A Comparative Study of AI-enabled DDoS Detection Technologies in SDN

Kwang-Man Ko , Jong-Min Baek , Byung-Suk Seo , Wan-Bum Lee *

Posted Date: 9 August 2023

doi: 10.20944/preprints202308.0700.v1

Keywords: SDN, DDoS attacks; Deep learning; Machine learning; Permutation Importance

Article

A Comparative Study of AI-Enabled DDoS Detection Technologies in SDN

Kwang-Man Ko [1,†], Jong-Min Baek [2,†], Byung-Suk Seo [3,†] and Wan-Bum Lee [4,*,†]

1 Dept. of Computer Engineering, SangJi University; kkman@sangji.ac.kr
2 Dept. of Computer Engineering, SangJi University; a01014@sj.sangji.ac.kr
3 Dept. of Computer Engineering, SangJi University; seobs@sangji.ac.kr
4 Dept. of Computer and Software Engineering, WonKwang University; lwbwon@wku.ac.kr
* Correspondence: lwbwon@wku.ac.kr
† Dr. Kwang-Man Ko, the first author, designed the entire system for DDoS detection and devised a specific solution based on various previous researches. The second author, Mr. Jong-Min Baek, investigated ML and DL learning models, conducted various experimental studies, and then performed initial writing. The third author, Dr. Byung-Suk Seo, analyzed the contents of the study coolly and designed the function of each component of the entire system. In addition, he read the written contents of the paper and performed detailed reviewing and editing. Dr. Wan-Bum Lee as a corresponding author, was in charge of funding acquisition this study and verifying the experimental results.

Abstract: Software Defined Networking (SDN) is positioning the standard for the management of networks due to its scalability and flexibility to program the network [1]. The SDN provides many advantages but it also involves some specific security problems take down the controller using cyber attack and in result the whole network will shut down which makes it a single point of failure. In this paper, the DDoS attacks in SDN were detected using AI-enabled, machine and deep learning, models with some specific features for data-set under normal and DDoS traffic. In our approach, initial data-set is collected from 84 features on kaggle and then the 20 top most features are selected using permutation importance algorithm. The data-set were learned and tested with AI-enabled 5 models. Our experimental results showed that the use of machine learning based random forest model has achieved the highest accuracy rate of 99.97%, in DDoS attack detection in SDN. Our contributions through this study are, first, we found highest 20 attacks that absolutely contributed to DDoS attacks. Secondly, it can reduce the time and cost of comparing various learning models and performance required for determining a learning model suitable for DDoS detection. Finally, various experimental methods for evaluating the performance of the learning model are presented so that related researchers can utilize them.

Keywords: SDN; DDoS attacks; deep learning; machine learning; permutation importance; random forest

1. Introduction

The SDN has gained the popularity due to the services and benefits it provides a scalability, flexibility, and monitoring [2]. In past decade, devices over internet were increased enormously which rises different problems with traditional networks and one of them is the management of networks. To resolve this issue a new paradigm Software Defined Networking was proposed. The whole network management and configuration is performed through controller which simplified the network management [3]. The separation of control and data plane plays a critical role by providing high performance in large-scale network systems, but the simplified network management comes with a cost of centralization. In an SDN environment, one can simply enforce policy and network configurations in real-time through the controller [4].

However, the controller is a single point which has a ability to control the whole network and if a controller is compromised then whole network is under attack. The SDN controller is the reason which reveals it to a variety of security threats, among them is DDoS attacks. A DDoS can make the

 doi:10.20944/preprints202308.0700.v1

controller or OpenFlow switch overwhelm if the network is not reasonably secured. In the DDoS attack a lot of boughs requests send to controller which makes the network slow and effect the legitimate traffic. The DDoS attacks also expose the data plane through the flow table. To protect the SDN controller from DDoS attacks, the intrusion detection systems have been used in the network to sniff the packets and alert the administrator when a DDoS attack is detected, and many researchers have done research on detecting and mitigating the DDoS attack using different techniques. Recently, The machine and deep learning-based approaches are more dynamic, efficient, and intelligent solutions for SDN management. In this paper we have studied the both approaches, which target to determine the most suitable artificial intelligent algorithm to detect a DDoS attack in the SDN.

The goal of this paper is to detect the DDoS attack in SDN environment without compromising the security and effecting the legitimate traffic for ensuring the high-level security against DDoS attacks on SDN. We have achieved this through applying different machine and deep learning algorithms. Firstly, the traffic is analyzed by SDN controller rules set by network administrator. It will further analyze the traffic and divide into legitimate and attack traffic using different algorithms. We have used machine learning based random forest, decision tree and naïve bayes algorithms and deep learning based convolution neural network, recurrent neural network for classification of traffic. For learning and testing the learning model, we used the Kaggle data set to check DDoS detection performance. Finally, the 20 data sets that affect DDoS detection were determined by applying the 2-step Permutation Importance Algorithm(PIA) to the original Kaggle data.

The rest of the paper is organized into the following sections. Section 2 have shown the research background and related work previously done regarding detection of DDoS attacks in SDN. Section 3 shows our methodology and model overview in detail for experiments. Section 4 shows the experiments and analysis, which also include experimental environment, and finally the section 5 is based on conclusion and future works, respectively.

2. Research Background and Motivations

2.1. Backgrounds

Since the Software defined networking has simplified the network management and providing the easiest way to manage the network infrastructure through SDN controller and programmable switches. Still SDN have some of major issues and researchers are continually trying to overcome them. We have discussed briefly related to SDN, DDoS detection in subsections. The main component of SDN is a controller which control the whole infrastructure of network and provide simplified network management. There are three layers in SDN architecture; Firstly, the application layer which provide the facility for the network to interact with application. Secondly, the control plane layer, it is considered as the brain of SDN because it controls the network flow. Lastly, the infrastructure layer, it is responsible for traffic forwarding. The preliminary step which took place before the actual packet forwarding in SDN network includes the discovery of topology [5]. The SDN controller kept the updated information related to the data plane using the OpenFlow discovery protocol. While all other network devices using Link Layer Discovery Protocol advertise their identities and neighbors in the network.

