DDoS Attack Preventing and Detection with the Artificial Intelligence Approach

Chapter · March 2022

DOI: 10.1007/978-3-030-98457-1_3

3 authors:

Tariqul Islam
Daffodil International University
16 PUBLICATIONS **3** CITATIONS
SEE PROFILE

Md. Ismail Jabiullah
Daffodil International University
53 PUBLICATIONS **121** CITATIONS
SEE PROFILE

Dm. Mehedi Hasan Abid
Daffodil International University
8 PUBLICATIONS **3** CITATIONS
SEE PROFILE

DDoS attack preventing and detection with the artificial intelligence approach

Tariqul Islam[1], Md. Ismail Jabiullah[2], Dm. Mehedi Hasan Abid[3]
Daffodil International University, Dhaka, Bangladesh
[1]tariqul15-2250@diu.edu.bd, [2]drismail.cse@diu.edu.bd, [3]mehedi15-226@diu.edu.bd

Abstract. DDoS attacks are a major Internet security concern with this large number of customers. Each attack sends a service request to a certain server, which limits the server's capacity to provide normal services. Since the attackers use legitimate packages and alter their package information, traditional methods are not very effective. The assault on DDoS is one of the most potent Internet hacking techniques. The hacker's basic weapon to take down and crash websites during these sorts of assaults is network trafficking. There are different sub-categories, each category explains how a hacker attempts to enter the network. In this paper, we define the DDoS attacks detection method based on artificial intelligence and explored with more than 96-percent accuracy a technique to detect a DDoS attacks assault danger using artificial intelligence (A.I). In addition to a secure or healthy network, authors have identified 7 separate sub-categories of DDoS attacks.

Keywords: DDoS, Cyber Security, Artificial Intelligence, Neural Networks, Data Visualization, Data Analytics

1 Introduction

The DDoS attack uses many distributed resource contra targets which will deprive authorized clients of service [1], [2]. DDoS attacks generate large volumes of traffic in a small time [3]. DDoS attacks can have a big impact on victims. Two types of DDoS attacks are IP spoofing, other is Flooding attacks. IP spoofing and Flooding attack differences are respectively impersonating a trusted source, sending exceeding packets to disrupt the services [4]. Also, DDoS attacks have three kinds of different flooding attacks [5], [6]. It is possible to use ML classification methods to differentiate between good and bad packages. Bayes classifiers are used in ML classification methods based upon the application of the Bayes theorem [7]. ANN can communicate with neurons and solve problems like the brain that is used in security fields [8]. Attack aims to jam the excessive traffic on the network or server. It is successful by the use, as a source of the attack, of several hacked systems. “SYN Flood, UDP Flood, MSSQL, LDAP, Portmap, NetBIOS” are among the sub-categories we identified via our study. Machine learning is one of the most frequent backbones of AI day. We utilize address issues with precise human performance in different fields. We have evaluated the A.I.

boundaries once again to discover the dangers in the field of cyber safety. In this study, the logs that were created during a DDoS assault were thoroughly analyzed using supervised and non-controlled threat detection approaches. Finally, we utilized deep learning to achieve more accuracy than 96-percent for classification and safe connection for distinct types of DDoS threats.

2 Related work

In this paper, Zhang et al. [9] describe that in last year's many DDoS attacks launched at a minimum cost. The cause says that DDoS attacks traffic is the same as normal traffic. Some ML algorithms are able to classify and detect traffic of DDoS attacks. They survey new papers on DDoS attack detection progress. Here [10] proposes NN (Neural network) based DDoS attacks detection that has five phases. They store traffic data in Hadoop distributed systems to detect DDoS attacks. The paper [11] proposes a mitigation model by ML algorithm for DDoS attacks detection. OMS for composing the model and monitoring DDoS attacks impact measurements. This model performance is increased than other ML algorithms. Authors [12] develop NN based DDoS attacks detection systems. Here [13] authors design an ANN-based model for detection of DDoS DNS amplification attacks. Ndibwile et al. [14] propose a network architecture for servers to distinguish DDoS attack traffic. Supervised learning machine-learning algorithm for customizing the ISP network gateway and decision tree used for malicious traffic, the random tree algorithm also used for avoiding false-positive traffic. Fouladi et al. [15] propose a frequency analysis method for detection, DWT gives the best accuracy but some features extracted from DWT and DFT then increase the accuracy of detection. Ramadan et al. [16] design an AIS system for flood attacks detection. There are four different phases of DCA (Dendritic cell algorithm). Peraković et al. [17] develop an ANN-based detection system, In ANN (Artificial Neural Network) model, traffics are classified into 4 types. The classification accuracy of UDP DDoS attacks is slightly lower than that of normal traffic. Kushnir et al. [18] propose an approach to enable completely automated. The solution is applicable with minimal configuration effort to a wide range of web applications without requiring access to the source code and without requiring the availability of a formal access control model. Our evaluation demonstrates that the solution can indeed be applied to various types of web applications. There are some limitations such as Focus on getting Requests and Evaluation scope is limited with four simple websites. Man et al. [19] Proposed a detecting method like distributed vulnerability, which method to improve the system performance and efficiency for large Java Web programs. Some limitations are that the model only exploits vulnerabilities that are supported by American Fuzzy lop (such as Code Execution, Denial-of-service, etc)". Model elements are targeted when the program decomposes, worker nodes perform guided fuzzing, and master nodes perform symbolic execution, creating constraint condition input for the test program. The result parameters are the efficiency of fuzzing and program coverage. Anagandula et al. [19] review the performance of 4 black box web- application scanners for store SQL, XSS vulnerability at the well-known testbed (i.e. Wackopicko, Scan it) by three renowned

