*I. J. Computer Network and Information Security*, **2022, 6, 89-97**
Published Online on December 8, 2022 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2022.06.07

# Detection of DDoS Attacks Using Machine Learning Classification Algorithms

**Kishore Babu Dasari***
Department of CSE, Acharya Nagarjuna University, Guntur, AP, India
E-mail: dasari2kishore@gmail.com
ORCID iD: https://orcid.org/0000-0001-6920-0674
***Corresponding Author

**Nagaraju Devarakonda**
School of Computer Science and Engineering, VIT-AP University, AP, India
E-mail: dnagaraj_dnr@yahoo.co.in
ORCID iD: https://orcid.org/0000-0003-4864-8482

Received: 20 November 2021; Revised: 02 March 2022; Accepted: 27 May 2022; Published: 08 December 2022

**Abstract:** The Internet is the most essential tool for communication in today's world. As a result, cyber-attacks are growing more often, and the severity of the consequences has risen as well. Distributed Denial of Service is one of the most effective and costly top five cyber attacks. Distributed Denial of Service (DDoS) is a type of cyber attack that prevents legitimate users from accessing network system resources. To minimize major damage, quick and accurate DDoS attack detection techniques are essential. To classify target classes, machine learning classification algorithms are faster and more accurate than traditional classification methods. This is a quantitative research applies Logistic Regression, Decision Tree, Random Forest, Ada Boost, Gradient Boost, KNN, and Naive Bayes classification algorithms to detect DDoS attacks on the CIC-DDoS2019 data set, which contains eleven different DDoS attacks each containing 87 features. In addition, evaluated classifiers' performances in terms of evaluation metrics. Experimental results show that AdaBoost and Gradient Boost algorithms give the best classification results, Logistic Regression, KNN, and Naive Bayes give good classification results, Decision Tree and Random Forest produce poor classification results.

**Index Terms:** DDoS Attacks, CIC-DDoS2019, Logistic Regression, Decision Tree, Random Forest, Ada Boost, Gradient Boost, KNN, Naive Bayes.

## 1. Introduction

Distributed Denial of Service (DDoS) [1,2,3] is a powerful weapon used by cyber attackers to disrupt the web servers or system resources of legitimate users by flooding or crashing them from multiple sources with too much traffic. DDoS attacks have increased substantially in both volume and intricacy as a result of the significant increase in internet usage, notably with the commencement of the COVID-19 pandemic. The number of DDoS attacks per day has risen exponentially in the first half of 2021 compared to Q4 of 2020. According to Cisco, the overall number of DDoS attacks globally would reach 15.4 million by 2023. According to NetScout, 50% of the DDoS attacks targeted financial organizations. Facebook, WhatsApp, and Instagram company servers shut down seven hours on 4th, October 2021, lost seven billion dollars, and Facebook CEO dropped one position in the billionaire's list. DDoS attacks not only bring down network resources, but they can also bring down the whole organization. It shows the importance of DDoS attacks detection and mitigation.

Botnets are used by attackers to launch DDoS attacks. A botnet is a short form of a robot network. It is a network of computers controlled by the attacking party and made up of malware-infected zombies or bots. Here the attacking party is referred to as the bot-herder or bot-master. Bot-herders choose some bots as command and control servers. Command and control servers carry the instructions from bot-herders to bots. When bot-herders send attack-launch instructions, the bots carry out massive traffic to the victim machine. Because all bots are legitimate users, the attackers can hide their identities by using bot-nets.

The three primary types of DDoS attacks are volume-based, protocol-based, and application layer-based attacks. In volume-based attacks, the attacker's primary goal is bandwidth saturation. Volume-based DDoS attacks comprise UDP and ICMP floods. In protocol-based attacks, the attacker's original objective is to get access to server resources. Protocol-based DDoS attacks comprise SYN-floods, Smurf, and the ping of death. The attacker's primary goal in application layer-based attacks are the application interface of web sites. Application layer-based DDoS attacks comprise HTTP-flood and slow loris.

