MACHINE LEARNING-DRIVEN DDOS ATTACKS DETECTION IN SOFTWARE-DEFINED NETWORKS

[1]Dori Lal, [2]Dr. Ram Milan
[1]Research Scholar, [2]Assistant Professor & Head
CSE Department
Sunderdeep Engineering College, Ghaziabad, UP, INDIA
dorilalagarwal@gmail.com, rammilan.in@gmail.com

ABSTRACT

Software Defined Networks (SDNs) provide a holistic network perspective by separating the control plane from the data plane. SDN aims to simplify network operations by centrally managing the entire network. Broadly implemented in data center networks, SDN uses software-based controllers for more flexible and efficient network management compared to traditional hardware-based approaches. The principal characteristics of SDN encompass the separation of Centralized Control, Data plane, Control Plane and Programmability via APIs, Flexibility, Agility, and Enhanced Network Management. SDNs, however, exhibit susceptibility to Distributed Denial of Service (DDoS), a perilous assault causing resource depletion and hindering service provision.

The primary controller represents a singular point of failure, and if compromised, malevolent entities could seize command of the whole network, manipulate traffic, and disrupt operations. Numerous scholars have proposed diverse methodologies for detecting DDoS attacks; nevertheless, these methodologies are afflicted with high instances of false positives, resulting in diminished accuracy, primarily due to the adoption of unqualified attributes and unrealistic datasets. InSDN dataset constitutes a comprehensive Software-Defined Network (SDN) repository for assessing Intrusion Detection Systems. It encompasses benign content and various forms of attacks that may manifest across distinct components of the SDN standard. InSDN encompasses an array of attacks, encompassing DoS, DDoS, brute force attacks, web application breaches, exploitations, probes, and botnets. Machine learning algorithms have witnessed widespread adoption in recent times for the identification of DDoS attacks. This study employs supervised Machine Learning (ML) algorithm Random Forest (RF) for DDoS detection. For dataset dimensionality reduction, the feature selection technique "Recursive Feature Elimination with Cross-Validation (RFECV)" is utilized, falling under the Wrapper method category.

Keywords: Dataset, Random Forest, DDoS, SDN, Intrusion, ML, Feature Selection

1. INTRODUCTION

Network traffic and reliance have grown because of the emergence of latest technologies like cloud computing [1], [2], the Internet of Things (IoT) [3], [4], 5G Technology [5], big data [6], [7], etc. Due to surged online work, study, education, teaching, research and entertainment, putting bigger expectations on network Availability and security. The quick accumulation of internet-linked devices and the growing dependence on digital services have made network security a critical concern for organizations worldwide. The challenges with conventional network architectures have become extra considerable and cannot fruitfully address users' requirements.

SDN [8], [9], [10], [11], [12] has emerged as a transformative technique to network management, offering enhanced flexibility, programmability and unified control. SDN decouples the data plane from the control plane, permitting for more dynamic and efficient network management. SDN Architecture is presented in Figure 1.