In [6,7] author said that "A DDoS is an attack on a server where a massive number of packets are sent to create an outage or service degradation for legitimate user". According to [8], the attacker's main focus is on the resource consumption and bandwidth reduction. So, detection of these attacks is necessary to protect the network. The attacker is attacking the two main targets such controllers and network elements. To protect the network from DDoS attacks, we must monitor and analyze the network traffic to identify the malicious traffic. A Detection of DDoS attack is very complicated because it pretends as normal traffic. Even if you identify the attack traffic you cannot block the attacker IP because of IP Spoofing. When the SDN suffers from DDoS attacks, the switch eventually loses its connection to the controller. It will look for another controller if there is any standby controller

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

available and then it will connect with it [9–11]. On the time-based techniques, the time characteristics is an important factor in attacks like DDoS. In [12] proposed technique that uses the time duration to detect DDoS attacks and used the time pattern to prevent attacks in future. The solutions like this are not the proper answer to the current attacks. The small window and short-term statistic can be used in SDN network as [13] have proposed an entropy-based DDoS detection that used in non-SDN network. This method uses the randomness to calculate the number of incoming packets to specific hosts [14] and then, it compares to a threshold value. The detection is based on a comparison results. In [15] have proposed an entropy-based technique to detect DDoS attack but there is no solution to mitigate it for future. There is need of some intelligent based algorithms to detect the attacks.

Table 1. Comparison of different methods used for DDoS attacks in SDN.

Method	Advantages	Issues
Time-Based DDoS Detection	Contains a mitigation process by creating time patterns to prevent future attacks. Reduces controller processing rate because the flow collector handles this task.	Causes time delay because the processor adds more processing for non-valid packets. Requires additional implementation
Entropy-Based DDoS Detection	Lightweight method. Detects DDoS attack at early stages. Flexible; can modify any parameter.	Additional overhead from window size. Unable to detect DDoS attacks on multiple hosts No prevention technique.
Machine Learning- based DDoS Detection	Introduces a high overhead reduction compared to other techniques. - Less CPU loads. - Fast detection time. - ability to monitor multiple points instead of one.	This technique is not implemented with normal switches. It requires an additional implementation.
Deep Learning- based DDoS Detectio	It provides accurate information on anomalous behavior. It is able to evaluate big sets of data, with less time and CPU load.	Can be implemented on network system programmable. Require additional functionalities

2.2. Related works and Research Motivations

Therefore, we have discussed some machine learning and deep learning based solutions. They [16] have the used different machine learning algorithms for detection of DDoS attack in SDN environment and compare the results. [17] they have provided the survey of current research related to the security of the SDN paradigm. In [18], proposed a lightweight DDoS attack detection mechanism. In [19] energy-based model was proposed, recurrent, conventional and fully connected and reported F-measure was 73.9% and 73.2%. In [20] Deep neural network is used for anomaly detection with 75% F-measure. In [21], three hidden layer and one output layer. Two hidden layers train the model using Autoencoder and classification was taking place in last layer using SoftMax. They reported 97 % F-measure. In [22] they have designed FCN, VAE, LSTM- Seq2Seq Structure for anomaly detection in network and stated that LSTM- Seq2Seq Structure yields 99% classification accuracy. In [23], a study was conducted to detect DDoS in a cloud environment by applying multiple regression analysis to the CICIDS 2017 benchmark data-set. In order to detect DDoS attacks in real time, big data access methods have been studied [24], and research to detect Botnet attacks has also progressed [25,26]. Recently, research on detecting and mitigating DDoS attacks based on machine learning [27–30] and deep learning [31–33] has been conducted.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

The motivations of our study are to confirm an excellent DDoS detection performance model by comparing existing machine learning and deep learning models in SDN environment. The fundamental intention of our study is to reduce the developer's initial burden by presenting a comparison of the detection performance of the existing model when developing a new AI-enabled DDoS detection model. Initially, we applying 84 DDoS candidate date-set in the original Kaggle data-set, the data-sets affecting detection were reduced to 64 and 20. After analyzing the DDoS data-set collected from the first Kaggle, 84 candidate data-set were manually selected. From the next step, the meaningful data-set for DDoS attack was determined step by step using the permutation importance algorithm. In this process, it was confirmed that even if various types of DDoS attacks occur, the top 20 attacks that affect the performance of SDN are affected. Through this, it can be used to construct a data-set that affects the design of a new learning model for DDoS attack detection. In many related studies, the DDoS detection performance and time are the mainstream. So, we conducted a study that corely compares the DDoS detection performance and time. These studies provide the basis for selecting an appropriate learning model for subsequent researchers. It also provides DDoS attack information that has a fatal impact on SDNs.

3. Model Overview and Methodology

3.1. Model Overview

The DDoS have become more complex and challenging with the emergence of new technologies. Generally, network traffic composes of abnormal and maliciously traffic. This traffic needs to be monitored and analyzed by organizations to prevent violation of policies and protect against attacks. Major approach that has been popular in recent years in the research community is the use of machine learning techniques in SDN. There are machine learning based techniques have been used to develop Network Intrusion Detection Systems. As the deep learning based technique has both supervised and unsupervised learning qualities. As you can see, convolution neural network and recurrent neural network as classifier model which also are being utilized for detection of DDoS attack. In this paper, in order to evaluate DDoS detection performance, feature was first selected from 84 data-set and 20 key attacks were finally selected as shown in Figure 1.

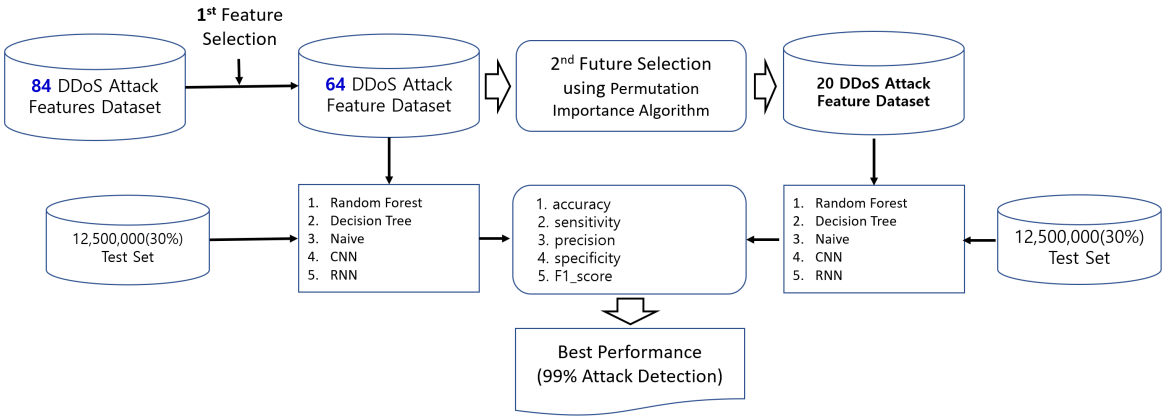

Figure 1. The process steps for applying the Feature Selection method using ML & DL model.