scanners i.e. Wireshark, Burp Suite, and Nessus. Abdullah Al Jumuah et al. [20] they developed a six-stage algorithm and used chaos theory to efficiently detect DDoS attacks. A mirror image of an actual network environment is used to initiate the learning process. The authors launched various DDoS attacks while legitimate traffic flows across the network. They distinguished DDoS attacks from real traffic using supervised and unsupervised ANN methods. They used up-to-date datasets and formed artificial neural networks into two learning methods and achieved over 95% accuracy in detecting DDoS attacks. In this paper [21], the authors propose hybrid ML model. They show results the hybrid ML model gives us good accuracy with detection rate compared to other normal ML models. Sabah Alzahrani et al. [22] propose detecting known and unknown DDoS Attacks systems, apply different IDS approaches, anomaly-based distributed ANN and signature-based approaches. This research paper [23] reviews the DDoS attacks survey to prevent recognition by data mining and use it to identify DDOS attack patterns and analyze patterns with ML algorithms. The result give the best accuracy by data mining algorithms for preventing DDoS attacks. Ghafarian et al. [24] present a platform-independent hybrid method without including further defensive code in the application. Mohammadi et al. [25] present a method for detecting vulnerabilities considerably earlier in the development cycle, as well as providing detailed feedback to developers on how to repair the flaw. The test evaluation is based on execution behavior, which can detect subtle vulnerabilities originating from internal browser decoding functions, and this method works with all currently available web languages. Ibarraflallos et al. [26] design a durable, reliable, and effective protection filter for reducing common web injection attacks. The author proposes a regular expressions-based input field validation filter and a sanitization method. Figueiredo et al. [27] developed a learning-based tool that is able to detect input validation vulnerability from the source code of a web application using static data flow analysis. This paper limitation is research focused on the PHP and Java Source Code, the study did not compare their outcome with other dictation tools, and MERLIN only considered the input validation vulnerability attack. Kao et al. [28] The paper provides an in-depth examination of the many types of SQLi assaults, their descriptions, and possible investigation strategies. However, there are significant limitations, such as the possible impact on SQL injection investigations and the examination of practical uses of the SIA framework in legal procedures, and the framework's accuracy is not thoroughly evaluated. Mokbal et al. [29] present large real-world data made up of 138,569 unique records for detecting XSS threats that has been built fully and distinctively. To offer training and testing datasets, a dynamic features extraction method is presented to extract data from a neural network model. When employed against XSS assaults, this approach is more resilient and platform-independent. The authors [30] during the session hijacking attack in VANETs, the main focus is on recognizing the malicious node that poses as a valid vehicle. The main focus of this research is on preventative approaches for deliberately hidden nodes. It does not, however, address security concerns such as message confidentiality and privacy. In this paper [31] the implications of SATs' actions while analyzing applications written in various coding styles and programming methods, as well as a discussion of the exploitability of SQLi vulnerabilities reported by SATs as true positives. Some result parameters are “TP, FP, FN, FFP (False False Positive)”. Here [32] authors

propose an LVQ neural network for host anomaly detection that uses the DDoS attacks detection method. That can improve the recognition rate of the detection system. Yuan et al. [33] investigate vulnerabilities analyze common web application vulnerabilities, research the basic characteristics of these vulnerabilities in depth, and comprehend the principles and remedies to these weaknesses. Moustafa et al. [34] suggest a method to automatically extract relevant features from web data and network traffic data to improve the effectiveness of threat detection techniques.

3 Research Methodology

3.1 Preprocessing Data

One of our initial problems has been the processing of data cause it [35] contained 88 characteristics. It was a really difficult challenge for us to process this enormous data inside a restricted RAM capacity. So, we lowered the attribute data type, hence reducing the data framework's memory use. Float64 data types have been reduced to "float32, int64 to int32, int32 to uint32" etc. Nearly 42 percent of the initial size has been decreased. The qualities or characteristics of our database were still almost unlimited, therefore, we also processed these data at the preprocessing stage. For instruments used throughout this research TensorFlow, Scikit Learn, Matplotlib, Seaborn.

3.2 The Distribution of Target Features

It can be observed that we have attempted to keep the goal feature evenly distributed together with the data set. While UDPLag is rather unequal in distribution to others, this situation has nevertheless been dealt with later on in this study.