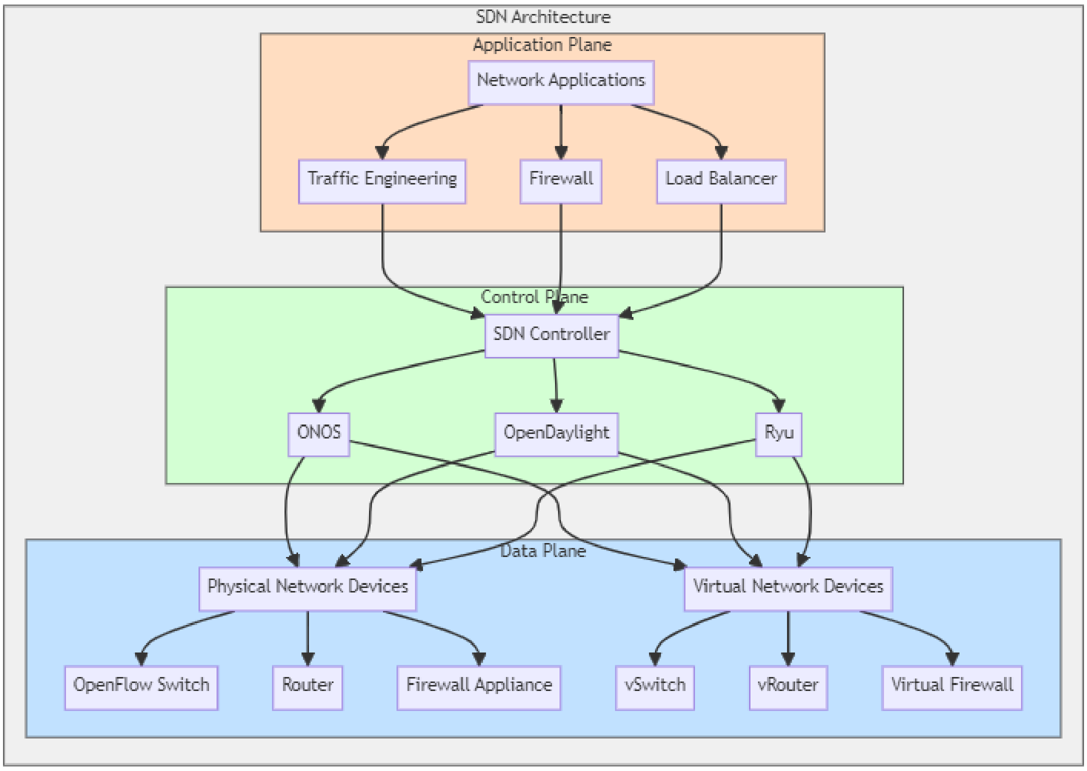


Figure 1: Software Defined Networking Architecture

The centralized nature of the SDN controller becomes a single point of failure, and attackers can exploit this to launch Multi-Vector DDoS attacks that can cripple the entire network infrastructure. One of the most significant threats to network availability and security is the Distributed Denial of Service (DDoS) attack [13], [14]. DDoS attacks aim to overwhelm network resources by flooding them with an excessive amount of traffic, rendering services unavailable to legitimate users. This scholarly article scrutinizes the utilization of ML algorithms to identify DDoS assaults in SDNs. The investigation aims to construct a resilient detection framework capable of precisely pinpointing DDoS attack trends and reacting promptly to alleviate their repercussions. By combining the SDN architecture with machine learning-based detection approaches, this research aims to increase the security and resilience of modern network infrastructures.

Therefore, the primary contributions are outlined below:

I. The adoption of a better method for spotting DDoS assaults in an SDN context. "Recursive Feature Elimination with Cross-Validation (RFECV)" is an attribute chosen strategy that falls under the Wrapper method and is used to reduce the dimensionality of datasets while maximizing classification accuracy.

II. This study presents InSDN, an extensive dataset created especially for assessing software-defined network intrusion detection technologies. The dataset offers a realistic and broad baseline for security evaluation because it comprises a range of attack types as well as typical traffic.

III. For DDoS identification, this research uses the supervised ML method Random Forest (RF). This algorithm was selected due to its excellent accuracy in classification tasks and their efficiency in processing huge datasets.

2. LITERATURE SURVEY

In recent years, there has been a burgeoning interest in exploring various machine learning techniques for the purpose of identifying distributed denial of service (DDoS) attacks within software-defined networks (SDNs). The primary objective of this comprehensive literature review is to offer a more profound comprehension of the existing methodologies.

Reference [15] employed a combination of diverse machine learning algorithms, including KNN, NB, SVM, and SOM, utilizing the Center for Applied Internet Data Analysis (CAIDA 2016) dataset to identify anomalous traffic patterns in SDN infrastructures. Upon amalgamating fundamental supervised algorithms with the unsupervised SOM algorithm, the outcomes suggest that the SVM-SOM amalgamation demonstrates superior performance levels in DDoS attack classification, achieving an accuracy rate of 98.12%.

In Reference [16], machine learning models such as SVM, Decision Tree, KNN, and Random Forest were applied to the SDN DDoS dataset for the purpose of detecting DDoS attacks within SDN environments. The final outcomes reveal distinct levels of accuracy for each model. Specifically, the SVM model exhibited an accuracy of 66.07%, the KNN model demonstrated an accuracy of 97.99%, the Decision Tree model showcased an accuracy of 99.95%, and the Random Forest model illustrated an accuracy of 99.99%.

Reference [17] initiates by cleansing and normalizing the CSE-CIC-IDS2018 dataset, followed by identifying the optimal feature subset through an enhanced binary grey wolf optimization algorithm. Subsequently, the optimal feature subset underwent training and testing across various machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (k-NN), Decision Tree, and XGBoost, leading to the identification of the most effective classifier for DDoS attack detection, which was then integrated into the SDN controller. The study proposes a method for DDoS detection in SDN grounded on feature engineering and machine learning, encompassing two modules: feature extraction and model selection, as well as DDoS attack detection. The feature extraction in Module 1 entailed the utilization of an enhanced binary grey wolf optimization algorithm, while five machine learning models—RF, SVM, XGBoost, Decision Tree, and k-NN—were employed to assess and designate the optimal classifier for both the original and feature-extracted datasets. The results indicated that XGBoost attained the highest accuracy of 0.969 on the original dataset; post feature extraction, the dataset's feature count decreased from 79 to 26, resulting in enhancements across all classifiers in various metrics. With scores of 99.13%, 99.92%, 98.43%, 99.13%, in f1_score, recall, precision and accuracy measures, respectively, the RF classifier demonstrated remarkable performance. To demonstrate the method's capacity to detect DDoS assaults and alert users, a selection of the most effective characteristics were used to deploy the superior classifier found in Module one to the controller for DDoS identification in Module 2.

Reference [18] introduces a novel dataset [19] that focuses on contemporary attack types, setting itself apart from prior studies. This dataset is categorized into five groups and comprises 25 attributes. The Multi-Vector attacks are conducted on the target server, with packet information extracted using Wireshark, a reliable tool known for producing authentic results that mirror real-world scenarios. The data collection includes various attack types aimed at the Application layer of the network. To facilitate classification, the dataset is partitioned into two groups namely testing group, Training group and purposes applying five different methods: SVM, K nearest neighbour, neural network, NB, RF. Processed datasets are adjusted to optimize the classification process. The analysis reveals that out of the five classification techniques, RF attain the maximum accuracy rate of 98.70% with a Weighted Average of 98.4%, demonstrating its effectiveness in accurately identifying DDoS attacks in future applications.

