacks by investigating real-time network traffic. Our suggested method takes place in four distinct periods: (i) teaching the adaptive sorting technique; (ii) mining of attributes of SDN-enabled Internet of Things (IoT) network traffic; (iii) real-time network traffic classification for DDoS identification; and (iv) DDoS limitation.

Prasad, A. et. al [6] Despite numerous attempts by the sector and researchers to strengthen online services, DDoS attacks continue to diminish their availability. Present technologies contain plenty of processing overhead and are examined employing a tiny pool of datasets. In this investigation, a physically cheap superior dataset developed from strident datasets spanning different VMFCVD approaches was generated by a significant struggle on feature choice, preceded by a systematic style to the construction and variety of the highest groups. Processing overhead for VMFCVD is the least when the server is hacked. Its effectiveness was investigated by way of multiple rounds of extensive challenges. However, according to the outcome of the experiment, VMFCVD performed more effectively in terms of accurate classification than prior investigations. When compared to all previous research, we have minimized the dataset to the highest degree achievable. While sustaining an accuracy of 99.99%, VMFCVD lowered the dataset by 98.2% in a few cases. For researchers to develop the model on an authentic server, we propose to deliver a generic DDoS and botnet dataset in future work. In the instance that a device initiates multiple malicious network messages, we must include a module to identify and eliminate the device.

Mittal, M.et. al [7] Differentiating between benign traffic and DDoS attacks with varying rates and patterns is a very difficult problem. Over the years, numerous effective deep learning techniques for DDoS attack detection have been put forth by other researchers. Unfortunately, though, these techniques only cover a very small area because attackers are always evolving their tactics and abilities to launch new, zero-day DDoS attacks with distinct traffic patterns. In this research, we reviewed the DDoS attack detection system based on DL techniques using the SLR protocol. The examination and judgment of the SLR protocol's final results are given below: Differentiating between benign traffic and DDoS attacks with varying rates and patterns is a very difficult problem. Over the years, an abundance of excellent algorithms for DDoS attack detection has been put forward by other scientists. Unfortunately, nevertheless, these strategies represent a tiny region while attackers are continually changing their strategy and competencies to launch novel and zero-day DDoS attacks with very different traffic patterns. In this investigation, we investigated the DDoS assault detection system based on DL strategies applying the SLR protocol.

Shaikh, J. et. al [8] DDoS attacks are merely one of the various unwanted breaches that have ravaged the Internet. The constant change of circuits and attack patterns have caused it more challenging to

determine denial-of-service (DDoS) attacks, particularly when implementing typical intrusion detection tools. This research proposed a novel strategy named the DL-based hybrid CNN-LSTM model to solve that issue. The DDoS-specific CICDDoS2019 dataset is used in this algorithm. Feature removal reduces the number of attributes from 86 to 30. The proposed model performs better than individual CNNs or LSTMs, and an accuracy of 99.86% is achieved. Automatic Encoding is used for dimensionality reduction, and SMOTE is used to balance the dataset. The suggested designs perform more efficiently beyond present-day benchmarks and previous investigations. This thesis yet continues to accept that DDoS attacks continue to be an existential threat to networks owing to their continuing challenges in being observed.

Elubeyd, H.et. al [9] In summary, our research has conveyed the benefit of deploying deep learning methods to detect and prevent DDoS attacks in software-defined networking (SDN) conditions. Our recommended hybrid deep learning framework, which blends a 1D convolutional neural network (CNN), a restricted recurrent unit (GRU), and a dense neural network (DNN), has presented more effective accomplishment juxtaposing to standard machine learning techniques, perfectly spotting DDoS attacks and promising efficient operation of SDN networks. Notably, our model may identify both short- and long-term trends in the input data and has proven particularly efficient at recognizing low-rate DDoS attacks. Even if the results of our research are beneficial, it remains essential to take the study's limitations into perspective. As an instance, our suggested approach had been evaluated using a particular dataset; additional testing on other datasets and network topologies are required to confirm its generalizability. Future studies ought to focus on how to implement inexpensive measures of mitigation when an attack has been spotted.