Figure 2 shows the structure and parameters of CNN and RNN.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

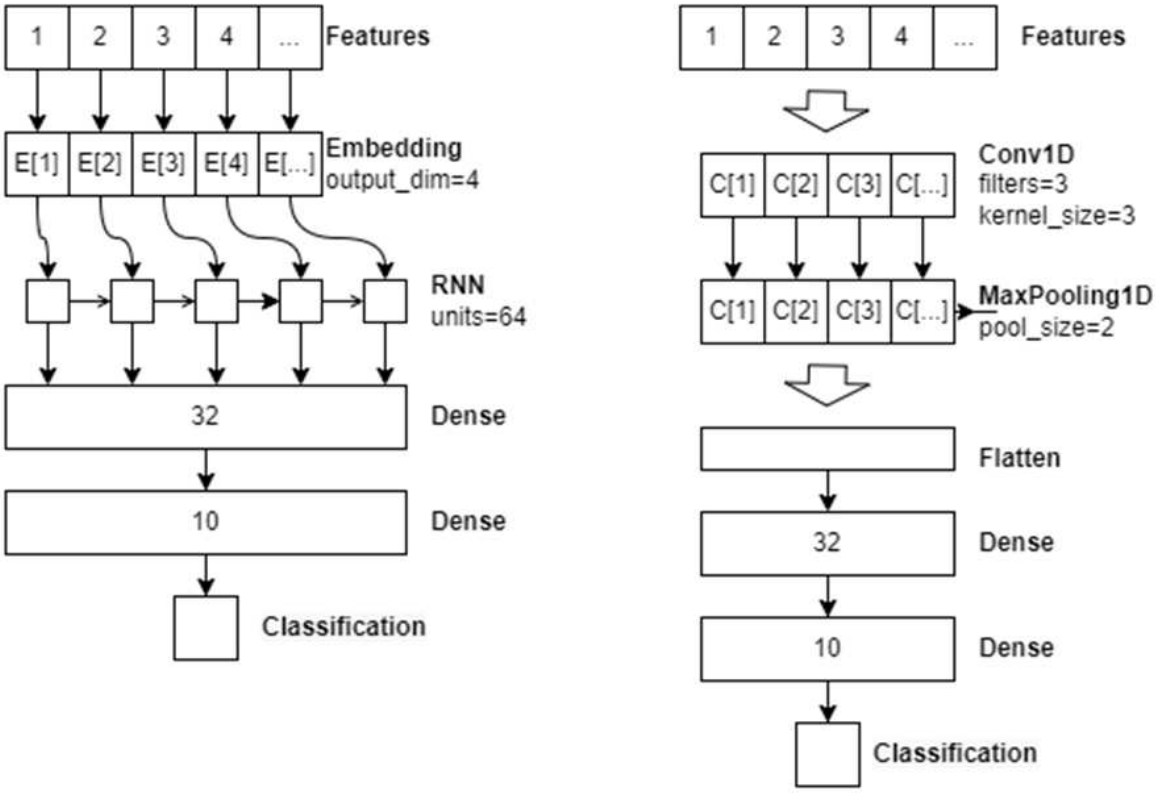

Figure 2. RNN and CNN Layer Structure.

3.2. Experimental Data-set

In this section, we describe how the data-set is obtained from the Kaggle [34] and the features of data-set are explained. In which DDoS based traffic containing 84 features, included 83 features and 1 classifier. To calculate every feature in every flow per second cause overhead to our model. This is why, we use our feature selection method, permutation importance algorithm, to reduce the data-set with high priority 20 topmost features with 12500,000 samples that are vital for SDN architecture for classification of traffic to our classifier model. In the literature, The recent researches suggest that there is no universal best model for classification tasks. In this study, we have analysis the performances of detecting of DDoS attacks from Random Forest, Decision Tree, Naïve Bayes, CNN and RNN on the obtained feature set with original and reduced features, respectively.

A DDoS and normal traffic data-set were captured to avoid confusion in labelling data-set. To detect DDoS attacks, the permutation importance algorithm is used for extracting features from data-set. The data-set used in the training test of machine learning and deep learning based models can contain a large number of features. The goal is to reduce the low impact features on the classification and to provide the highly effective features. The data-set of DDoS traffic, we got from Kaggle. In which DDoS based traffic containing 84 features, included 83 features and 1 classifier. Using permutation importance algorithm, we selected the best 20 features [35] from our 84 data-set to train in our classifier models, respectively shown in Table 2. The raw data capture undergo pre-processing to make it suitable to train the proposed model and avoid over fitting.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

Table 2. The Top 20 Features using Permutation Importance algorithm.

#	Feature	Importance	Description
1	fl_iat_avg	0.5334	Two flows average time
2	fw_iat_max	0.3351	Maximum time of two packets sent
3	fw_win_byt	0.3248	Number of bytes
4	fw_iat_tot	0.3231	Total time of two packets sent
5	fl_dur	0.3162	Time of duration
6	fl_iat_min	0.2778	Minimum time of two flows
7	bw_iat_min	0.2655	Minimum time of two packets sent
8	fl_iat_max	0.2426	Maximum time of two flows
9	fw_iat_avg	0.2203	Mean time of two packets sent
10	Bw_pkt_l_max	0.2139	Maximum size of packet
11	bw_iat_max	0.2077	Maximum time of two packets sent
12	bw_win_byt	0.1944	Number of bytes sent
13	bw_iat_tot	0.1942	Total time of two packets sent
14	fw_iat_min	0.1520	Minimum time of two packets sent
15	bw_iat_avg	0.1513	Mean time of two packets sent
16	idl_max	0.1139	Flow maximum time before becoming active
17	bw_seg_avg	0.0926	Average size observed
18	Bw_pkt_l_avg	0.0922	packet average size
19	fw_pkt_l_avg	0.0913	Packet average size
20	pkt_size_avg	0.0910	Packet average size

3.3. Experimental Environment

Our DDoS detection model is based on an SDN based topology which is built on Mininet for simulating the results of experiments as following in Figure 3. Our system topology is consisting of six PCs, one OpenFlow Switch and an RYU SDN controller. Among the six PCs, one is an attacker PC and other is an attacker PC, while other four PC generate normal traffic. To detect the attacker traffic, we use RF, NB, DT, CNN and RNN classifier to implement on controller for testing and training the traffic. The deep learning is basically derived from machine learning and the researchers found that machine learning performs better in DDoS attack scenarios.