Reference [20] aims to enhance the recognition of DDoS assaults by means of research endeavors. Experimental assessments were carried out utilizing the two datasets namely- CIC-DDoS2019 and CIC-IDS2017, which encompass pertinent data concerning DDoS attacks. The process of feature selection entailed the utilization of Mutual Information (MI), Random Forest Feature Importance (RFFI) methodologies to pinpoint crucial features. These identified features were subsequently integrated into ML models such as Logistic Regression (LR), KNN, Weighted Voting Ensemble (WVE) RF, Gradient Boosting (GB). Notably, the Random Forest algorithm demonstrated a noteworthy predictive accuracy of 99.993% with sixteen features and 99.9977% with nineteen features, surpassing alternative approaches. The comprehensive outcomes underscore the efficacy of LR, KNN, WVE, GB, and RF when employing RFFI and Mutual Information for the selection of optimal features. Subsequent research endeavors could delve into the implementation of wrapper attribute chosen techniques, like sequential, in conjunction with the neural networks to bolster DDoS and additional assault identification capabilities.

Reference [21] explores the effectiveness of ML models like DT, Convolutional Neural Networks, RF, NGBooST and Stochastic Gradient Descent (SGD) on the CIC-DDoS2019 dataset for DDoS assault recognition within SDN setting. The outcomes reveal varying levels of accuracy for each model, with values of 0.99, 0.91, 0.98, 0.96, and 0.93, respectively.

The emphasis of reference [22] is the important contribution of finding new traits for identifying DDoS assaults. These distinct attributes are stored in file format CSV to create a dataset, which is then utilized in ML models training on the SDN dataset. Prior research on the identification of DDoS attacks has mostly used non-SDN datasets, with little study data available to the general public. The classification process makes use of a brand-new cross-breed ML architecture. The outcomes show how well the cross-breed model, which combines a SVM classifier and Random Forest classifier to classify network data, performs in testing, attaining an amazing accuracy of 98.8% with minimum false positive rate.

In Reference [23], the article presented a novel methodology for recognizing and mitigating DDoS assaults within a SDN framework. An advanced approach was taken by employing a ML model SVM classifier to identify potential threats. The model was further enhanced by integrating algorithm Genetic with algorithm Kernel Principal Component Analysis (KPCA) to enhance accuracy and minimize testing duration. KPCA was utilized for extracting essential attributes out of the DDoS dataset, while GA was implemented to optimize parameters for the SVM classifier. Moreover, Non-linear Radial Basis Function (N-RBF) was introduced to expedite the training process. Results from experiments indicated that KPCA surpassed Principal Component Analysis (PCA) in performance on the DDoS dataset, achieving an impressive accuracy of 98.907% and outperforming other existing models. By integrating a kernel function into PCA, a more substantial reduction in principal components can be attained, thereby enhancing overall system performance.

In Reference [24], various ML algorithms were trained utilizing the CIC-IDS2018 dataset, which was constructed utilizing CICFlowMeter tool on a vast collection of PCAP files. The tool extracted 84 statistical attributes related to traffic flow, encompassing details such as IP Address of initiator**,** flow duration, IP address of target, and statistical values like standard deviation, mean, highest and lowest of packet sizes. This research focuses specifically on segments of the dataset containing DDoS traffic exclusively, originating from known tools of DDoS assault like HOIC, Hulk, Golden-Eye. Erroneous data points containing NaN values and negatives were identified and removed, resulting in a dataset with over 11 million data points. This dataset was then split into two groups for training and testing labeled as IDS-Train and IDS-Test, respectively. Furthermore, feature chosen was performed using Chi-square method to determine the most suitable attributes for training on IDS-Train, leading to the retention of a total of 67 features. Key features included metrics such as "No. packets have data in forward flow," "Total length of forward packets," "Total backward packets per second," "Window bytes of initial forward flow," and "Average segmentation size in forward flow." Prior to training, normalization of IDS-Train was carried out employing the Euclidean norm technique to prevent overfitting. ML algorithms incorporating NB, RF, Linear SVM, and DT were applied to the public dataset CICIDS2018 for identifying intrusion attacks within an SDN network. Finally it indicated varying levels of accuracy for each model, with percentages of 95.67%, 67.69%, 99.97%, and 99.83%, respectively.

Reference [25] examined seven distinct machine learning models on dataset CICDDoS2019, with each model evaluated using 30, 20, and ten attributes. XGBoost demonstrated the highest accuracy (99.99996%) among models with 30 attributes. The Random Forest model showed the top accuracy (99.99999%) for models with 20 features, with a precision of 1, while KNN achieved an accuracy of (99.98%). For models with ten features, both XGBoost and RF displayed identical accuracies of (99.99%). CNN models with 30 attributes yielded an accuracy of (84.75%). The study found that XGBoost and RF were the most fruitful ML models. Reducing the number of features to five can result in time and cost savings. Analyzing 30 or 20 features for Anti-DDoS solutions is impractical due to the significant time required to distinguish between attacks and benign activity. Precision, indicating the model's frequency of accurately predicting outcomes, was examined for each machine-learning model. The XGBoost model with a 30-feature set achieved the highest precision level (100%). Conversely, the RF model achieved optimal precision with a 20-feature cluster (100%), while KNN demonstrated the best precision with a 20-feature set (99.99%), and the CNN model showed superior precision with a 20-feature group (98.99%). Thus, the Random Forest model with a 20-feature set provided the most precise outcomes.