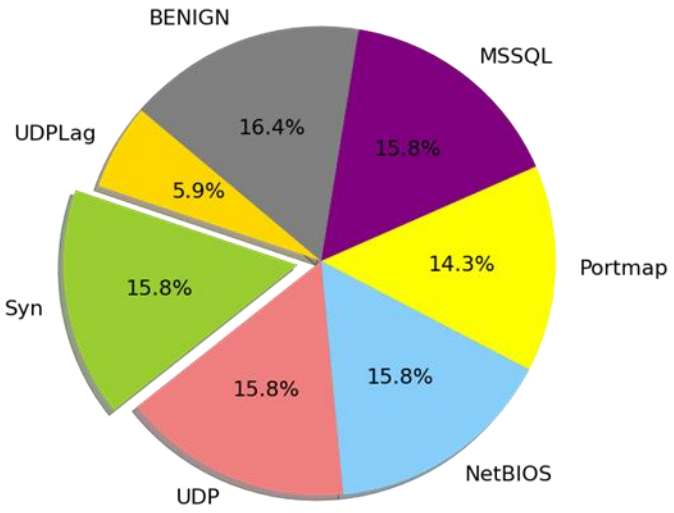

Figure 1: Target Features

3.3 Data analysis

Data analysis is the process of examining and describing a huge number of data points with care. This element of the research process is usually completed in stages [36]. Typically, researchers collect data during the whole data collection procedure. The purpose of this step is to identify the patterns and procedures that will allow them to examine the data. To preserve the integrity of the data, it is critical to strictly follow the norms and procedures for statistical analysis.

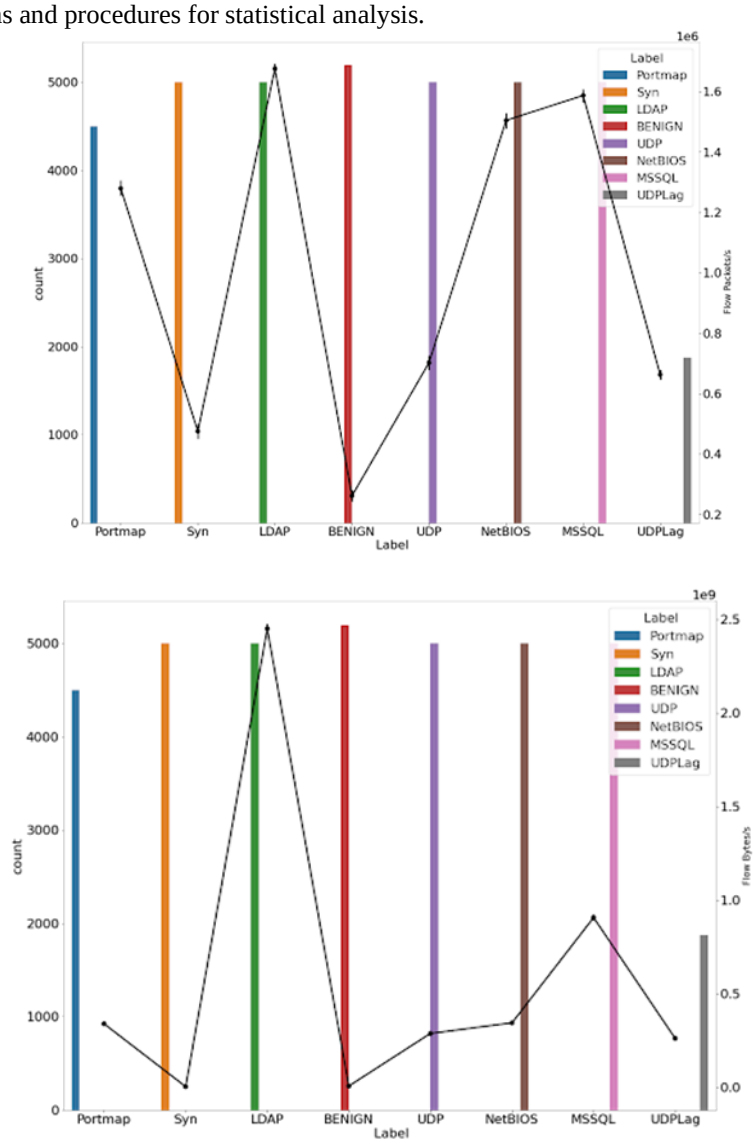

Figure 2: The Bit and Packet Flow Drift

In the two above assessments, bit and packet flow drift in a DDoS assault in comparison with a Benign or a Security connection may be observed.