Akgun, D.et. al [10] In this investigation, we provide a unique deep learning model-based intrusion detection system for distributed denial of service violations. We availed benefits of the 12-class CICDDoS 2019 dataset, which contains one benign class. We investigated an assortment of deep learning models for various levels per layer, which involves CNN, LSTM, and DNN. Another benefit is that we upgraded the system by leveraging methodologies for preprocessing such as feature selection and elimination. This helped us to drastically reduce a sequence of 88 features to 40 significant ones. We produced a new homogeneous assortment of data by choosing a comparable quantity of individual samples from every single attack type with random subset picking. After that, we abolished duplicate records to acquire an abundant, non-repetitive data set that was neglected by a wide range of pertinent studies. In this context, this freelance introduces two different data sets to the repository of research that possess a direct consequence on how well training techniques constructed making use of the CIC-DDoS 2019 data set perform.

Kareem, M. I. et. al [11] By utilizing two CIC datasets, we assessed the accuracy of five supervised ML methodologies. In this effort, four indicates as colleagues (accuracy, F-score, precision, and recall) were implemented as enactment extents. For five algorithm models, testing time had an impact on the sum of correlated capabilities, clusters of recorded attributes and development lumps; on the other hand, testing time was prompted by the number of estimators, acquisition rate, maximum depth, and the total amount of leaves in the J48. The degree of precision of REPT, RT, RF, and J48 was outstanding. The quickest testing session was essential for detection models that were speedy and reliable. The results highlighted why the J48 technique remains the best strategy for real-

time DoS and DDoS identification. The CICDDoS2019 dataset was utilized to test the proposed classifier, and preliminary findings highlighted that it could precisely identify both SYN and UDP attacks with a short test time. Consequently, the provided concept is viable for real-time implementation and is lightweight.

3. Methods and Materials

Several machine learning techniques that have been employed for DDoS attack detection in over the past a few decades. The key concept of machine learning (ML) is to routinely absorb from a group of statistics with the objective to accept particular configurations such DDoS attacks. Defense systems can identify if a certain user is an attacker or a regular user with the assistance of machine learning computations. An synopsis of the ML-based DDoS attack exposure mechanism is explained here. First, new linkage sachets are added to the database and evaluated employing straining policies. Carefully chosen (such as source and destination addresses, protocol names, and port numbers) are taken from the catalogue during the feature extraction progression. Some features are normalized in order to enhance the training process' performance. The training phase can be carried out by the machine learning algorithms to make sure they may detect patterns in the dataset. An incoming packet is determined regardless of both a legitimate user or a DDoS attack in accordance with the learning parameters. In the ultimate procedure, the structure gets rid of the detected DDoS wallets and impacts its purification strategy with the goal it will spread over to the new arriving traffic. Supervised, unsupervised, and semi-supervised educational practices comprise the different groups of available machine learning techniques. It also sets forth how distinct ML procedures can be organized depending on the techniques for education that are currently employed. Additional property data on each category can be located in the subsection that precedes.

3.1 Supervised Learning

Supervised learning is an instance of ML where procedures absorb to foresee output variables (Y) by employing input variables (X) as the equivalent of a teacher or supervisor. Since a procedure can gain knowledge from a designated teaching dataset and the education progression finishes as soon as the model reaches a suitable performance level, this category is sometimes referred to as supervised learning. The support vector machine (SVM), logistic regression, naive Bayes, linear regression, and k-nearest neighbor (KNN) algorithms are the most commonly implemented supervised learning algorithms. Further, the two distinct types of supervised learning exist:

Classification: Using a labeled training dataset, the software learns new observation classification talents. The classification's output can be either multi-class, suggesting it consists of more than two classes, or bi-class, signifying it is capable of determining whether a specific consumer is an attacker or a regular user, or if an email is spam or legit.

Regression: A regression task is transported by the algorithm. In mandate to calculate the linear or continuous output (y), it is trained to estimate the mapping function (f) using a training dataset (x). A stock price or the price of an asset, for example can be estimated using such kind of algorithm [12].