Reference [26] utilized the deep learning algorithm CNN on the public dataset InSDN for identifying DDoS assaults in SDN specific networks. This research investigates the application of CNN for Intrusion Detection Systems and proposes techniques like L2 regularization and dropout to enhance its performance. While the achieved accuracy (93.01%) of the proposed CNN approach may not be adequate for real-world SDN environments, efforts are underway to improve the algorithm's performance by addressing Overfitting challenges and enhancing outlier detection capabilities.

Detection of DDoS assaults in software-defined environments commonly involves ML algorithms and network-based strategies. Algorithms such as XGBoost, RF, SVM, KNN, and DT have been employed in DDoS assault identification in SDN specific networks, showing promising results in accurately identifying such attacks based on various performance metrics.

3. METHODOLOGY

This comprehensive research focus to develop an accurate DDoS attack detection algorithm with minimum false-alarm rate. To train this model, InSDN dataset was used. This dataset contained 84 features. The feature set of the training dataset should be condensed, which is accomplished using the attribute chosen algorithm RFECV. This section outlines the methodology employed in identifying DDoS assaults in SDN specific environments using RF algorithm. The approach includes multiple steps like attribute chosen, data preprocessing, model training and evaluation. Public dataset InSDN is being utilized in both training and testing the model. Z-score normalization is applied for data standardization, and Recursive Feature Elimination with Cross-Validation (RFECV) is utilized for attribute chosen. Machine Learning process diagram is shown in Figure 2.