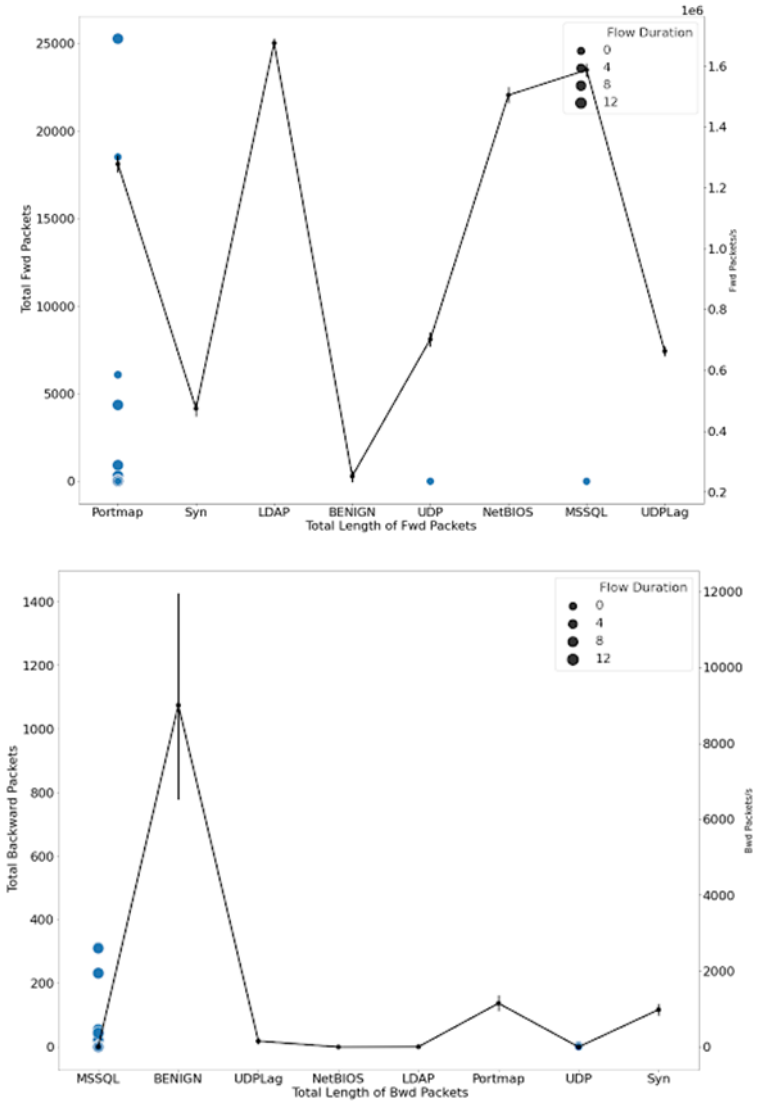

Figure 3: Protocol and Incoming

In each kind of protocol and incoming, we have also studied the distribution of each threat category. The diagrams showing the analysis are shown below.