To minimize the damage caused by DDoS attacks, it is necessary to detect them quickly and accurately. According to research on DDoS attacks, it is difficult because DDoS attacks are distributed in nature [4], they comprise large data sets, which makes detection more challenging [5,6]. The detection approaches are effective if the network traffic data is distributed normally, but data from computer networks have a non-Gaussian distribution. Machine learning classification algorithms [7] can handle these types of challenges more effectively than statistical and data mining methods.

This section introduces the DDoS attacks, key participants of the DDoS attack architecture, types of DDoS attacks and the research motivation and objective for detecting DDoS attacks. The related work is explained in part II of this paper, as well as the reason for using machine learning classification algorithms and the CICDDoS2019 dataset to detect DDoS attacks in this study. In section III of this paper, the methodology is explained, including datasets, preprocessing, and machine learning classification algorithms. The results and discussion are explained with experimental results in section IV of this paper. The study's conclusion is found in Section V of this paper.

## 2. Related Work

Distributed Denial of Service (DDoS) is a network security threat that attempts to overload target networks with malicious traffic. Despite the fact that numerous DDoS attack detection methods have been developed by researchers for DDoS attack detection, one of the primary concerns remains the development of an accurate detection method. On the other hand, the availability of well-designed DDoS attack datasets are important for the evaluation of new DDoS attack detection methods.

Laura et al. [8] proposed using entropy and frequency-sorted distributions of choosing a packet attribute to detect DDoS attacks. Based on aggregated NetFlow statistics, Hussein et al. [9] proposed solution uses the Z-score and co-variance measurements to detect DDoS traffic as a divergence from regular traffic. For efficient detection of DDOS attacks, Benamar et al.[4] proposed a detection method based on the continuous ranked probability score statistical measure and the exponential smoothing scheme. The detection efficiency of these statistical detection methods is good, but the model construction takes a long time.

To address the aforementioned issues, the detection of DDoS attacks using machine learning algorithms has gradually become the focus of research. Tanaphon et al.[10] proposed a method for detecting DDoS attacks using KNN, MLP, and SVM machine learning algorithms, and tested it on the KDD CUP 1999 and NSL-KDD datasets. Karan et al.[11] proposed a method for detecting DDoS attacks using SVM machine learning algorithms, and tested it on the KDD CUP 1999 dataset. The accuracy of the machine learning method for detecting DDoS attacks is high, but the detection efficiency is low because the tested DDoS attack datasets do not contain all network features.

This study uses the CICDDoS2019 dataset, which has all required network traffic features, to handle the aforementioned issue. This study tested logistic regression, decision tree, random forest, Ada boost, Gradient boost, KNN and Naïve Bayes machine learning classification algorithms to detect DDoS attacks on CICDDoS2019 dataset.

## 3. Methodology

The research aims to use machine learning classification algorithms to detect DDoS attacks.

### *3.1. Data Set*

The CICDDoS2019 data set is used in this study. It is a DDoS evaluation data set collected from the Canadian Institute for Cyber Security. It contains eleven different DDoS attack data sets with more than eighty features. Each data set contains millions of records. In this research, collected two lakhs records from each attack data set and combined them into a single data set.

### *3.2. Preprocessing*

To improve model generalizability, data pre-processing [12,13] is performed at the start of machine learning experiments. In Pre-processing, first removed socket features, next removed missing values records. Encoding the label data with ‘0’ for BENIGN class labels and ‘1’ for DDoS attacks class labels. Split the data 80:20 percentage of training and test the models accordingly. Standardize the training data for efficient classification.

### *3.3. Classification Algorithms*

Machine learning algorithms [14,15] are used to handle many network security issues because of their substantial advantages and advanced features, such as high accuracy, continuous development, and the ability to manage a vast amount of data. In this research, applied Logistic Regression, Decision Tree, Random Forest, Ada Boost, Gradient Boost, K Nearest Neighbors, and Naive Bayes classification algorithms on the data set to detect DDoS attacks.