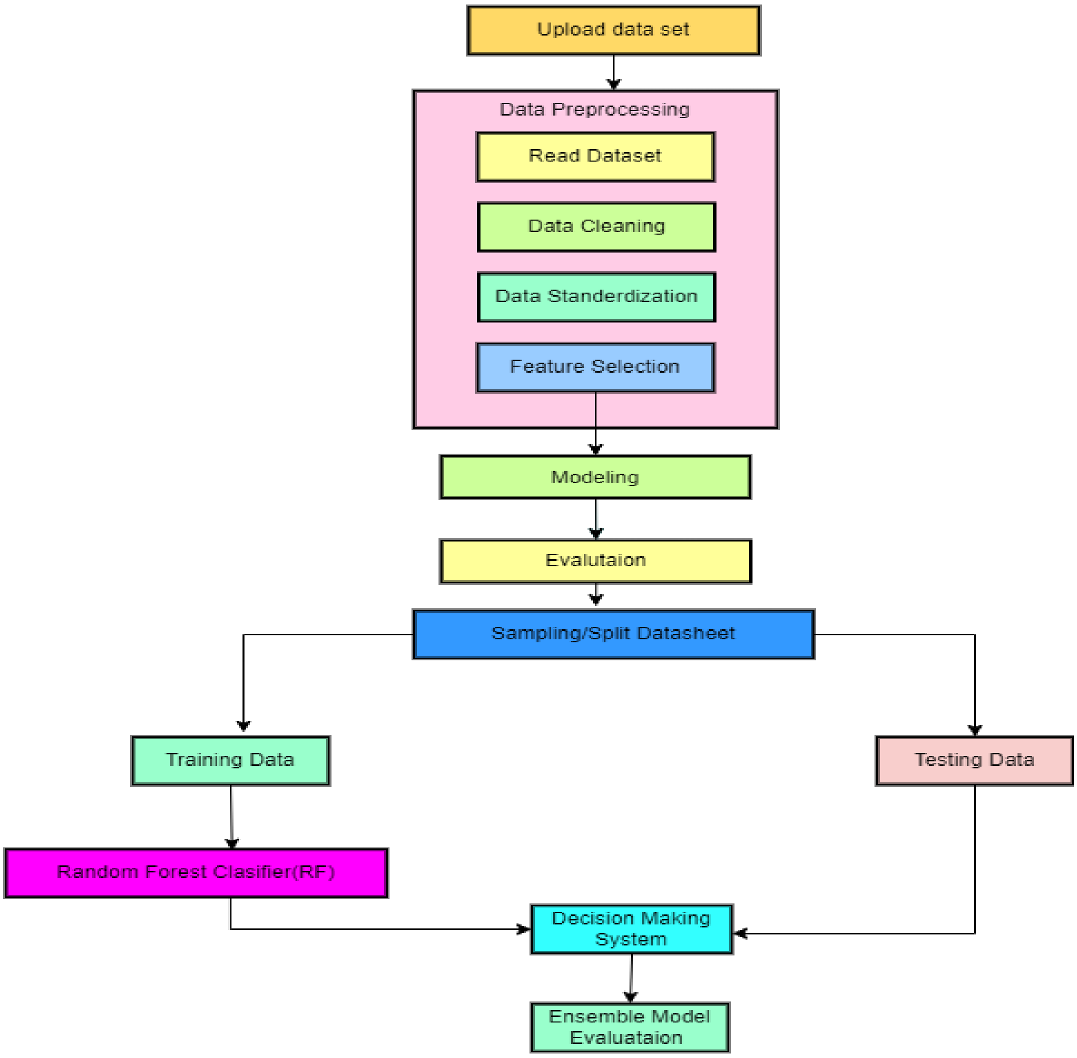


Figure 2: Machine Learning Process Diagram

3.1 READING THE DATASET

In this study, the InSDN dataset is employed in assessment of proposed Ensemble ML model. The InSDN dataset enables scholars to explore and formulate models for detection of intrusions tailored to SDN environments. It comprises normal network traffic such as secure shell, HTTP, FTP, HTTPS, E-mail, Domain name system as well as various attacks targeting traditional and SDN-specific networks. The intrusion categories consist of U2R, DoS, Probe, DDoS, BFA, Web based assaults and Bot network. This intrusion dataset was generated utilizing a simulated SDN environment composed of four different virtual machines. One machine assumes the role of the attacker, while the second machine operates as the susceptible Linux server Metasploitable-2. The remaining pair of machines are allocated to representing the Open vSwitch and Controller. Moreover, this dataset encompasses a variety of attacks that stem from diverse origins, both internal and external, in order to replicate authentic attack scenarios. This dataset can be found in both formats, namely PCAPs and CSVs, and is categorized into three distinct groups. The initial group, referred to as the OVS, pertains to attacks directed from external sources towards the inside network of SDN. Group number two comprises attacks against the Metasploitable-2 machine, while the final group indicates regular network activities. This ultimate dataset comprises 84 different features derived using the CICFlowMeter tool. For this research, a subset of 33 optimal features is exclusively utilized for training the Ensemble learning model. Additional insights regarding these features are elaborated in [27]. Furthermore, the attack instances utilized during the testing phase differ in distribution from those employed in training. The training and testing datasets consist of 124,288 and 82,858 data records, respectively.

3.2 DATA PREPROCESSING

3.2.1 Load and Read dataset using Python code.

32.2 Understand the structure of data by checking it.

3.2.3 Check for missing values and no missing values are found.

3.2.4 Check all categories along with count in Target column. In this dataset there are six unique categories.

3.2.5 Concatenation of multiple datasets. In this research we have combined Normal_data.csv and OVS.csv files.

3.2.6 Drop Bad columns with the help of domain knowledge.

3.2.7 Separate target and features.

3.2.8 Split dataset into training and testing.

3.2.9 Compute the correlation matrix and drop feature columns having correlation greater than 70 %.

3.2.10 Initialize a model with Random Forest and also initialize RFECV. Fit RFECV to training data.

3.2.11 Reduce both the groups namely training and Test to selected features.

3.2.12 Using normalization technique name Z-Score for standardizing feature values.

3.2.13 Combine pre-processing steps for numeric and categorical features.

3.2.14 Finally pre-process the Training data.

4. DDOS DETECTION MODEL

4.1 Random Forest (RF)

RF algorithm functions by operating as an ensemble technique in machine learning, combining forecasts originating from a variety of decision trees. This method is suitable for tasks involving classification as well as regression. The predictions generated by each decision tree in a RF are brought together to produce the final projection. The individual decision trees in a Random Forest are trained utilizing unique subsets of the available data. By employing ensemble strategies, the RF algorithm improves the precision of individual decision trees, thus establishing a higher level of dependability especially in the context of DDoS attack identification.

4.1.1 Classification: When formulating a prediction, every individual tree within the forest generates a classification label. The ultimate prediction is determined through a collective decision-making process amongst all the trees. The class that is most commonly observed is identified as the anticipated class.

For an input x, each tree T_i gives a predicted class $\widehat{y_i}$. The endmost forecast $\hat{y}$ is the mode of entire forecasts:

$$\hat{y}= \text{mode}\ (\{\widehat{y_1},\widehat{y_2},\ldots\ldots,\widehat{y_B}\})$$

4.1.2 Regression: In regression tasks, every tree generates a numerical output. The ultimate forecast is computed as the mean of the various predictions produced by the individual trees.

For an input x, each tree T_i gives a predicted class $\widehat{y_i}$. The endmost forecast $\hat{y}$ is the average of entire forecasts: $\hat{y} = \frac{1}{B}\sum_{i=1}^{B}\hat{y}_i$

5. PROPOSED WORK

Here we purpose a combination of ML based model as RF with RFECV (RF-RFECV). Table 1 display the comparison among previous studies and our purposed model.

References	Methodologies	Feature Selection method	Dataset	Accuracy (%)
[15]	SOM+SVM, SVM, KNN, NB	Self-Organizing Map (SOM)	CAIDA 2016	SVM-SOM Best Classifier 98.12
[16]	SVM, Decision tree, KNN, Random Forest	Not Mention in Research	SDN DDOS dataset	RF Best Classifier 99.99
[17]	RF, SVM, XGBoost, DT, KNN	Improved Binary Grey Wolf Optimization (BGWO)	CSE-CIC-IDS2018	RF Best Classifier 99.13
[18]	NN, NB, RF, KNN and SVM	Analysis of Variance (ANOVA), F-test	Customized	RF Best Classifier 98.70
[20]	Gradient Boosting, RF, WVE, KNN and Linear Regression	MI, RFFI	CIC-IDS2017, CIC-DDoS2019	RF Best Classifier 99.99
[21]	RF, DT, SGD, CNN, and NGBooST	Filter- correlation and chi-square, SelectBest method, univariate selection strategies	CICDDoS2019	RF Best Classifier 99
[22]	SVM-RF	PCA, t-SNE	Customized	98.8
[23]	SVM	KPCA, GA	NSL-KDD	98.9
[24]	Random Forest, Linear SVM, NB, DT	Chi-square test	CIC-IDS2018	RF-99.83, DT-99.97
[25]	XGBoost, RF, KNN, CNN, NB, Logistic Regression, AdaBoost	ANOVA, F-test ExtraTreeClassifier, Logistic Regression	CIC-DDoS2019	RF Best Classifier 99.99
[26]	Convolutional neural network (CNN)	L2 regularization and the dropout methods	InSDN	93.01
Proposed Work	Random Forest	Z-Score normalization, RFECV	InSDN	RF Best Classifier 99.99