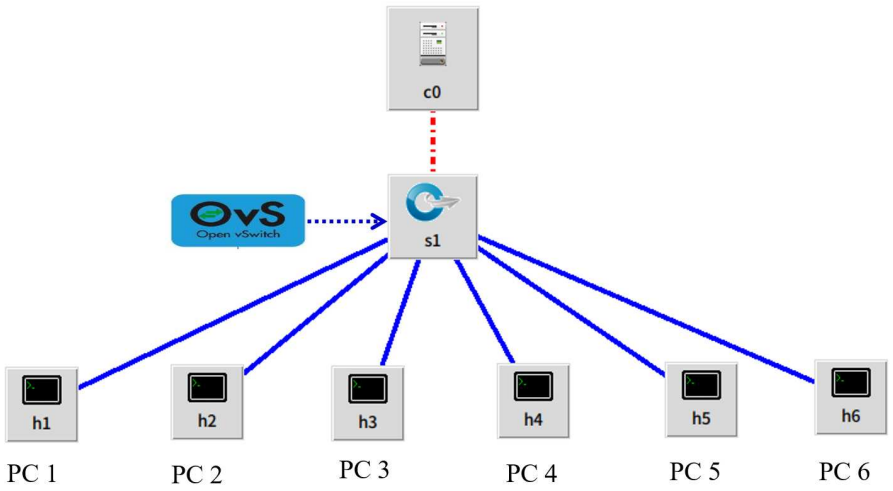

Figure 3. An SDN topology on Mininet for DDoS attack testing.

Mininet is a network emulator which provide the functionality to create a virtual network environment which have ability to communicate with virtual devices using virtual links [36]. It also provides capabilities to integrate with different SDN controller such as RYU SDN controller based on python. Mininet provide some of amazing advantages which makes it ideal choice such as it supports for OpenFlow protocol and is capable of running linux application in virtual environment. RYU

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

controller is an open source component based SDN framework and provides software components with well-defined API that make it easy for developers to create new network management and control applications [37]. RYU supports different version of OpenFlow protocols, also it supports NETCONF and OF-config protocols [38]. Ryu uses scripts and OpenFlow protocol to communicate and manage the switches [39]. The entire experiments are carried out on Ubuntu (18.10) virtual machine setup on VMware with a 2GB of RAM, 200GB hard drive space and Mininet (ver2.3) is used with a RYU (ver4.3) controller. The SDN network contains the following units that RYU controller, OpenFlow switch, 6 PCs, in which one is attacker PC, and another one is Victim PC and all other PCs are normal PCs as shown in Figure 3. In the experimental first stage, classifiers were trained and tested with using all features in the dataset. In the second stage, permutation importance algorithm was used to select the most effective features in the entire dataset. The performance ratio was determined by RF, NB, DT, CNN, and RNN algorithms on the basis of selected features.

4. Experiments Analysis and Verification

4.1. Experimental Results

In order to obtain the high priority features, which have a high impact on the prediction, the feature selection method permutation importance was applied. The normal traffic data and a DDoS attack data were analyzed on our SDN architecture with the total dataset of 12,794,627 samples. The training dataset used is 7,676,776 samples which is 60% of total dataset. In testing the 5,117,851 (40% of total dataset) is used. The reduced feature obtained by the feature selection was applied to our classifier models and the parameters were the same as we used in our previous study. As the reduced feature set of 64, 15, 14, 13, 12,11, 10, 9, 8, 7, 6, 5 and 4 respectively were selected as per given classifier model. The features were trained by classifier algorithm based on RF, NB, DT,CNN and RNN model.

The obtained results are presented in Table 3. The parameters used for analyzing the traffic are Evaluation Time,Accuracy, Sensitivity, Precision and F1_Score. The F_Score F1 is the weighted average of Precision and Sensitivity. Finally, the specificity is the ability to assess unequivocally the analyse in the presence of components. The results obtained are showing an interesting fact that if we reduce the features cannot guarantee us the performance increase in parameters like, accuracy, sensitivity, precision, specificity and F1_score, however the evaluation is decreased by reducing features. To summarize the results obtained, we can say that obtaining the best features which have high impact on detection of DDoS attack can enhance our performance. The results of our DDoS detection model based on Random Forest model has achieved by 5 feature sets using permutation importance algorithm with 99.976% of accuracy and F1-Score.

Table 3. Performance Comparision of the machine learning and deep learning models.

Model	Features	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
Random Forest	5	5.09 s	99.976 %	99.974 %	99.978 %	99.978 %	99.976 %
Decision Tree	7	0.44 s	99.842 %	99.829 %	99.852 %	99.856 %	99.840 %
Naive Bayes	8	0.69 s	93.645 %	96.293 %	91.317 %	91.059 %	93.739 %
CNN	13	4.13 s	96.654 %	94.412 %	98.760 %	98.843 %	96.537 %
RNN	15	4.28 s	98.723 %	98.200 %	99.207 %	99.234 %	98.701 %

In the second part of study, we have analyzed each model performance to briefly provide the results of individual models with different features selected from dataset. We have started from 64 feature sets and then according to classifier model features are selected which range from 15 to 4. The classifier model based on machine learning and deep learning both have their own impact on the results. The results of individual classifier models are showed following Tables 4–8 respectively. The obtained results from the different classifier models with different feature sets are quite interesting. Every classifier model result show that the importance of selecting features. As we have seen from

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

Table 2 that Random Forest perform better with 5 feature set in detecting and mitigating the DDoS attacks in SDN environment. The individual results show that machine learning based models are performing better than deep learning models. The decision tree and random forest results outclass other classifier models with different feature sets. The deep learning-based model highest accuracy and F1-Score is 98.723% and 98.70% with RNN and with CNN model is 96.654% and 96.537% respectively. However, the machine learning-based Random Forest and Decision tree models have above 99% accuracy rate and above 99% F1_score.

Table 4. Performance of the Random Forest (RF) model with Permutation Importance algorithm.

Feature Set	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
64	7.85 s	99.979 %	99.978 %	99.980 %	99.980 %	99.979 %
12	5.66 s	99.986 %	99.985 %	99.986 %	99.987 %	99.986 %
11	5.57 s	99.986 %	99.985 %	91.986 %	99.987 %	99.986 %
10	5.07 s	99.985 %	99.984 %	98.986 %	98.986 %	99.985 %
9	5.07 s	99.978 %	99.977 %	99.978 %	99.979 %	99.978 %
8	4.88 s	99.976 %	99.975 %	99.977 %	99.978 %	99.976 %
7	4.94 s	99.976 %	99.975 %	99.977 %	99.978 %	99.976 %
6	5.09 s	99.977 %	99.975 %	99.978 %	99.978 %	99.976 %
5	5.09 s	99.976 %	99.974 %	99.978 %	99.978 %	99.976 %
4	4.59 s	99.908 %	99.862 %	99.952 %	99.953 %	99.907 %

Table 5. Performance of the Decision Tree (DT) model with Permutation Importance algorithm.