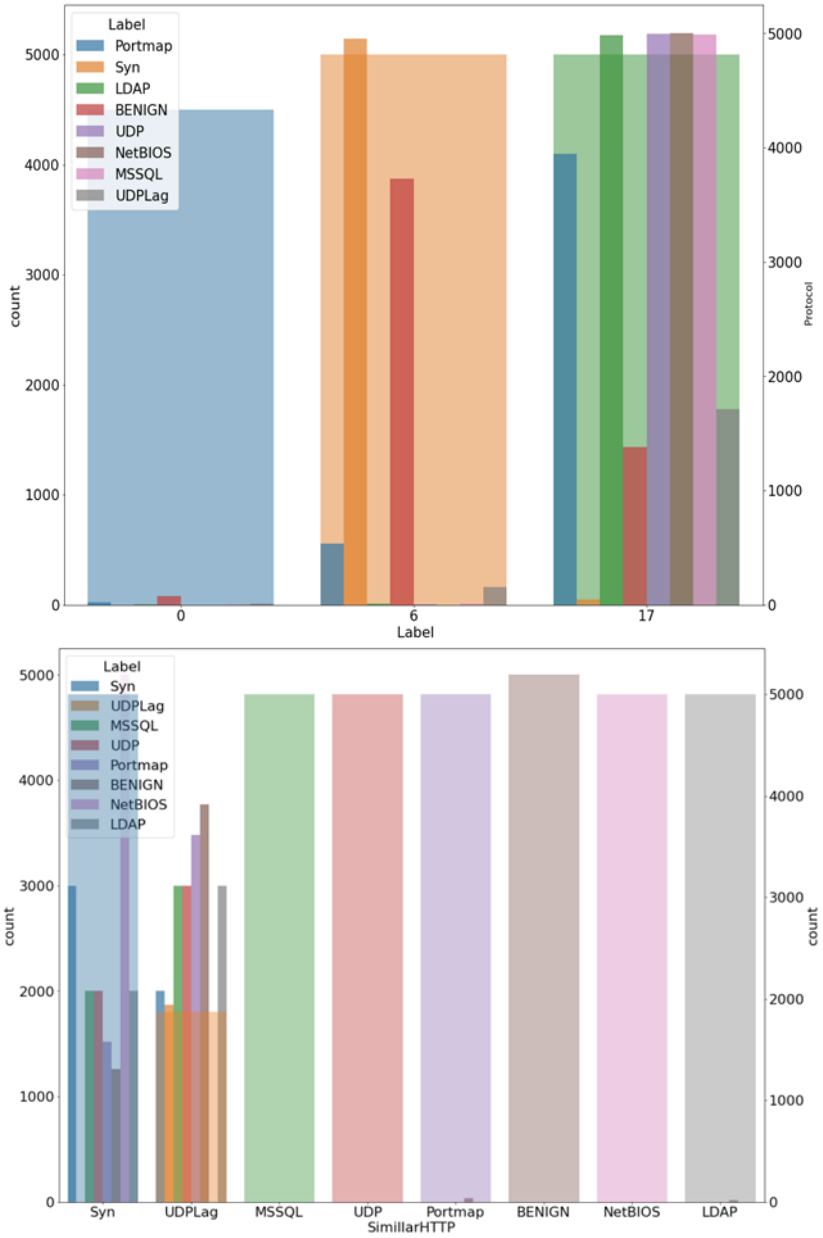

Figure 4: Threat Analysis

4 The Approach for detection of the threat in unsupervised method

In our unsupervised method, we don't allow our model to learn from the target variables, but instead force our algorithms to know from the input data. Before training, pre-processing. Some of the functions of our data have been deleted. “Flow Packets” and "Flow Bytes/s" have been deleted as they have been converted to excessively big values for float64 and NaN after normal scaling. By conventional scaling and normalization, we have scaled our data. Use the main dimension, reduction-component analysis, and decrease the size to two-dimensional data.

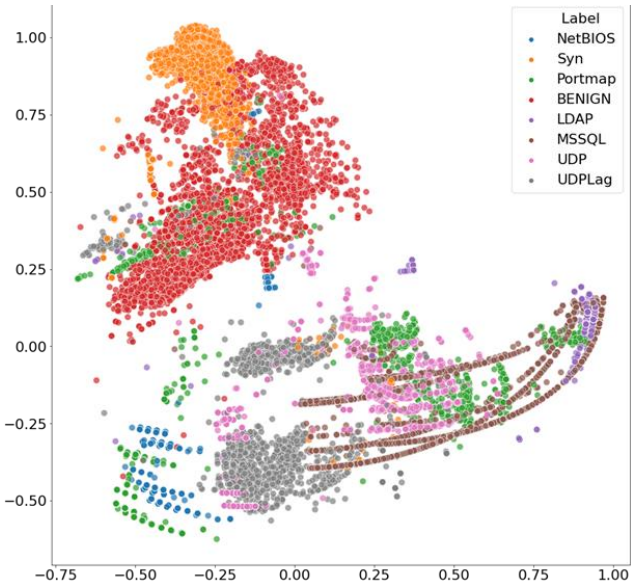

Figure 5: Visualizing DDoS attacks

Thus, it can be readily noticed from the above two visualizations that the distinct threats may be clustered in some way by our method. See how the produced clusters might be identified with our unattended model.

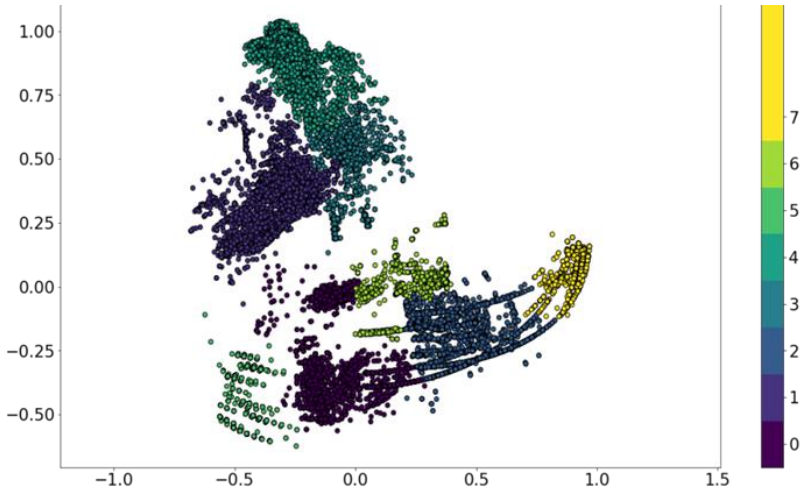

Figure 6: Visualizing DDoS attacks after Agglomerative Clustering

It seems that our uncontrolled model has successfully identified a trend in the data and could to some extent segment our target variable. Non-controlled learning provides a detailed perspective of the data's shape and structure, as well as analysis. If the form and structure of the data change as they are not informed of the target information, the prediction of the grouping and the label of the target will vary. No way controlled machine learning is better adapted to real-life situations and It's not accurate. It is also one of the reasons that uncontrolled formed models are not adapted to production deployments.

5 The Approach for detection of the threat in supervised method

Unlike the supervised approach, we allow our model to learn from the target variable that allows our model to understand the model from target tags. For the unsupervised approach, identical pretreatment data were used. We utilized profound learning in this example to train our model.