Table 1: Comparison among earlier and proposed researches

6. PERFORMANCE EVALUATION

A confusion matrix works as a fundamental instrument used in the domain of ML for the appraisal of a classification model's effectiveness. This matrix takes the form of a tabular display, encapsulating a concise overview of the anticipated outcomes vis-à-vis the factual results stemming from a classification assignment. By offering a comprehensive portrayal of the algorithm's efficacy, confusion matrix delves into the nuances of the errors it encounters during its operation. In case of binary classification challenge, confusion matrix is a 2x2 table, which includes four key elements.

True Positive (TP): The unit of occurrences precisely forecasted as belonging to the positive category.

True Negative (TN): The unit of occurrences precisely forecasted as the negative category.

False Positive (FP): The quantity of occurrences inaccurately forecasted as the positive category (commonly referred to as Type I errors).

False Negative (FN): The tally of occurrences that are erroneously categorized as the negative category, also known as Type II errors.

Confusion Matrix		
Scenarios	DDoS	Not a DDoS
DDoS Assaults	**TP**	**FP**
Not a DDoS Assaults	**FN**	**TN**

Table 2: Confusion Matrix

Derived Metrics

For performance measurement of the final model, several metrics are acquired with the help of confusion matrix.

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

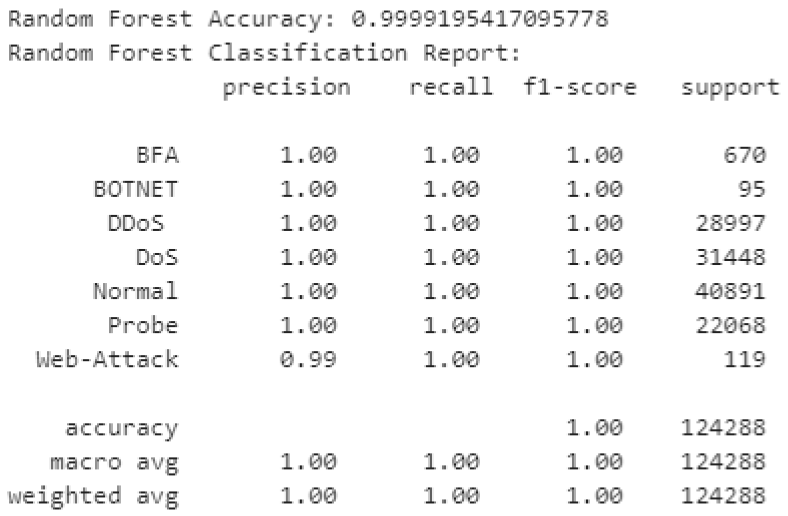
Precision $= \frac{TP}{TP+FP}$

Recall $= \frac{TP}{TP+FN}$

F1-S $= 2 \times \frac{Precision \times Recall}{Precision+Recall}$

7. RESULTS AND DISCUSSION

Performance of ML algorithm in the present research is reviewed applying F1 Score, Accuracy, Precision and Recall parameters. In the proposed work our RF model has achieved 99.9920 % Accuracy. Category wise Classification report is shown in Figure 3**.** Figure 4 is indicating confusion matrix. Error rate is observed 0.008045 %.



```
Random Forest Accuracy: 0.9999195417095778
Random Forest Classification Report:
              precision    recall  f1-score   support

         BFA       1.00      1.00      1.00       670
      BOTNET       1.00      1.00      1.00        95
        DDoS       1.00      1.00      1.00     28997
         DoS       1.00      1.00      1.00     31448
      Normal       1.00      1.00      1.00     40891
       Probe       1.00      1.00      1.00     22068
  Web-Attack       0.99      1.00      1.00       119

    accuracy                           1.00    124288
   macro avg       1.00      1.00      1.00    124288
weighted avg       1.00      1.00      1.00    124288
```