### A. Logistic Regression

The logistic function is used in the Logistic Regression (LR) classification algorithm [16,17]. The Logistic function also calls the sigmoid function. Sigmoid function is

$$\varnothing(z) = \frac{1}{1+e^{-z}} \tag{1}$$

Here z is the input which is the linear combination of weights and features.

$$z = w^T x \; = \; w_0 + w_1 x \; + w_2 x^2 + \; \ldots\ldots + w_n x^n$$

$\phi(z)$ values limits in the range [0,1]. It indicates that if z goes to infinity, the function becomes one, and if z goes minus infinity, the function reaches zero.

### B. Decision Tree

The Decision Tree classification algorithm [16,18] works as a human thinking ability while making a decision. The classification model is created by the decision tree algorithm, which generates a decision tree. Each branch descending from that node represents one of the possible values, and each node in a decision tree represents a test for a feature. Because the core structure of Decision Trees is unaffected by the values taken by each feature, they can function efficiently with unnormalized datasets.

### C. Random Forest

When dealing with a dataset with a huge number of features, the decision tree algorithm is prone to overfitting, which complicates the model and learning process. Random Forest (RF) classification algorithm [19,20] is an ensemble Decision Tree classification algorithm that incorporates several weaker models to build a more accurate one.

### D. Ada Boost

Adaptive Boosting (Ada Boost) [16,21] is the first boosting algorithm. It is an iterative ensemble classification algorithm that means weak learners grow sequentially and become strong ones. The classifier should be interactively trained using a variety of weighted training instances. It tries to provide an excellent fit for these instances in each iteration by minimizing training errors.

### E. Gradient Boost

Gradient Boost [22,23] is an ensemble boosting classification algorithm that combines several weak learners into strong learners. Gradient Boosting classification algorithm depends on the loss function. The gradient descent optimization procedure is used to determine the contribution of the weak learner to the ensemble.

### F. K-Nearest Neighbors

K-Nearest Neighbors (KNN) [19,16] is a non-parametric, the lazy classification algorithm that memorizes class labels rather than learning how to discriminate them. Based on the distance metric, this algorithm finds the k samples which are closest to the point to classify. It uses the ‘Minkowski’ distance. It's defined as

$$d(x,y) = \sqrt[p]{\sum_k | x_k - y_k |^p} \tag{2}$$

### G. Naive Bayes

The Naive Bayes (NB) classification algorithm [2,24] is a statistical classification technique based on the Bayes theorem. The NB classifier implies that each feature is independent of the others and that they do not interact so that each feature contributes equally to the probability of a sample associated with a particular class. The NB classifier is easy to use and fast to compute, and it works well with huge datasets with high dimensionality.

Carry out the experiments in this study with the Python programming language using sklearn, pandas and numpy libraries for classification algorithms processing and matplotlib and seaborn libraries for visualization of ROC-AUC curve on a Google Collab with 25 GB of RAM and a TPU environment. CICFlowMeter is network traffic flow generator tool used in this study to generate CSV files from extracted pcap files which are network traffic packet capture files.

## 4. Results & Discussion

A collection of performance metrics is used to evaluate the effectiveness of classification algorithms [25] for detecting DDoS attacks to see how well they classify attacks and benign classifications.

### 4.1. *Confusion Matrix*

The actual and predicted values of label classes are displayed in a confusion matrix. It shows the four key values that are True Positive, False Negative, False Positive, and True Negative. These values are used to calculate the evaluation metrics.

TRUE POSITIVE (TP): The amount of DDoS attacks properly identified by the classifier.
TRUE NEGATIVE (TN): The number of BENIGN class labels accurately detected by the classifier.
FALSE POSITIVE (FP): The number of BENIGN class labels, classified as DDoS attacks by the classifier.
FALSE NEGATIVE (FN): The number of DDoS attack labels, classified as BENIGN class labels by the classifier.