Feature Set	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
64	1.61 s	99.984 %	99.982 %	99.985 %	99.985 %	99.984 %
12	0.49 s	99.983 %	99.981 %	99.985 %	99.985 %	99.983 %
11	0.50 s	99.853 %	99.841 %	99.862 %	99.865 %	99.852 %
10	0.49 s	99.853 %	99.839 %	99.863 %	98.866 %	99.851 %
9	0.48 s	99.842 %	99.829 %	99.852 %	99.855 %	99.841 %
8	0.47 s	99.842 %	99.829 %	99.851 %	99.855 %	99.840 %
7	0.44 s	99.842 %	99.829 %	99.852 %	99.856 %	99.840 %
6	0.38 s	97.740 %	99.960 %	99.660 %	95.572 %	97.763 %

Table 6. Performance of the Naïve Bayes (NB) model with Permutation Importance algorithm.

Feature Set	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
64	7.85 s	99.979 %	99.978 %	99.980 %	99.980 %	99.979 %
12	0.89 s	93.637 %	96.300 %	91.297 %	91.036 %	93.732 %
11	0.88 s	93.645 %	96.293 %	91.317 %	91.059 %	93.739 %
10	0.82 s	93.645 %	96.293 %	91.317 %	91.059 %	93.739 %
9	0.71 s	93.645 %	96.293 %	91.317 %	91.059 %	93.739 %
8	0.69 s	93.645 %	96.293 %	91.317 %	91.059 %	93.739 %
7	0.59 s	92.108 %	93.124 %	91.100 %	91.117 %	92.101 %

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

Table 7. Performance of the Convolution Neural Network model (CNN) with Permutation Importance algorithm.

Feature Set	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
64	5.67 s	96.244 %	95.161 %	97.178 %	97.301 %	96.159 %
15	4.38 s	95.450 %	95.266 %	95.513 %	95.630 %	95.389 %
14	4.26 s	96.471 %	94.103 %	98.694 %	98.784 %	96.344 %
13	4.13 s	96.654 %	94.412 %	98.760 %	98.843 %	96.537 %
12	4.10 s	95.349 %	91.718 %	98.780 %	98.894 %	95.118 %
11	4.12 s	96.184 %	93.226 %	98.841 %	98.971 %	95.951 %
10	4.17 s	95.537 %	96.233 %	94.812 %	94.858 %	95.517 %
9	4.23 s	89.137 %	86.972 %	90.661 %	91.251 %	88.778 %

Table 8. Performance of the Recurrent Neural Network model (CNN) with Permutation Importance algorithm.

Feature Set	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
64	2.09 s	95.874 %	97.830 %	94.056 %	93.963 %	95.906 %
15	4.28 s	98.723 %	98.200 %	99.207 %	99.234 %	98.701 %
14	3.78 s	97.407 %	95.579 %	99.141 %	99.192 %	97.328 %
13	3.63 s	97.449 %	95.611 %	99.197 %	99.244 %	97.371 %
12	3.38 s	97.366 %	95.477 %	99.161 %	99.211 %	97.284 %
11	3.28 s	97.454 %	95.601 %	99.218 %	99.264 %	97.376 %
10	3.07 s	97.440 %	95.589 %	99.199 %	99.247 %	97.361 %
9	3.14 s	97.061 %	94.672 %	99.348 %	99.394 %	96.954 %
8	3.27 s	97.059 %	94.756 %	99.257 %	99.308 %	96.954 %

In the third part of study, we have compared the performance of our models on the basis of detection time of attack. In this study, we have analyzed the lowest time taken by our classifiers models to detect DDoS attack in SDN environment. The graph is present in Figure 4. The results are showing that the Decision Tree model has the lowest time taken for detecting the DDoS attack and then the Naïve Bayes. The Random Forest shows the highest time taken for detecting DDoS attack. In the fourth part of study, the accuracy and F1-Score has been compared to analyse the performance of individual classifier. The graph is present in Figure 5. The graph is showing the results occur after using classifier models. The accuracy and F1_score are the important parameters in detecting and mitigating DDoS attacks. The results depicting the highest percentage of accuracy and F1_score obtained by random forest and decision tree models. The CNN and RNN model's performance is less than the machine learning-based model except the Naïve Bayes model. The Random Forest and Decision tree models have achieved the accuracy and F1_score above 99%. The performance of both models is showing how much useful the both can be in detecting and mitigating DDoS attacks in SDN network.

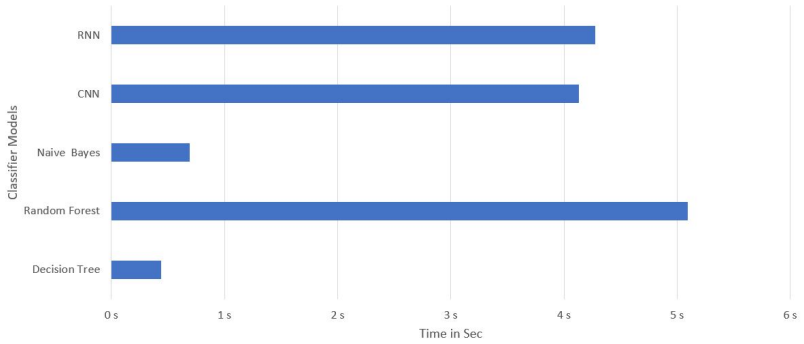

Figure 4. Performance comparison of Classifier models on detecting DDoS attack in SDN network.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

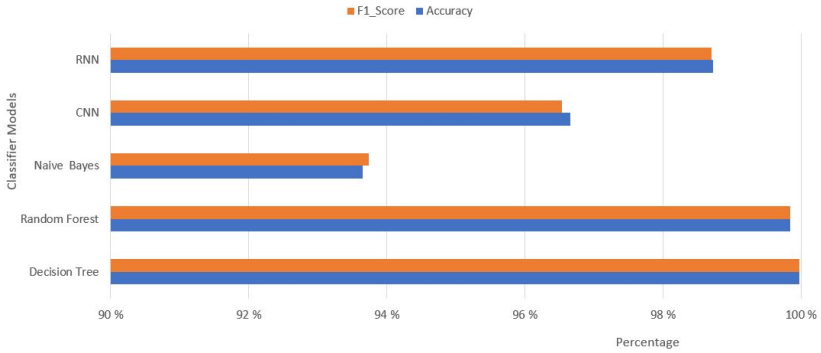

Figure 5. Performance comparison of Classifier models on accuracy and F1-Score.