3.2 CICDDoS2019 Dataset

One of the most complex challenges affecting machine learning (ML) intrusion detection methods is the absence of datasets. The main causes for why there are not enough datasets in the intrusion detection sector are privacy and legal challenges. Network traffic holds extremely confidential data that might reveal users' identities in addition to the firm's secrets if it comes out to the public or even gets closer into direct proximity with customers. Twelve distinct DDoS attacks are comprised in the dataset, which can be started in the application layer by deploying transport layer protocols which involve User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). The ranking of attacks is completed in terms of reflection-based and exploitation-based crimes. The dataset was obtained across an interval of two days. Every single physical assault category, however, was submitted in distinct PCAP and CSV formats. 87 flow characteristics are contained in the dataset, which was acquired using CICFlowMeter tools. This dataset's most notable characteristic is that it's up to date, representing both unrelated traffic and the most recent breaches.

3.3 Recommended Architecture

The intrusion detection scheme for the suggested strategy is created in tandem with the three primary methods for the identification and segmentation of incoming network data, as demonstrated in Figure 3.1. TCP, UDP, or an amalgam of the two transport layer protocols can be utilized to carry out these attacks through the application layer. The system architecture, gathering information methodologies, and classification strategy are all described in this section. This overcomes the challenge by fitting the cost analyst and utilizing the minimum-maximum normalization procedures to correctly train the model.

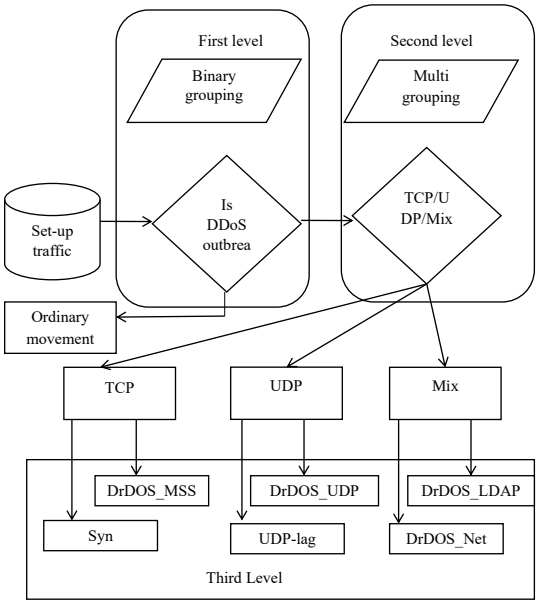

Fig. 3.1. Overall Structure of the Model

3.4 Preprocessing of Datasets

Standardizing to produce a consistent data set is the forthcoming step. First, the parameters of the socket are taken out: All socket houses, comprising source and destination ports, IP addresses, timestamps, and flow IDs, were removed. Although these qualities vary from network to network, we are forced to apply packet characteristics to train the model. As a consequence, we can acquire 80 different characteristics for the CIC-DDoS2019 model input. Next, grooming the data: The original data includes the majority missing (nan) and unlimited values; all of these figures have been extracted from the data. In addition, classifier construction freedom is granted by the min-max normalization. The main benefit of the normalization procedure is that it minimizes divergence by precisely conserving all of the associations among the data. As a result, every attribute is secret the apt level of the classifier when implementing the min-max normalization tactic, and the basic distribution of the associated features is unmodified. Attack data sets generally have combined continuous and discrete value characteristics. The choice of feature morals will be shifted when isolated and nonstop principles are amalgamated. Many ML algorithm training methods use min-max normalization to normalize each attribute and only accept certain sorts of input. The conversion formula is as outlined below:

$$y' = \frac{y-min}{max-min} \quad (1)$$

By including two numerical columns, we gained the ability to train our model for both multi-class and binary classification. Thus, we classify all DDoS clusters as outbreak categories, separately daily. Hence, we scramble the string quantities intended for the standard and violence signs to the binary values of 0 and 1, correspondingly, in a single column. As TCP, UDP, or mixed-based attacks, we also convert the sequence charge associated with the assault form to numbers from 1 to 3.