### 4.2. *Accuracy*

Accuracy shows how much the classifier classifies class labels truly. The logistic regression, Gradient Boost, and Naive Bayes models achieved the best classification accuracy of 99.58%. The decision tree model provides poor accuracy. Figure 1 shows the accuracy values bar-chart of the classification algorithms for DDoS attack detection.

$$ACCURACY = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

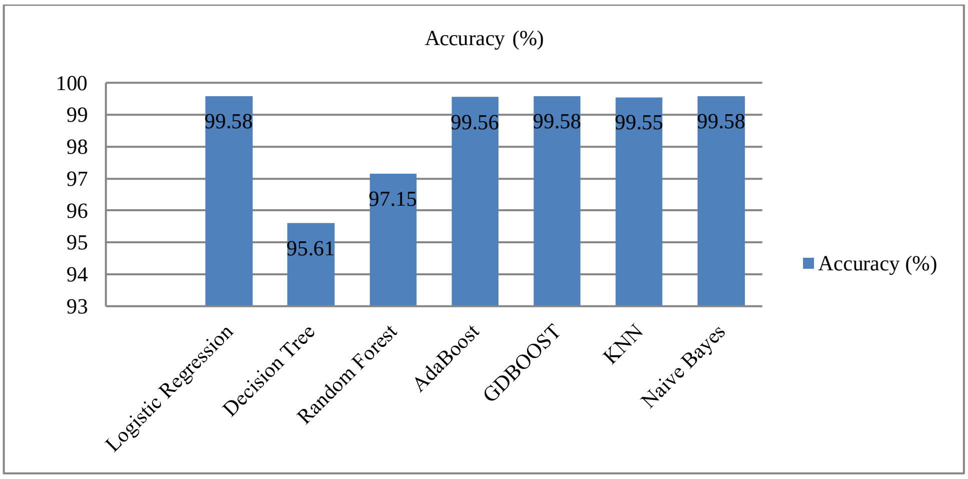

Fig.1. Bar-chart of accuracy values of classification algorithms for DDoS attack detection

### 4.3. *Precision*

Precision shows that the ratio of the truly detected DDoS attacks out of the total DDoS attacks class labels classified by the classifier. Logistic Regression and Naive Bayes provide the best precision. Decision Tree and Random Forest provide poor precision values.

$$PRECISION = \frac{TP}{TP + FP} \tag{4}$$

### 4.4. *Recall*

Recall also known as sensitivity, shows that the ratio of the truly detected DDoS attacks out of total actual DDoS attacks. Ada Boost provides the best recall value. Gradient Boost and KNN provide almost the same recall value as the Ada Boost method. Decision Tree  provides a poor recall value.

$$RECALL = \frac{TP}{TP + FN} \tag{5}$$

### 4.5. *F1 Score*

The F1 score is the weighted average  precision and recall. Logistic Regression, Gradient Boost, and Naive Bayes provide the best F1-score value. Ada Boost and KNN almost provide the best F1-score value. Decision Tree also provides a poor F1-score value.

$$F1 - SCORE = \frac{2 * PRECISION * RECALL}{PRECISION + RECALL} \tag{6}$$

Table 1 displays the Precision, Recall, and F-Score values of the classification algorithms for detecting DDoS attacks. Figure 2 illustrates a bar chart of the classifiers' classification report for detecting DDoS attacks.

Table 1. Classifier precision, recall, and F-Score values for detecting DDoS attacks.

| | Logistic Regression | Decision Tree | Random Forest | AdaBoost | GD BOOST | KNN | Naive Bayes |
|---|---|---|---|---|---|---|---|
| Precision | 0.9981 | 0.9969 | 0.9969 | 0.9969 | 0.9974 | 0.9973 | 0.9981 |
| Recall | 0.9976 | 0.9586 | 0.9742 | 0.9986 | 0.9984 | 0.9982 | 0.9976 |
| F-score | 0.9979 | 0.9774 | 0.9854 | 0.9978 | 0.9979 | 0.9977 | 0.9979 |