5.1 DL model structure

Deep learning is a method for teaching computers how to filter and classify data. Here our authors design the best DL model for increasing fit with the dataset and input given. Authors give input as Keras learning because it wraps the effective digital computing libraries TensorFlow and allows you to define and form neural network models.

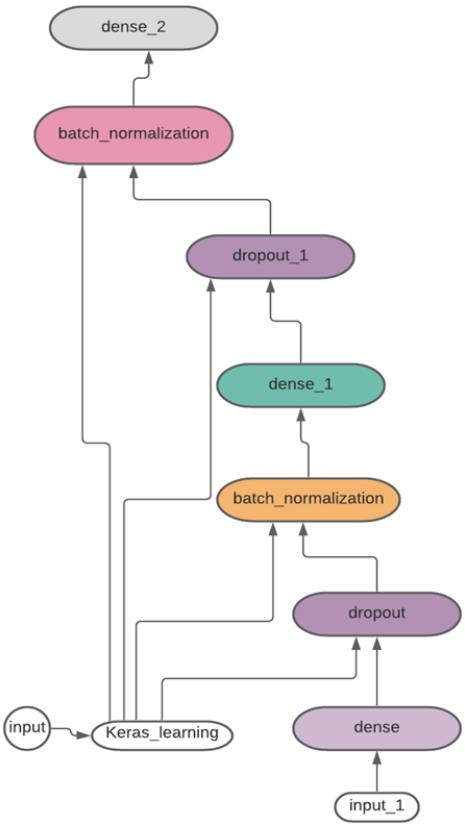

Figure 7: The DL model

As our objective variable was unbalanced, a stratified K-fold was used to form and assess our data on each fold. The distribution of training and the validation of an unbalanced feature is balanced. To evaluate the model's performance, we used Adam as the base optimizer and ROC AUC. An area where the ROC AUC score is calculated from the prediction scores below the characteristic curve of the recipient. We developed and validated our 10-step model and achieved ROC Auc scores of 96% or more on an average of 97% or more for threat detection.

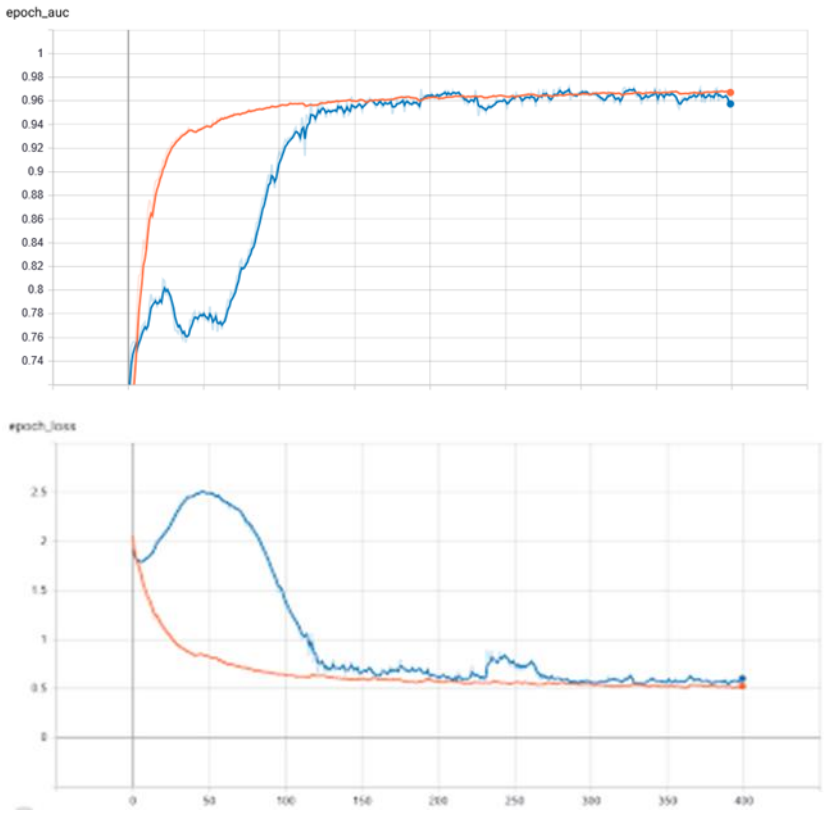

Figure 8: Characteristic Curve

6 Result

The evaluation procedure' results were achieved using 10-fold cross-validation. This technique ensures that the data is never from the same person. A different split of the original sample is used each time the test and training data are used.

Table 1: Ten folds classifications report

	Precision	Recall	f1-score	Support
BENIGN	0.98	0.99	0.98	519
Portmap	0.94	0.97	0.96	500
NetBIOS	0.86	0.85	0.86	500
LDAP	0.60	0.90	0.72	500
UDP	0.91	0.10	0.18	500
Syn	0.99	0.99	0.99	500
MSSQL	0.66	0.75	0.70	500
UDPLag	0.55	0.88	0.67	188
Accuracy			0.80	3707

Macro avg	0.81	0.80	0.76	3707
Weighted avg	0.83	0.80	0.77	3707

In table 1, Syn was detected with greatest precision, recall, and f1-score of 99 percent false positives per volume in table 1, resulting in a positive prediction. On a 2.2 GHz dual-core laptop computer, the whole processing time was between 9.9 and 13.0 seconds on average, with the majority of the time spent on disk candidate discovery.