3.4.1 First level

The first-level XGBoost-LGBM paradigm for application-layer DDoS attacks. Following data preprocessing in the initial phases of building the CatBoost classifier model, the proposed method comprises two components: deciding on features and training and evaluating the classifier. A selection planning based on the feature significance score calculated using XGBoost and LGBM is applied in the feature selection segment. The CatBoost classifier (CBC) undergoes training and testing according to the second portion of the processed data. We provide an approach to feature selection (Feature Selection Based on LGBM and Feature Selection Based on XGBoost) FSBLGXG to create feature rankings based on XGBoost and LGBM algorithms that rapidly recognize the perfect attribute combination. This is done by developing a new list of features using union selection from both. This way leverages classification efficiency as a measure of measurement and two algorithms to discover vital characteristic combinations.

3.4.2 Second level

Following the DDoS attack is associated at initial level, it can be further separated into three subsets at the second level: TCP, UDP, and Mix. This is achieved by using a voting classifier whose design depends upon the two procedures, XGBoost and LGBM.

3.4.3 Third level

When a hacking attempt is caught by the preliminary level and categorized as a DDoS attack, pre-trained classifiers in the third level will figure out the type of attack and pick out the most suitable plan of action based on whether it's is a TCP, UDP, or Mixed attack. Here an entire set of the most carefully determined criteria and appropriate weights for all forms of attack is present. Accuracy and speed have to be the guiding principles for the most vital elements at this level. We identified two characteristics to differentiate among UDP and UDP-lag attacks, twofold attributes to categorize amongst LDAP and NetBIOS assaults, and two parameters to separate between MSSQL and Syn crimes [13].

3.5 C-Support Vector Classification

When it relates to different uses like spam filtering, pattern recognition, and intrusion detection, SVM has become one of the most frequently employed supervised learning algorithms. For distribution estimation, regression, and classification, there are numerous SVM formulations.

As mentioned earlier, the primary goal of this study is to declare each IP address as either normal or fraudulent. Thus, for training and testing datasets, we pick c-support vector classification (C-SVC). The SVM Classification Confusion Matrix employing Dataset 1 is shown in Table 1.

Let $y_j \in S^o, j = 1,2, \dots m$, where m be the number of sample scenarios, and let $z \in S^m$ be the indicator vector including training vectors for $z_j \in \{-1,1\}$. Because we have three datasets with different quantities of features, the dimension parameter o in the trials comprises from three to five. The following optimization problem is tackled by employing C-SVC. We use -1 to represent "Normal IP" for IP addresses from the victim pool and +1 to represent "Attacker IP" for IP addresses from the attacker pool.

$$\underset{\varphi,c,\delta}{\text{minimize}} \frac{1}{2}\varphi^U\varphi + D\sum_{j=1}^{m}\delta_j \quad subject\ to\ z_j\left(\varphi^U\partial\left(y_j\right) + c\right) \geq 1 - \delta_j, \delta_j \geq 0,\ j = 1,2, \dots, m \tag{2}$$

where D is the regularization variable, which must to be greater than zero, $\partial\left(y_j\right)$ and conveys y_j into a higher dimensional space. We answer the following quadratic programming problem considering the vector variable α could be incredibly dimensional.

$$\min_{\beta} \frac{1}{2}\beta^U R\beta - e^U\beta \quad subject\ to\ z^U\beta = 0, 0 \leq \beta_j \leq D, j = 1, \dots, m \tag{3}$$

when R is a m by m positive semidefinite matrix, $R_{jk} \equiv z_j z_k L\left(y_j, y_k\right)$ is the kernel function, and $L\left(y_j, y_k\right) \equiv \tau\left(y_j\right)^U \tau(y_k)$ is the vector alongside all one components, $f = [1, \dots m]^U$. An ideal answer to equation (3) that employs a main dual tandem to solve it satisfies equation (4), and the decision function is (5).