The performance on the basis of remaining parameters like, sensitivity, precision, specificity has been evaluated and shown in graph in Figure 6. The results again showing the highest percentage of Random Forest and Decision Tree classifier models. After, analyzing overall and individual performance of our classifier models we can say that the machine learning-based Random Forest and Decision tree classifier model have shown the better performance in detecting and mitigating the DDoS attack in SDN network.

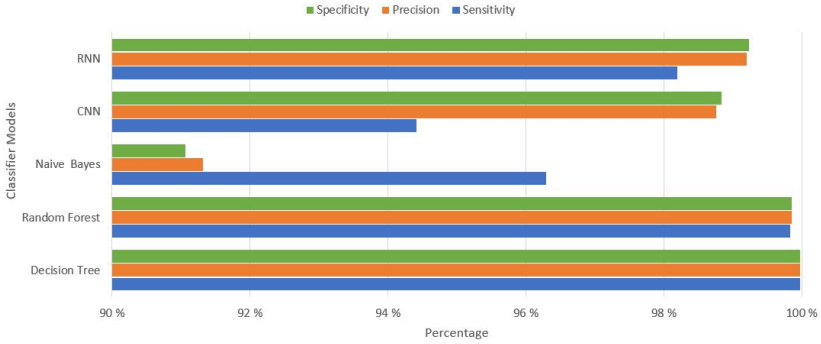

Figure 6. Performance comparison of Classifier models on sensitivity, Precision and Specificity.

As, we have analyzed the performance of classifier models with different feature sets. Now we will examine and analyze the performance of classifier models with 10 feature sets. At first, we will see the overall performance of our classifier models which is shown in Table 9. The results showing the same behavior and random forest is better than other classifier models.

Table 9. Performance of the machine learning and deep learning models with 10 selected features.

Model	Features	Evaluation Time	Accuracy	Sensibility	Precision	Specificity	F1_Score
Random Forest	10	0.49 s	99.853 %	99.839 %	99.863 %	99.866 %	99.851 %
Decision Tree	10	5.07 s	99.985 %	99.984 %	99.986 %	99.986 %	99.985 %
Naive Bayes	10	0.82 s	93.645 %	96.293 %	91.317 %	91.059 %	93.739 %
CNN	10	4.17 s	95.537 %	96.233 %	94.812 %	94.858 %	95.517 %
RNN	10	3.07 s	97.440 %	95.589 %	99.199 %	99.247 %	97.361 %

Now, we will present the performance of our classifier models according to evaluation time, accuracy and F1-Score, and with sensitivity, specificity, and precision as we have studied prior in the section. The purpose of selecting 10 features and showing the results is that to describe there is no huge impact on the performances of classifier models. The performance graphs of classifier models with respect to evaluation time, accuracy and F1-Score, and with sensitivity, specificity, and precision have been presented in Figures 7–9 respectively. However, the training carried out with the CNN model for detecting and mitigating the DDoS attack in SDN network, we have presented the performance of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

network with the help of graphs. The graphs show the parameters as bandwidth and time to show the performance of our traffic, without attack, with attack but without mitigation, and with attack but without mitigation, respectively. The purpose of presenting the graphs is that to show the impact of our CNN model for detecting and mitigating the DDoS attack in an SDN environment. The graphs are shown in Figures 10–12 respectively.

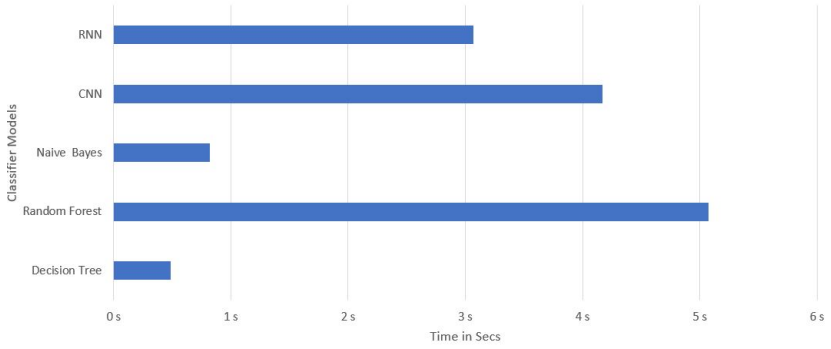

Figure 7. Performance comparison of Classifier models on detecting DDoS attack in SDN network with 10 features.

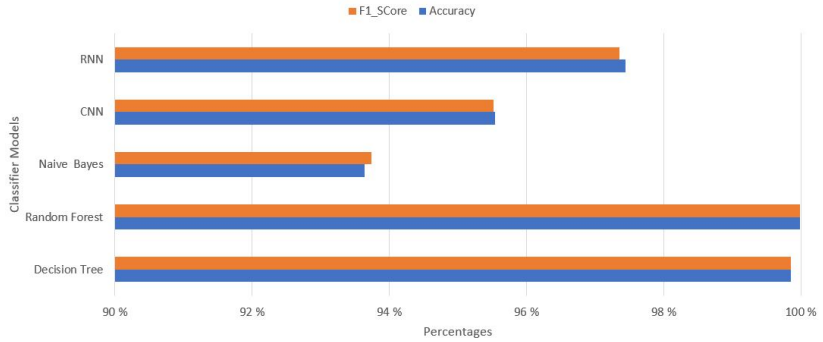

Figure 8. Performance comparison of Classifier models on accuracy and F1-Score with 10 features.

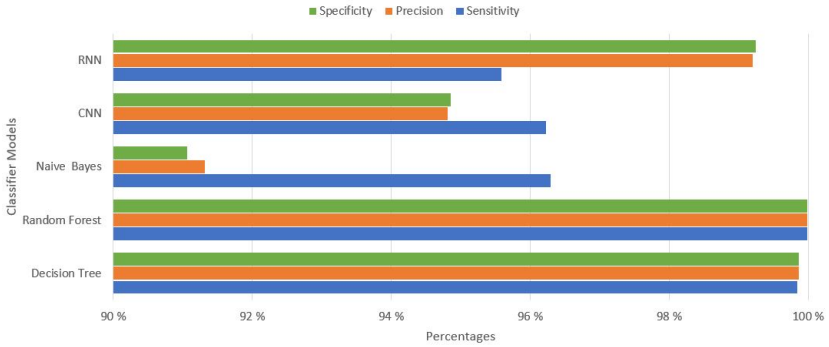

Figure 9. Performance comparison of Classifier models on sensitivity, Precision and Specificity with 10 Features.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 August 2023 doi:10.20944/preprints202308.0700.v1

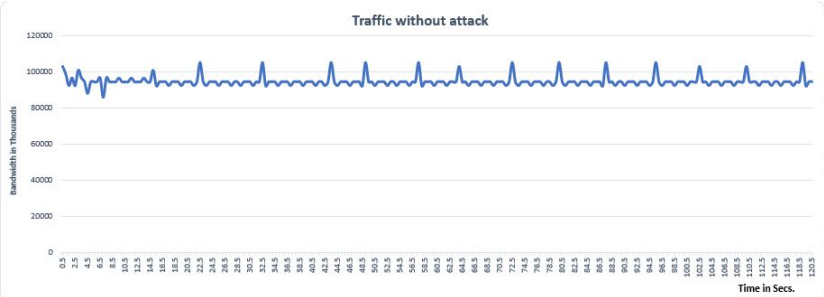

Figure 10. The performance of traffic without attack in SDN Network.