Figure 3: Category wise classification report

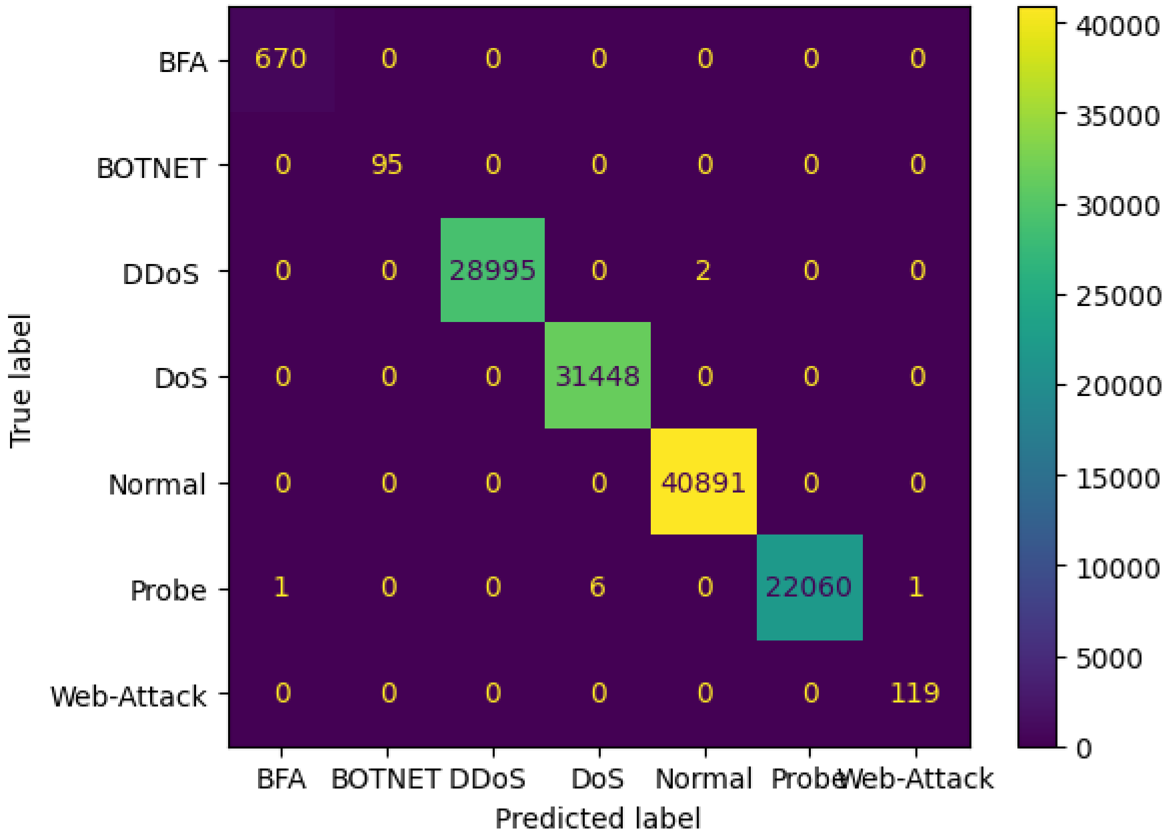


Figure 4: Confusion Matrix

8. CONCLUSION AND FUTURE WORK

This investigation showcases the effectiveness of supervised ML technique RF in conjunction with feature selection methodology like Recursive Feature Elimination with Cross-Validation (RFECV). Through the integration of these methodologies within SDN framework, the study reveals promising outcomes in bolstering detection precision and adaptability against advancing DDoS tactics. RF-RFECV model has achieved 99.991954 % accuracy with an error rate of 0.008045 %. In our prospective endeavors, we aim to enhance and broaden the existing techniques for recognizing DDoS assaults in SDN based networks. This encompasses the application of sophisticated feature-selection methodologies and ML frameworks, like Deep Learning, for the detection of Slow Loris and HTTP-centric DDoS intrusions. Furthermore, our investigation will delve into strategies for mitigating hostile intrusions to enhance the resilience and flexibility of the methodology. Moreover, we envisage the integration of this approach with more technologies in the field of security to formulate a more all-encompassing and thorough security resolution.

REFERENCES

[1] R. Sahba, “A Brief Study of Software Defined Networking for Cloud Computing,” *World Autom. Congr. Proc.*, vol. 2018-June, pp. 6–9, 2018, doi: 10.23919/WAC.2018.8430419.

[2] A. Mohamed *et al.*, “Software-defined networks for resource allocation in cloud computing: A survey,” *Comput. Networks*, vol. 195, p. 108151, 2021, doi: 10.1016/j.comnet.2021.108151.

[3] M. B. Yassein, S. Aljawarneh, M. Al-Rousan, W. Mardini, and W. Al-Rashdan, “Combined software-defined network (SDN) and Internet of Things (IoT),” *2017 Int. Conf. Electr. Comput. Technol. Appl. ICECTA 2017*, vol. 2018-Janua, pp. 1–6, 2017, doi: 10.1109/ICECTA.2017.8252003.

[4] S. Bera, S. Misra, and A. V. Vasilakos, “Software-Defined Networking for Internet of Things: A Survey,” *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1994–2008, 2017, doi: 10.1109/JIOT.2017.2746186.

[5] C. N. Tadros, M. R. M. Rizk, and B. M. Mokhtar, “Software Defined Network-Based Management for Enhanced 5G Network Services,” *IEEE Access*, vol. 8, pp. 53997–54008, 2020, doi: 10.1109/ACCESS.2020.2980392.

[6] A. Tarek, B. Mohammed, R. Barakat, Y. Eid, S. El-Kaliouby, and N. AbdElbaki, “Software-Defined Networks Towards Big Data: A Survey,” *Adv. Intell. Syst. Comput.*, vol. 1339, no. November, pp. 626–636, 2021, doi: 10.1007/978-3-030-69717-4_59.

[7] L. Cui, F. R. Yu, and Q. Yan, “When big data meets software-defined networking: SDN for big data and big data for SDN,” *IEEE Netw.*, vol. 30, no. 1, pp. 58–65, 2016, doi: 10.1109/MNET.2016.7389832.