$$\varphi^U = \sum_{j=1}^{m} z_j\beta_j\partial\left(y_j\right) \tag{4}$$

$$sgn(\varphi^U\partial(y) + c) = sgn\left(\sum_{j=1}^{m} z_j\beta_j L\left(y_j, y_k\right) + c\right) \tag{5}$$

Table 1. SVM Classification Confusion Matrix employing Dataset 1

		SVM classification	
		Attacker IP	Normal IP
Classification according to the data	Attacker IP	0.9396	0.0076
	Normal IP	0.2826	0.8846

In the model for prediction, we include $z_j\beta_j \ \forall j$, b, packaging, support vectors, and extra information which may include kernel parameters. The RBF (Gaussian) kernel that comes next has been employed as the kernel function in this research [14].

$$L(y_j, y_k) = exp\left(-\rho\|y_j - y_k\|^2\right) \quad (6)$$

4. Implementation and Results

As evident in Figure 4.1, we deployed Naïve Bayes, Decision Stump, Logistic Model Tree, Naïve Bayes Updateable, Naïve Bayes Multinomial Text, AdaBoostM1, Attribute Selected Classifier, Iterative Classifier Optimizer, and OneR to calculate the accuracy of our outlined MLP classification model.

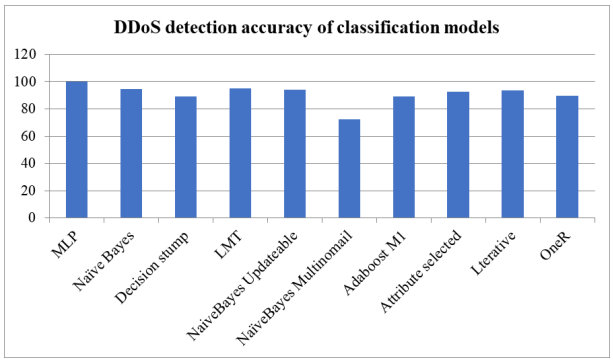

Fig. 4.1. Comparison of the specified MLP classifier's accuracy to a Replacement Rlassification Models

Using an competence of 98.99%, the outcomes demonstrated render visible that the MLP classifier performs more efficiently than any other classification model. Combining our indicated MLP classifier, we may detect DDoS strikes at the application level with velocity. We may differentiate between authorized clients and fraudulent users by implementing its recommended MLP classifier. On the flip side, a handful of the suggested IP addresses do not correspond with the characteristics of an ordinary user or an aggressor. In this research, we examined the potential of our suggested strategy by spending it to recognize breaches in real-world DDoS attack datasets, such as our dataset, our company's site logs from 2019, and CTU-13 (2011). Table 2a and Table 2b depicts the analysis of ten classifiers' detection accuracy [15].

Table 2a. The Proposed Approach (MLP) is Contrasted with other Possible Models

Criteria	NB	DS	LMT	NBU	NBMT
CM $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$	$\begin{bmatrix} 530 & 56 \\ 33 & 20 \end{bmatrix}$	$\begin{bmatrix} 500 & 50 \\ 46 & 300 \end{bmatrix}$	$\begin{bmatrix} 546 & 44 \\ 50 & 336 \end{bmatrix}$	$\begin{bmatrix} 520 & 55 \\ 36 & 345 \end{bmatrix}$	$\begin{bmatrix} 425 & 30 \\ 27 & 344 \end{bmatrix}$
Accuracy	0.9442	0.8922	0.9502	0.9434	0.8244
TP	0.9742	0.8812	0.9511	0.9111	0.7811
FP	0.543	0.6781	0.1887	0.1781	0.8889