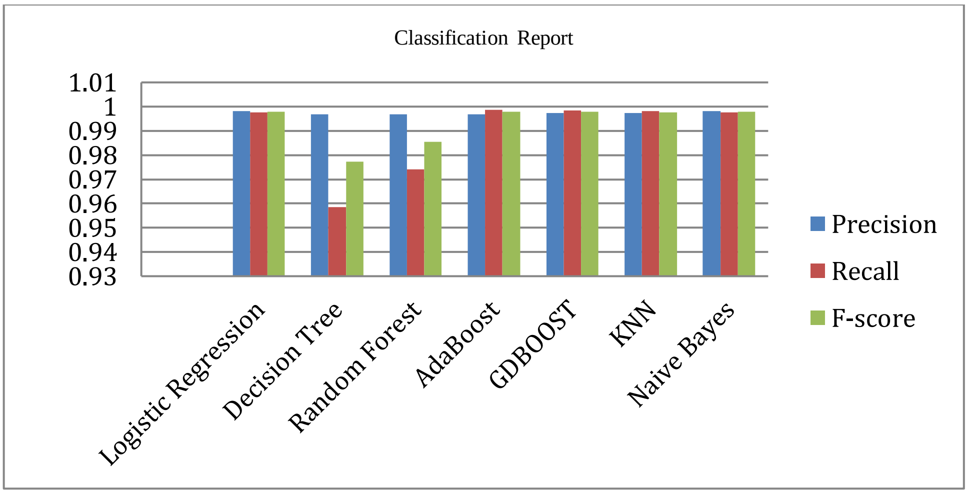

Fig.2. Bar-chart of classifiers Precision, Recall and F-Score values for DDoS attack detection

## 4.6. *Specificity*

Specificity is the ratio of the truly classified BENIGN class labels out of the total actual BENIGN class labels. Logistic Regression and Naive Bayes provide the best specificity of 0.82. Random forest and Ada boost provide a poor specificity value of 0.70. Decision Tree gives an almost poor specificity value of 0.71. Figure 3 shows the bar chart of classification algorithm specificity values.

$$SPECIFICITY = \frac{TN}{TN + FP} \tag{7}$$

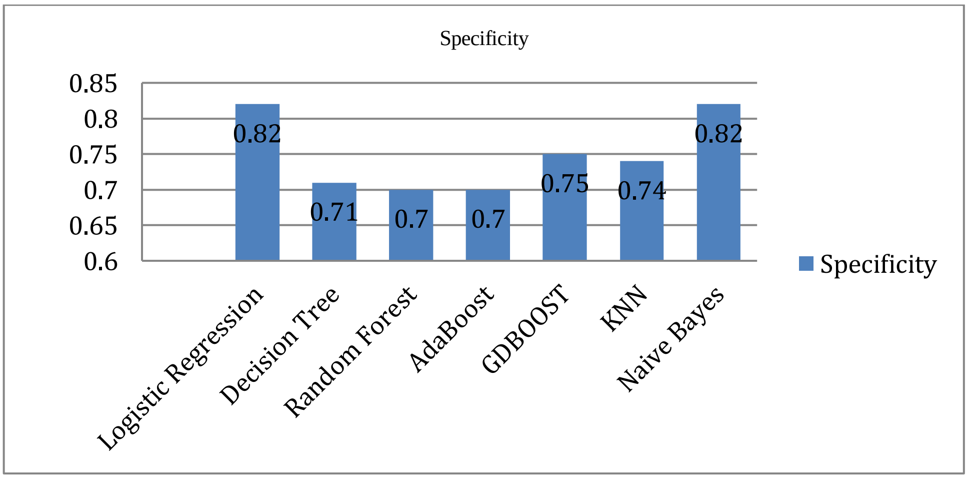

Fig.3. Bar-chart of specificity of classifiers for DDoS attack detection

## 4.7. *ROC-AUC Score*

A ROC curve is a graph that shows a classification model's performance overall decision threshold. A decision threshold is a value used to translate a probabilistic prediction into a class label. Scores between 0 and 1 on the ROC-AUC. When the ROC-AUC value is 1, the classifier correctly classifies all labels. When the ROC-AUC value is zero, the classifier classifies all labels not accordingly, that is, it classifies TRUE labels as FALSE labels and FALSE labels as TRUE labels. With a ROC-AUC score of 0.9680, AdaBoost is the best option. Gradient Boost has about the same performance as the best ROC-AUC score of 0.9676. The decision tree gives a poor ROC-AUC score of 0.8487. Table 2 shows the classification algorithms' ROC-AUC scores for detecting DDoS attacks. Figure 4 shows the ROC curves of the classification algorithms for detecting DDoS attacks.