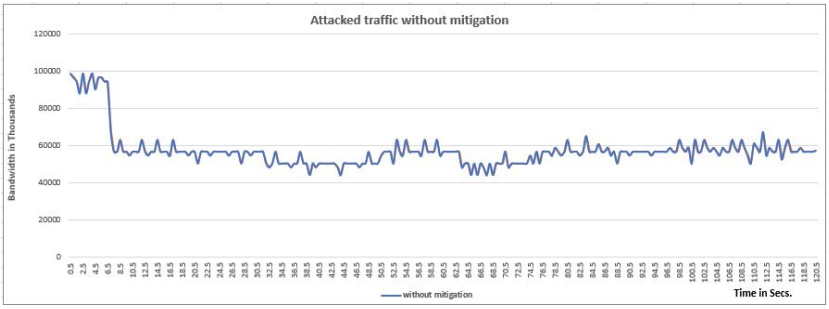

Figure 11. The performance of attacked traffic without mitigation in SDN Network.

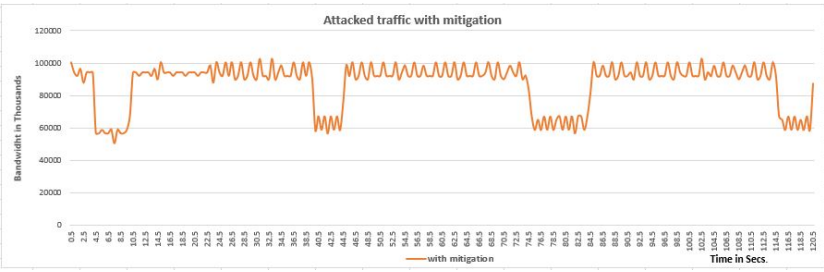

Figure 12. The performance of attacked traffic with mitigation in SDN Network.

The results show that the SDN architecture can be the best solution in terms of detecting DDoS attacks with machine learning techniques as Random Forest model. With the planned approach, a secure and efficient SDN architecture can be developed. In SDN topology, the location of the controllers is important at this point. We have shown with our results that machine learning-based Random Forest model has achieved the best performance by classifying the traffic from attacked to normal traffic. We hope to implement our model on multi-controller SDN network to detect and mitigate the DDoS attack. The random forest model is the best among the models with the created dataset using permutation importance algorithm.

5. Conclusions

In this paper, we have represented a study of implementing machine learning and deep learning-based models for detecting the DDoS attack in SDN environment. The results obtained from our study shows that random forest has achieved the highest performance in detecting the DDoS attacks because of the centralized nature of the controller. The basic information related to network traffic can be obtained by the controller and can be evaluated by the machine learning-based Random Forest detection module. We have achieved the accuracy, and F1_score especially the percentage of above 99% with Random Forest classifier model . We have analyzed the traffic with the flexibility of the SDN structure, we have used Permutation Importance algorithm to extract best high ranked

topmost 20 features that contain more valuable information for our CNN classifier model related to the type of attacks, in our case DDoS.

Our results shows the required performance with high accuracy of 99.985%, precision above 99% and F1_score of 99.985%. We have implemented our model on Mininet based SDN environment. In future, we will also try to implement this approach in the real SDN environment with real network traffic, and evaluate the performance of the whole network in terms of other parameters, such as spoofing, latency, and throughput. Firstly, our experimental verification is the identification of 20 data-sets that have a significant impact on DDoS detection. It was confirmed that the random forest model has excellent DDoS detection performance in SDN networks.

Through this, Firstlly, it is possible to reduce the amount and time of collecting DDoS attack data-sets that affect the performance of the learning model. Secondly, it can reduce the time and cost of comparing various learning models and performance required for determining a learning model suitable for DDoS detection. we are verified that it is possible to reduce detection time of DDoS and appropriately utilize it when determining a detection model. Finally, various experimental methods for evaluating the performance of the learning model are presented so that related researchers can utilize them. Based on the results of this study, we are currently developing a new learning model for DDoS detection and mitigation in a blockchain network environment, and are conducting experiments and verification in real environments.

Acknowledgments: This paper was supported by Wonkwang University in 2021.

Sample Availability: Samples of the compounds ..... are available from the authors.

References

1. Ko, K.M. A DDoS Attack Detection Technique through CNN Model in Software Define Network. Journal of Korea Institute of Information, electronics, and communication technology. KIISE, 2020, pp. 605–610.
2. Imran, M.; Durad, M.H.; Khan, F.A.; Derhab, A. Toward an optimal solution against denial of service attacks in software defined networks. *Future Generation Computer Systems* **2019**, *92*, 444–453.
3. Rahman, O.; Quraishi, M.A.G.; Lung, C.H. DDoS attacks detection and mitigation in SDN using machine learning. 2019 IEEE World Congress on Services (SERVICES). IEEE, 2019, Vol. 2642, pp. 184–189.
4. Tselios, C.; Politis, I.; Kotsopoulos, S. Enhancing SDN security for IoT-related deployments through blockchain. 2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN). IEEE, 2017, pp. 303–308.
5. Hamid Tahaei, Rosli Bin Salleh, M.F.A.R.K.M.K.; Anuar, N.B. Cost Effective Network Flow Measurement for Software Defined Networks: A Distributed Controller Scenario. *IEEE Access* **2018**.
6. Rahman Obaid, Quraishi Mohammad Ali Gauhar, L.C.H. DDoS Attacks Detection and Mitigation in SDN Using Machine Learning. IEEE 2019 IEEE World Congress on Services (SERVICES). IEEE, 2019, pp. 1–8.
7. Smith-perrone, J.; Sims, J. Securing cloud, SDN and large data network environments from emerging DDoS attacks. 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence. IEEE, 2017, pp. 466–469.
8. Douligeris, C.; Mitrokotsa, A. DDoS attacks and defense mechanisms: a classification. Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No. 03EX795). IEEE, 2003, pp. 190–193.
9. Fonseca, P.; Bennesby, R.; Mota, E.; Passito, A. A replication component for resilient OpenFlow-based networking. 2012 IEEE Network operations and management symposium. IEEE, 2012, pp. 933–939.
10. Wang, J.; Wang, L. SDN-Defend: A Lightweight Online Attack Detection and Mitigation System for DDoS Attacks in SDN. *MDPI Sensors* **2022**.
11. Pedro Manso, Jose Moura, C.S. SDN-Based Intrusion Detection System for Early Detection and Mitigation of DDoS Attacks. *IEEE Access* **2019**.
12. Dharma, N.G.; Muthohar, M.F.; Prayuda, J.A.; Priagung, K.; Choi, D. Time-based DDoS detection and mitigation for SDN controller. 2015 17th Asia-Pacific Network Operations and Management Symposium (APNOMS). IEEE, 2015, pp. 550–553.