[8] N. Mckeown, “Software-defined Networking,” no. April, 2009.

[9] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, “A Survey on Software-Defined Networking,” *IEEE Commun. Surv. Tutorials*, vol. 17, no. 1, pp. 27–51, 2015, doi: 10.1109/COMST.2014.2330903.

[10] H. Kim and N. Feamster, “Improving network management with software defined networking,” *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 114–119, 2013, doi:

10.1109/MCOM.2013.6461195.

[11] B. A. A. Nunes, M. Mendonca, X. N. Nguyen, K. Obraczka, and T. Turletti, "A survey of software-defined networking: Past, present, and future of programmable networks," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 3, pp. 1617–1634, 2014, doi: 10.1109/SURV.2014.012214.00180.

[12] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14–76, 2015, doi: 10.1109/JPROC.2014.2371999.

[13] N. Innab and A. Alamri, "The Impact of DDoS on E-commerce," *21st Saudi Comput. Soc. Natl. Comput. Conf. NCC 2018*, pp. 1–4, 2018, doi: 10.1109/NCG.2018.8593125.

[14] L. F. Eliyan and R. Di Pietro, "DoS and DDoS attacks in Software Defined Networks: A survey of existing solutions and research challenges," *Futur. Gener. Comput. Syst.*, vol. 122, pp. 149–171, 2021, doi: 10.1016/j.future.2021.03.011.

[15] V. Deepa, K. M. Sudar, and P. Deepalakshmi, "Design of Ensemble Learning Methods for DDoS Detection in SDN Environment," *Proc. - Int. Conf. Vis. Towar. Emerg. Trends Commun. Networking, ViTECoN 2019*, no. February 2023, 2019, doi: 10.1109/ViTECoN.2019.8899682.

[16] R. Anusuya, C. Prathima, M. Ramkumar Prabhu, and J. R. Arun Kumar, "Detection of TCP, UDP and ICMP DDOS attacks in SDN Using Machine Learning approach," *Journal of Survey in Fisheries Sciences*, vol. 10, no. 4S. p. 2023.

[17] Z. Liu, Y. Wang, F. Feng, Y. Liu, Z. Li, and Y. Shan, "A DDoS Detection Method Based on Feature Engineering and Machine Learning in Software-Defined Networks," *Sensors*, vol. 23, no. 13, Jul. 2023, doi: 10.3390/s23136176.

[18] A. Maslan, K. M. Bin Mohamad, and F. B. Mohd Foozy, "Feature selection for DDoS detection using classification machine learning techniques," *IAES Int. J. Artif. Intell.*, vol. 9, no. 1, pp. 137–145, 2020, doi: 10.11591/ijai.v9.i1.pp137-145.

[19] M. Alkasassbeh, G. Al-Naymat, A. B.A, and M. Almseidin, "Detecting Distributed Denial of Service Attacks Using Data Mining Techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 1, pp. 436–445, 2016, doi: 10.14569/ijacsa.2016.070159.

[20] M. Alduailij, Q. W. Khan, M. Tahir, M. Sardaraz, M. Alduailij, and F. Malik, "Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method," *Symmetry (Basel).*, vol. 14, no. 6, pp. 1–15, 2022, doi: 10.3390/sym14061095.

[21] M. S. Raza, M. N. A. Sheikh, I.-S. Hwang, and M. S. Ab-Rahman, "Feature-Selection-Based DDoS Attack Detection Using AI Algorithms," *Telecom*, vol. 5, no. 2, pp. 333–346, Apr. 2024, doi: 10.3390/telecom5020017.

[22] N. Ahuja, G. Singal, D. Mukhopadhyay, and N. Kumar, "Automated DDOS attack detection in software defined networking," *J. Netw. Comput. Appl.*, vol. 187, no. May, p. 103108, 2021, doi: 10.1016/j.jnca.2021.103108.

[23] K. S. Sahoo *et al.*, "An Evolutionary SVM Model for DDOS Attack Detection in Software Defined Networks," *IEEE Access*, vol. 8, pp. 132502–132513, 2020, doi: 10.1109/ACCESS.2020.3009733.

[24] T. K. Luong, T. D. Tran, and G. T. Le, "DDoS attack detection and defense in SDN based on machine learning," *Proc. - 2020 7th NAFOSTED Conf. Inf. Comput. Sci. NICS 2020*, pp. 31–

35, 2020, doi: 10.1109/NICS51282.2020.9335867.

[25] A. Golduzian, “Predict And Prevent DDOS Attacks Using Machine Learning and Statistical Algorithms,” vol. 12, p. 2022, 2022.

[26] M. S. Elsayed, H. Z. Jahromi, M. M. Nazir, and A. D. Jurcut, “The Role of CNN for Intrusion Detection Systems: An Improved CNN Learning Approach for SDNs,” *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST*, vol. 382, no. May, pp. 91–104, 2021, doi: 10.1007/978-3-030-78459-1_7.

[27] M. S. Elsayed, N. A. Le-Khac, and A. D. Jurcut, “InSDN: A novel SDN intrusion dataset,” *IEEE Access*, vol. 8, pp. 165263–165284, 2020, doi: 10.1109/ACCESS.2020.3022633.