Table 2. ROC-AUC scores of classification algorithms for DDoS attack detection

| Classification algorithms | ROC-AUC SCORE |
|---|---|
| Logistic Regression | 0.9574 |
| Decision Tree | 0.8487 |
| Random Forest | 0.9072 |
| Ada Boost | 0.9680 |
| Gradient Boost | 0.9676 |
| KNN | 0.9081 |
| Naive Bayes | 0.9529 |

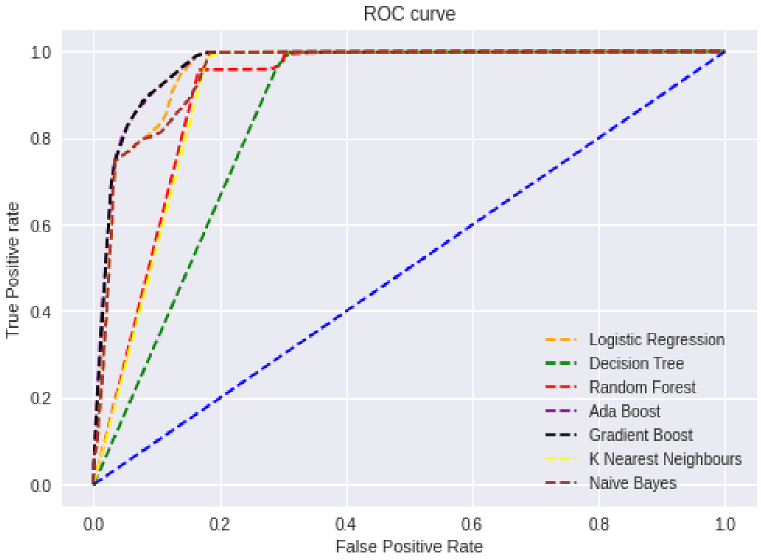

Fig.4. ROC curves of classifiers for DDoS attack detection

### 4.8. *Log Loss*

The most important probability-based classification metric is log loss. The lower the log-loss number, the better the predictions; the log loss value is 0 for a perfect model. Gradient Boost, Naive Bayes, and Logistic Regression provide the best log loss value while the Decision Tree and Random Forest give the poor log loss values. Ada Boost and KNN give good log loss values. Table 3 shows the log loss values of the classification algorithms for DDoS attack detection.

$$Log - Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i \ln p_i + (1-y_i)\ln(1-p_i)] \tag{8}$$

Where N is the number of observations, p is the prediction probability and y is the actual value.

Table 3. Log loss values of the classifiers for DDoS attack detection

| Classification algorithms | Log Loss value |
|---|---|
| Logistic Regression | 0.1459 |
| Decision Tree | 1.5170 |
| Random Forest | 0.9856 |
| Ada Boost | 0.1527 |
| Gradient Boost | 0.1440 |
| KNN | 0.1552 |
| Naive Bayes | 0.1458 |

### *4.9. Cross-Fold Validation*

Cross-fold validation is a statistical method for evaluating machine learning classification models. A test set should still be kept aside for final evaluation when employing Cross-validation, but the validation set is no longer required. The training set is partitioned into k smaller sets in a k-Cross-fold validation. The training data for a model is taken from k-1 folds. After that, the model is tested against the remaining data. Here k-fold cross validation calculated cross-fold validation accuracy and standard deviation values. Gradient Boost, Naive Bayes, KNN, Logistic regression, and AdaBoost provide the best cross-fold validation accuracy values. Decision Tree provides poor cross-fold validation accuracy value. Ada Boost prides the best standard deviation value of 0.0074. Gradient Boost and KNN also provide good standard deviation values. The random forest provides poor standard deviation value. Table 4 shows the cross-fold validation accuracy and standard deviation values in cross-fold validation of classification algorithms for DDoS attack detection.