Table 2: Ten folds classifications report

	Accuracy
BENIGN	0.97687861
Portmap	0.206
NetBIOS	0.96
LDAP	0.862
UDP	0.138
Syn	0.982
MSSQL	0.498
UDPLag	0.91489362

Precision refers to everything that is relevant, whereas recall refers to everything that is genuinely relevant. The recall of your model is also known as its sensitivity, whereas precision is known as Positive Predicted Value. The precision of the table 2 discovered folds classification has been justified.

7 Conclusion

In this paper authors have relatively little labeled data in the real situation compared to unlabeled data, there are ways to achieve amazing performance, such as semi-monitored learning and self-management. The Model Fairness Indicator may also be used as one of the TensorFlow tools to improve model assessment and performance scaling. The authors have identified 7 separate sub-categories of DDoS attacks threats. And we get 96-percent accuracy to detect DDoS attacks and assault danger using artificial intelligence.

References

1. X. Yuan, C. Li, and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," 2017 IEEE International Conference on Smart Computing (SMARTCOMP), Hong Kong, 2017, pp. 1-8. DOI: 10.1109/SMARTCOMP.2017.7946998
2. M. Guri, Y. Mirsky, and Y. Elovici, "9-1-1 DDoS: Attacks, Analysis and Mitigation," 2017 IEEE European Symposium on Security and Privacy (EuroS&P), Paris, France, 2017, pp. 218-232. DOI: 10.1109/EuroSP.2017.23

3. C. J. Hsieh and T. Y. Chan, "Detection DDoS attacks based on neural-network using Apache Spark," 2016 International Conference on Applied System Innovation (ICASI), Okinawa, 2016, pp. 1-4. doi: 10.1109/ICASI.2016.7539833
4. B. S. Kiruthika Devi, G. Preetha, G. Selvaram and S. Mercy Shalinie, "An impact analysis: Real time DDoS attack detection and mitigation using machine learning," 2014 International Conference on Recent Trends in Information Technology, Chennai, 2014, pp. 1-7. doi: 10.1109/ICRTIT.2014.6996133
5. Irom Lalit Meitei, Khundrakpam Johnson Singh, and Tanmay De. 2016. Detection of DDoS DNS Amplification Attack Using Classification Algorithm. In Proceedings of the International Conference on Informatics and Analytics (ICIA-16). ACM, New York, NY, the USA, Article 81, 6 pages. DOI: https://doi.org/10.1145/2980258.2980431
6. G. Ramadhan, Y. Kurniawan and Chang-Soo Kim, "Design of TCP SYN Flood DDoS attack detection using artificial immune systems," 2016 6th International Conference on System Engineering and Technology (ICSET), Bandung, 2016, pp. 72-76. doi: 10.1109/ICSEngT.2016.7849626
7. Rish, I. "An empirical study of the naive Bayes classifier." Journal of Universal Computer Science 1.2(2001):127
8. Ahmad, Iftikhar, A. B. Abdullah, and A. S. Alghamdi. "Artificial neural network approaches to intrusion detection: a review." Wseas International Conference on Telecommunications and Informatics World Scientific and Engineering Academy and Society (WSEAS), 2009:200-205
9. Zhang, Boyang, Tao Zhang, and Zhijian Yu. "DDoS detection and prevention based on artificial intelligence techniques." 2017 3rd IEEE International Conference on Computer and Communications (ICCC). IEEE, 2017
10. C. J. Hsieh and T. Y. Chan, "Detection DDoS attacks based on neural-network using Apache Spark," 2016 International Conference on Applied System Innovation (ICASI), Okinawa, 2016, pp. 1-4. doi: 10.1109/ICASI.2016.7539833
11. B. S. Kiruthika Devi, G. Preetha, G. Selvaram and S. Mercy Shalinie, "An impact analysis: Real time DDoS attack detection and mitigation using machine learning," 2014 International Conference on Recent Trends in Information Technology, Chennai, 2014, pp. 1-7. doi: 10.1109/ICRTIT.2014.6996133
12. T. Zhao, D. C. T. Lo and K. Qian, "A Neural-Network Based DDoS Detection System Using Hadoop and HBase," 2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems, New York, NY, 2015, pp. 1326-1331. doi: 10.1109/HPCC-CSS-ICESS.2015.38
13. Irom Lalit Meitei, Khundrakpam Johnson Singh, and Tanmay De. 2016. Detection of DDoS DNS Amplification Attack Using Classification Algorithm. In Proceedings of the International Conference on Informatics and Analytics (ICIA-16)
14. J. D. Ndibwile, A. Govardhan, K. Okada and Y. Kadobayashi, "Web Server Protection against Application Layer DDoS Attacks Using Machine Learning and Traffic Authentication," 2015 IEEE 39th Annual Computer Software and Applications Conference, Taichung, 2015, pp. 261-267. doi: 10.1109/COMPSAC.2015.240
15. R. F. Fouladi, C. E. Kayatas and E. Anarim, "Frequency based DDoS attack detection approach using naive Bayes classification," 2016 39th International Conference on Telecommunications and Signal Processing(TSP), Vienna, 2016, pp.104 107. doi:10.1109/TSP.2016.7760838
16. G. Ramadhan, Y. Kurniawan and Chang-Soo Kim, "Design of TCP SYN Flood DDoS attack detection using artificial immune systems," 2016 6th International Conference on System