Table 2b. The Proposed approach (MLP) is Contrasted with other Possible Models

Criteria	ABM1	ASC	ICO	OneR	MLP (Proposed)
CM $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$	$\begin{bmatrix} 436 & 45 \\ 32 & 323 \end{bmatrix}$	$\begin{bmatrix} 509 & 33 \\ 26 & 322 \end{bmatrix}$	$\begin{bmatrix} 520 & 52 \\ 38 & 320 \end{bmatrix}$	$\begin{bmatrix} 536 & 55 \\ 38 & 328 \end{bmatrix}$	$\begin{bmatrix} 598 & 60 \\ 40 & 50 \end{bmatrix}$
Accuracy	0.8912	0.9256	0.9365	0.8946	0.9988
TP	0.8722	0.8811	0.8911	0.8566	0.9888
FP	0.9988	0.8786	0.8898	0.9789	0.1322

4.1 CICDDoS2019 Dataset Analysis

The CICDDoS2019 dataset, a benchmark in cybersecurity research permitted by the Canadian Institute for Cybersecurity, conveys a comprehensive safeguard containing both DDoS violence instances and benign network traffic samples. Recognized for its massive grouping of network traffic data, it delivers an extensive spectrum of rich traits that surpasses the capacity and breadth of present-day datasets. With twelve different kinds of DDoS attacks characterized by 88 special qualities proposed from flow-based features, this dataset significantly raises the value for extensive research and trustworthy DDoS attack detection. Several metrics have been identified to measure the performance of the proposed model on this dataset, including as accuracy, precision, recall, and F1-score. Exceptional performance indicators stemmed from the rigorous evaluation: 98.70% accuracy, 98.78% precision, 98.81% recall, and a 98.78% F1-score. These superb outcomes reflect the method's capacity to pinpoint and categorize DDoS attacks within this unique dataset [16]. Precision, recall, f1-score, and other evaluation criteria are the ones employed in a complement to the accuracy score to evaluate and grade each classifier. Table 2a and Table 2b exhibit the overall accuracy of each classification algorithm for the disproportionate dataset, even though Table 3 exhibits the output observations for the balanced dataset. The best values from the five rounds of observations are utilized to decide on the data. All of the classification procedures had tremendously precise results given that the unbalanced dataset is skewed towards the attack class. However, none of these criteria assist us in identifying the top-performing methodology for DDoS attack assessment. Here, every technique works astonishingly well including unbalanced data, except Naïve Bayes. Meanwhile, we discovered that the accuracy altered little for the balanced dataset. Table 3 illustrates that distance-based classification approaches such as K-NN and tree-based algorithms like Random Forest and Decision Tree perform smoothly although Naïve Bayes yields excellent precision but the

additional classification techniques, SVM and Logistic Regression, perform worse. For each kind of classification algorithm, Figure 4.2 distinguishes the accuracy rates of the unequal and balanced datasets. Furthermore, Figures 4.3, 4.4, and 4.5 demonstrate the F-1 score, Precision, and Recall of the balanced and unbalanced datasets, respectively. Soon after investigating outcomes, it was determined that distance-based and tree-based classification approaches, such as Random Forest and Decision trees, yielded nearly full accuracy on both forms of datasets. These three classifiers function the best even when extra variables are also taken into consideration. However, when the constraints for each classifier are modified, a tiny variation in performance is recognized. Here, we made a commitment to decide which method behaved the best overall [17].

Table 3. Well-adjusted dataset outcomes

Poised dataset	Accuracy	Precision	Recall	F1 score
Decision tree	100	1	1	1
Naive Bayes	97.36483	0.95	0.95	0.95
Logistic Regression	78.45296	0.96	0.88	0.89
SVM	51.21265	0.36	0.6	0.44
k-NN	200	1	1	1
RF	200	1	1	1