Table 4. Cross-fold validation accuracy with standard deviation of classifiers.

| Classification algorithms | cross-fold validation Accuracy(%) | Standard Deviation |
|---|---|---|
| Logistic Regression | 99.5769 | 0.0113 |
| Decision Tree | 96.3575 | 0.0251 |
| Random Forest | 97.2008 | 0.0260 |
| Ada Boost | 99.5540 | 0.0074 |
| Gradient Boost | 99.5792 | 0.0092 |
| KNN | 99.5563 | 0.0095 |
| Naive Bayes | 99.5771 | 0.0114 |

### *4.10. Run Time*

The KNN has taken more time for execution than other models to classify the labels. Naive Bayes has taken very little time for execution to classify labels. Logistic regression, Decision Tree, and Ada Boost models were also executed very fast. Random forest and Gradient Boost took more execution time, but that is less than the KNN model. Table 5 shows the classification algorithm's run times in seconds for DDoS attack detection.

Table 5. Run-time in seconds of the classification algorithms for DDoS attack detection

| Classification algorithms | Run time (Sec) |
|---|---|
| Logistic Regression | 4.1140 |
| Decision Tree | 12.7741 |
| Random Forest | 686.2687 |
| Ada Boost | 54.2303 |
| Gradient Boost | 202.6695 |
| KNN | 3390.3411 |
| Naive Bayes | 0.3533 |

### *4.11. Discussion*

The logistic regression, Gradient Boost, and Naive Bayes models achieved the best overall classification accuracy and their cross-fold validation accuracy values also same with small standard deviation values. Among these Gradient Boost gives very small standard deviation value. Logistic Regression, Gradient Boost, and Naive Bayes also provide the best F1-score and log-loss values. Logistic Regression and Naive Bayes provide the best precision and specificity values. Gradient Boost gives better precision and specificity values. Gradient Boost gives better recall and ROC-AUC score values than Logistic Regression and Naive Bayes which produce same good recall and ROC-AUC score values. Naive Bayes and Logistic regression executed faster than Gradient Boost. Ada Boost provides the best recall and ROC-AUC score values while it produces poor specificity value. Ada Boost provides better accuracy, precision, F1-score and log-loss value. Ada Boost execution time also good. Ada Boost cross-fold validation accuracy same as its overall classification accuracy. Ada Boost gives best standard deviation value in cross-fold validation. KNN produces better values in accuracy, recall. KNN produces good values in precision, F1-score specificity and log-loss values. KNN produces poor values in the ROC-AUC score and execution time. KNN produces same overall accuracy and cross-fold validation accuracy with small standard deviation value. Random Forest and Decision tree classification algorithms produce a poor classification results than other classification in this study.

## 5. Conclusions

This quantitative research proved that machine learning classification algorithms detect DDoS attacks accurately in very less time by evaluating the effectiveness of the classification algorithms for detecting DDoS attacks on the

CICDDoS2019 dataset. Logistic regression, Gradient Boost, and Nave Bayes produced the best accuracy, F-score and specificity values in the experiments. The best precision value is obtained using logistic regression and Nave Bayes. The best levels of precision and ROC AUC values can be found with AdaBoost. The log loss value produced by the Gradient boosting algorithm is the best. Naïve Bayes executed very fast. Gradient Boost produces the best accuracy, according to cross-fold validation. This research concludes AdaBoost and Gradient Boost algorithms give the best classification results, Logistic Regression, KNN, and Naive Bayes give good classification results. When compared to other classification methods, Decision Tree and Random Forest produce poor classification results. For DDoS attack detection in future work concentrate additionally on computation space to detect DDoS attacks accurately in less time and less computation space, feature selection approaches [26] will be integrated with machine learning classification algorithms to achieve the target in future study.