13. Oshima, S.; Nakashima, T.; Sueyoshi, T. Early DDoS detection method using short-term statistics. 2010 International Conference on Complex, Intelligent and Software Intensive Systems. IEEE, 2010, pp. 168–173.
14. Zubaydi Haider Dhia, A.M.C.Y.W. Review on Detection Techniques against DDoS Attacks on a Software-Defined Networking Controller. IEEE 2017 Palestinian International Conference on Information and Communication Technology (PICICT). IEEE, 2017, pp. 22–31.
15. Wang, R.; Jia, Z.; Ju, L. An entropy-based distributed DDoS detection mechanism in software-defined networking. 2015 IEEE Trustcom/BigDataSE/ISPA. IEEE, 2015, Vol. 1, pp. 310–317.
16. Ashraf, J.; Latif, S. Handling intrusion and DDoS attacks in Software Defined Networks using machine learning techniques. 2014 National Software Engineering Conference. IEEE, 2014, pp. 55–60.
17. Sultana, N.; Chilamkurti, N.; Peng, W.; Alhadad, R. Survey on SDN based network intrusion detection system using machine learning approaches. *Peer-to-Peer Networking and Applications* **2019**, *12*, 493–501.
18. Braga, R.; Mota, E.; Passito, A. Lightweight DDoS flooding attack detection using NOX/OpenFlow. IEEE Local Computer Network Conference. IEEE, 2010, pp. 408–415.
19. Zhai, S.; Cheng, Y.; Lu, W.; Zhang, Z. Deep structured energy based models for anomaly detection. *arXiv preprint arXiv:1605.07717* **2016**.
20. Tang, T.A.; Mhamdi, L.; McLernon, D.; Zaidi, S.A.R.; Ghogho, M. Deep learning approach for network intrusion detection in software defined networking. 2016 International Conference on Wireless Networks and Mobile Communications (WINCOM). IEEE, 2016, pp. 258–263.
21. Potluri, S.; Diedrich, C. Accelerated deep neural networks for enhanced intrusion detection system. 2016 IEEE 21st international conference on emerging technologies and factory automation (ETFA). IEEE, 2016, pp. 1–8.
22. Malaiya, R.K.; Kwon, D.; Kim, J.; Suh, S.C.; Kim, H.; Kim, I. An empirical evaluation of deep learning for network anomaly detection. 2018 International Conference on Computing, Networking and Communications (ICNC). IEEE, 2018, pp. 893–898.
23. Sambangi, S.; Gondi, L. A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression. *Proceedings* **2020**.
24. Awan, M.J., F.U.B.H.Y.A.N.H.H.M.H.O.Z.A. Real-Time DDoS Attack Detection System Using Big Data Approach. *Sustainability* **2021**.
25. Nakip, M.; Gelenbe, E. Mirai botnet attack detection with auto-associative dense random neural network. in IEEE Global Communications Conference. GLOBECOM, 2021, pp. 1–6.
26. Nakip, M.; Gelenbe, E. Botnet attack detection with incremental online learning. 2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC). Springer, 2022, pp. 51–60.
27. Polat H., P.O.; Cetin. Detecting DDoS Attacks in Software-Defined Networks Through Feature Selection Methods and Machine Learning Models. *Sustainability* **2020**.
28. T. A. Tuan, H. V. Long, R.K.I.P.N.T.K.S.e.a. Performance evaluation of botnet ddos attack detection using machine learning,. 2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC). Evolutionary Intelligence, 2019, pp. 1–12.
29. Khashab, Fatima, e.a. DDoS attack detection and mitigation in SDN using machine learning. 2021 IEEE 7th International Conference on Network Softwarization (NetSoft). IEEE, 2021.
30. Sanjeetha R, Anita Kanavalli, A.G.A.P.S.A. Real-time DDoS Detection and Mitigation in Software Defined Networks using Machine Learning Techniques. *International Journal of Computing* **2022**.
31. Tang, Tuan A., e.a. Deep learning approach for network intrusion detection in software defined networking. 2016 international conference on wireless networks and mobile communications (WINCOM). IEEE, 2016.
32. Hasan Alkahtani, T.H.H.A. Developing Cybersecurity Systems Based on Machine Learning and Deep Learning Algorithms for Protecting Food Security Systems: Industrial Control Systems. *MDPI Electronics* **2022**.
33. Theyazn H. H. Aldhyani, H.A. Attacks to Automatous Vehicles: A Deep Learning Algorithm for Cybersecurity. *MDPI Sensors* **2022**.
34. Devendra. "DDoS Dataset- Kaggle. https://www.kaggle.com/devendra416/ddos-datasets, 2020. Accessed: 2020-07-20.
35. Yeo M., Koo Y., Y.Y.H.T.R.J.S.J.P.C. Flow-based malware detection using convolutional neural network. IEEE 2018 International Conference on Information Networking (ICOIN). Korean Society for Internet Information(KSII), 2018, pp. 1–26.

36. Team, M. "Mininet Overview- Mininet. http://mininet.org/overview/, 2018. Accessed: 2020-07-15.
37. Yuh-Shyan Chen, Y.T.T. A Mobility Management Using Follow-Me Cloud-Cloudlet in Fog-Computing-Based RANs for Smart Cities. MDPI Sensors. MDPI, 2018, pp. 1–26.
38. Ryu, A. Component-based Software-defined Networking Framework, 2013.
39. Asadollahi, S.; Goswami, B.; Sameer, M. Ryu controller's scalability experiment on software defined networks. 2018 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC). IEEE, 2018, pp. 1–5.