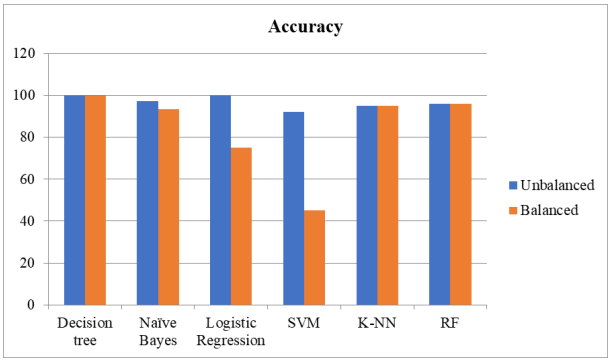

Fig. 4.2. Grouping algorithm's accuracy score

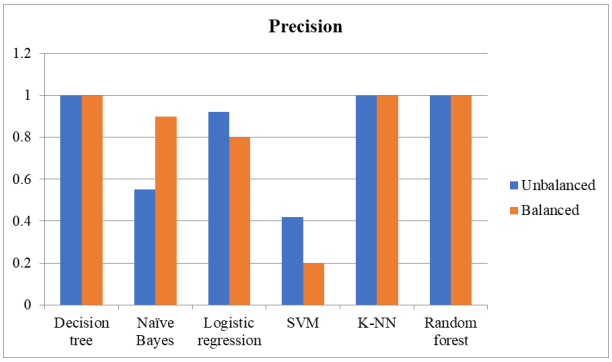

Fig. 4.3. Precision Scores for Various Algorithms

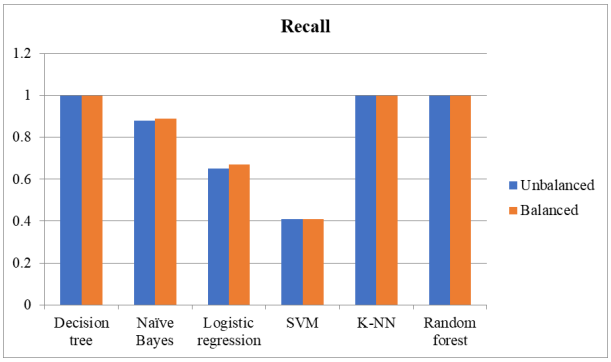

Fig. 4.4. Recall scores

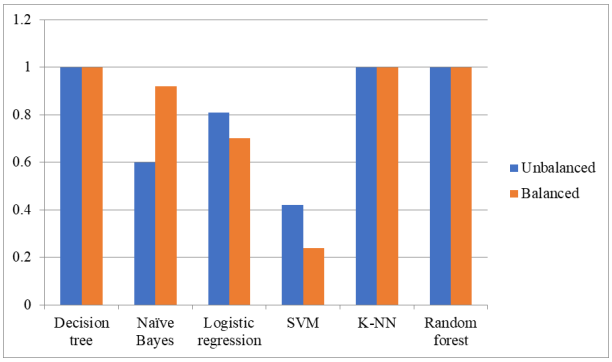

Fig. 4.5. Graph for F-1 score

5. Conclusion

In contrast with prior techniques, this type of technology displays higher adaptability with emerging traffic patterns and can deal with an assortment of DDoS attack types. Flexible thinking can be guaranteed by modular structures, which understand the demand for numerous types of protection and facilitate independent enhancements in multiple parts. The versatile nature of the model is enhanced by its compatibility with numerous SDN controllers. In SDN-based networks, the proposed approach enables a real-time solution for DDoS attack detection and reduction. Future areas for investigation include testing the model in live environments, analyzing deep learning models, determining the model with real-world network topologies, and upgrading the model's capabilities to cloud-based SDN environments for overall safety in hybrid network infrastructures. The current research offers a machine learning-based DDoS exposure methodology for SDN. The DDoS attack detection module and the characteristic extraction and model selection module make up the strategy. The grades recommend that the anticipated methodology can spot DDoS assaults and educate handlers. We will similarly examine methods to shorten oppositional attacks to promote the system’s toughness and flexibility. To create a more complete and completely inclusive network security solution, we also aim to integrate this kind of approach with other technologies. The CICDDoS2019 dataset, which embraces the most recent DDoS attack signatures, is a legitimately recent dataset that we utilized for this research. Major supervised classification strategies have been deployed during the assessment to accurately classify the attack from the legitimate flows. K-NN, random forest, and decision tree algorithms surpassed all other classifiers when the results were compared. Even though the preliminary conclusions are uplifting we want to improve the experiment by emphasizing numerous DDoS attack types and utilizing a broader dataset. We'll emphasize our attempts in these domains continuing forth.