Engineering and Technology (ICSET), Bandung, 2016, pp. 72-76. doi: 10.1109/ICSEngT.2016.7849626
17. D. Peraković, M. Periša, I. Cvitić and S. Husnjak, "Artificial neuron network implementation in detection and classification of DDoS traffic," 2016 24th Telecommunications Forum (TELFOR), Belgrade, 2016, pp. 1-4. doi: 10.1109/TELFOR.2016.7818791
18. Kushnir, Malte, et al. "Automated black box detection of HTTP GET request-based access control vulnerabilities in web applications." ICISSP 2021, online, 11-13 February 2021. SciTePress, 2021. Man, Hongpeng, et al. "JSEFuzz: Vulnerability Detection Method for Java Web Application." 2018 3rd International Conference on System Reliability and Safety (ICSRS)
19. Anagandula, Karthik, and Pavol Zavarsky. "An Analysis of Effectiveness of Black-Box Web Application Scanners in Detection of Stored SQL Injection and Stored XSS Vulnerabilities." 2020 3rd International Conference on Data Intelligence and Security (ICDIS). IEEE, 2020.
20. Aljumah, Abdullah, and Tariq Ahamad. "A novel approach for detecting DDoS using artificial neural networks." International Journal of Computer Science and Network Security 16.12 (2016): 132-138
21. Deepa, V., K. Muthamil Sudar, and P. Deepalakshmi. "Detection of DDoS attack on SDN control plane using hybrid machine learning techniques." 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, 2018
22. S. Alzahrani and L. Hong, "Detection of Distributed Denial of Service (DDoS) Attacks Using Artificial Intelligence on Cloud," 2018 IEEE World Congress on Services (SERVICES), 2018, pp. 35-36, doi: 10.1109/SERVICES.2018.00031
23. Bandara, K. R. W. V., et al. "Preventing DDOS attack using data mining algorithms." International Journal of Scientific and Research Publications 6.10 (2016): 390
24. Ghafarian, Ahmad. "A hybrid method for detection and prevention of SQL injection attacks." 2017 Computing Conference. IEEE, 2017
25. Mohammadi, Mahmoud, et al. "Automatic web security unit testing: XSS vulnerability detection." 2016 IEEE/ACM 11th International Workshop in Automation of Software Test (AST). IEEE, 2016
26. Ibarra-Fiallos, Santiago, et al. "Effective Filter for Common Injection Attacks in Online Web Applications." IEEE Access 9 (2021): 10378-10391
27. Figueiredo, Alexandra, Tatjana Lide, and Miguel Correia. "Multi-Language Web Vulnerability Detection." 2020 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW). IEEE, 2020
28. Kao, Da-Yu, Chung-Jui Lai, and Ching-Wei Su. "A Framework for SQL Injection Investigations: Detection, Investigation, and Forensics." 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2018
29. Mokbal, Fawaz Mahiuob Mohammed, et al. "MLPXSS: an integrated XSS-based attack detection scheme in web applications using multilayer perceptron technique." IEEE Access 7 (2019): 100567-100580
30. Jeevitha, R., and N. Sudha Bhuvaneswari. "Malicious node detection in VANET Session Hijacking Attack." 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2019
31. Medeiros, Ibéria, and Nuno Neves. "Effect of Coding Styles in Detection of Web Application Vulnerabilities." 2020 16th European Dependable Computing Conference (EDCC). IEEE, 2020
32. Li, Jin, Yong Liu, and Lin Gu. "DDoS attack detection based on a neural network." 2010 2nd international symposium on aware computing. IEEE, 2010

33. Yuan, Hui, et al. "Research and Implementation of Security Vulnerability Detection in Application System of WEB Static Source Code Analysis Based on JAVA." The International Conference on Cyber Security Intelligence and Analytics. Springer, Cham, 2019
34. Moustafa, Nour, Gaurav Misra, and Jill Slay. "Generalized outlier gaussian mixture technique based on automated association features for simulating and detecting web application attacks." IEEE Transactions on Sustainable Computing (2018)
35. The UNIVERSITY OF New BRUNSWICK DDoS Evaluation Dataset (CIC-DDoS2019) https://www.unb.ca/cic/datasets/ddos-2019.html
36. Shamoo, Adil E., and David B. Resnik. Responsible conduct of research. Oxford University Press, 2009