References

[1] Beitollahi, H., Sharif, D. M., & Fazeli, M. (2022). Application layer DDoS attack detection using cuckoo search algorithm-trained radial basis function. *IEEE Access*, *10*, 63844-63854.

[2] Yadav, S., & Selvakumar, S. (2015, September). Detection of application layer DDoS attack by modeling user behavior using logistic regression. In *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)* (pp. 1-6). IEEE.

[3] Li, Q., Meng, L., Zhang, Y., & Yan, J. (2019). DDoS attacks detection using machine learning algorithms. In *Digital TV and Multimedia Communication: 15th International Forum, IFTC 2018, Shanghai, China, September 20–21, 2018, Revised Selected Papers 15* (pp. 205-216). Springer Singapore.

[4] Sarraf, S. (2020). Analysis and detection of ddos attacks using machine learning techniques. *Am. Sci. Res. J. Eng. Technol. Sci*, *66*(1), 95-104.

[5] Aslam, M., Ye, D., Tariq, A., Asad, M., Hanif, M., Ndzi, D., ... & Jilani, S. F. (2022). Adaptive machine learning based distributed denial-of-services attacks detection and mitigation system for SDN-enabled IoT. *Sensors*, *22*(7), 2697.

[6] Prasad, A., & Chandra, S. (2022). VMFCVD: an optimized framework to combat volumetric DDoS attacks using machine learning. *Arabian Journal for Science and Engineering*, *47*(8), 9965-9983.

[7] Mittal, M., Kumar, K., & Behal, S. (2023). Deep learning approaches for detecting DDoS attacks: A systematic review. *Soft computing*, *27*(18), 13039-13075.

[8] Shaikh, J., Butt, Y. A., & Naqvi, H. F. (2024). Effective Intrusion Detection System Using Deep Learning for DDoS Attacks. *The Asian Bulletin of Big Data Management*, *4*(1).

[9] Elubeyd, H., & Yiltas-Kaplan, D. (2023). Hybrid deep learning approach for automatic Dos/DDoS attacks detection in software-defined networks. *Applied Sciences*, *13*(6), 3828.

[10] Akgun, D., Hizal, S., & Cavusoglu, U. (2022). A new DDoS attacks intrusion detection model based on deep learning for cybersecurity. *Computers & Security*, *118*, 102748.

[11] Kareem, M. I., & Jasim, M. N. (2022). Fast and accurate classifying model for denial-of-service attacks by using machine learning. *Bulletin of Electrical Engineering and Informatics*, *11*(3), 1742-1751.

[12] Aljuhani, A. (2021). Machine learning approaches for combating distributed denial of service attacks in modern networking environments. *IEEE Access*, *9*, 42236-42264.

[13] Kanber, B. M., Noaman, N. F., Saeed, A. M., & Malas, M. (2022). DDoS Attacks Detection in the Application Layer Using Three Level Machine Learning Classification Architecture. *International Journal of Computer Network & Information Security*, *14*(3).

[14] Kato, K., & Klyuev, V. (2014). An intelligent ddos attack detection system using packet analysis and support vector machine. *IJICR*, *14*(5), 3.

[15] Ahmed, S., Khan, Z. A., Mohsin, S. M., Latif, S., Aslam, S., Mujlid, H., ... & Najam, Z. (2023). Effective and efficient DDoS attack detection using deep learning algorithm, multi-layer perceptron. *Future Internet*, *15*(2), 76.

[16] Alashhab, A. A., Zahid, M. S., Isyaku, B., Elnour, A. A., Nagmeldin, W., Abdelmaboud, A., ... & Maiwada, U. (2024). Enhancing DDoS Attack Detection and Mitigation in SDN using an Ensemble Online Machine Learning Model. *IEEE Access*.

[17] Gohil, M., & Kumar, S. (2020, December). Evaluation of classification algorithms for distributed denial of service attack detection. In *2020 IEEE Third International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)* (pp. 138-141). IEEE.