## References

[1] Kaur, G. (2020). A comparison of two hybrid ensemble techniques for network anomaly detection in spark distributed environment. *Journal of Information Security and Applications*, *55*, 102601. https://doi.org/10.1016/J.JISA.2020.102601

[2] Kishore Babu Dasari, Dr Nagaraju Devarakonda. (2018). Distributed denial of service attacks, tools and defense mechanisms. *International Journal of Pure and Applied Mathematics*, *120*(6), 3423–3437. https://acadpubl.eu/hub/2018-120-6/3/247.pdf

[3] Lu, K., Wu, D., Fan, J., Todorovic, S., & Nucci, A. (2007). Robust and efficient detection of DDoS attacks for large-scale internet. *Computer Networks*, *51*(18), 5036–5056. https://doi.org/10.1016/J.COMNET.2007.08.008

[4] Bouyeddou, B., Kadri, B., Harrou, F., & Sun, Y. (2020). DDOS-attacks detection using an efficient measurement-based statistical mechanism. *Engineering Science and Technology, an International Journal*, *23*(4), 870–878. https://doi.org/10.1016/j.jestch.2020.05.002

[5] Nilesh Vishwasrao Patil, C. Rama Krishna, Krishan Kumar, Distributed frameworks for detecting distributed denial of service attacks: A comprehensive review, challenges

[6] X. Liang and T. Znati, "An empirical study of intelligent approaches to DDoS detection in large scale networks," 2019 International Conference on Computing, Networking and Communications (ICNC), 2019, pp. 821-827, doi: 10.1109/ICCNC.2019.8685519.

[7] Suresh, M., & Anitha, R. (2011). Evaluating Machine Learning Algorithms for Detecting DDoS Attacks *. In *CCIS* (Vol. 196).

[8] L. Feinstein, D. Schnackenberg, R. Balupari and D. Kindred, "Statistical approaches to DDoS attack detection and response," *Proceedings DARPA Information Survivability Conference and Exposition*, 2003, pp. 303-314 vol.1, doi: 10.1109/DISCEX.2003.1194894.

[9] Majed, H., Noura, H. N., Salman, O., Malli, M., & Chehab, A. (2020). Efficient and secure statistical DDoS detection scheme. *ICETE 2020 - Proceedings of the 17th International Joint Conference on e-Business and Telecommunications*, 153–161. https://doi.org/10.5220/0009873801530161

[10] T. Roempluk and O. Surinta, "A Machine Learning Approach for Detecting Distributed Denial of Service Attacks," *2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON)*, 2019, pp. 146-149, doi: 10.1109/ECTI-NCON.2019.8692243.

[11] K. B. V., N. D. G. and P. S. Hiremath, "Detection of DDoS Attacks in Software Defined Networks," *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*, 2018, pp. 265-270, doi: 10.1109/CSITSS.2018.8768551.

[12] Batchu, R. K., & Seetha, H. (2021). A generalized machine learning model for DDoS attacks detection using hybrid feature selection and hyperparameter tuning. *Computer Networks*, *200*, 108498. https://doi.org/10.1016/J.COMNET.2021.108498

[13] Y. Chen, X. Ma and X. Wu, "DDoS Detection Algorithm Based on Preprocessing Network Traffic Predicted Method and Chaos Theory," in IEEE Communications Letters, vol. 17, no. 5, pp. 1052-1054, May 2013, doi: 10.1109/LCOMM.2013.031913.130066.

[14] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," in IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, Second quarter 2016, doi: 10.1109/COMST.2015.2494502.

[15] W. Zhijun, L. Wenjing, L. Liang and Y. Meng, "Low-Rate DoS Attacks, Detection, Defense, and Challenges: A Survey," in IEEE Access, vol. 8, pp. 43920-43943, 2020, DOI: 10.1109/ACCESS.2020.2976609.

[16] Dasari, K.B., Devarakonda, N. (2021). Detection of different DDoS attacks using machine learning classification algorithms. Ingénierie des Systèmes d'Information, Vol. 26, No. 5, pp. 461-468. http://dx.doi.org/10.18280/isi.260505

[17] Sambangi, S., & Gondi, L. (2020). A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression. *Proceedings*, *63*(1), 51. https://doi.org/10.3390/proceedings2020063051

[18] Santos, R., Souza, D., Santo, W., Ribeiro, A., & Moreno, E. (2020). Machine learning algorithms to detect DDoS attacks in SDN. *Concurrency and Computation: Practice and Experience*, *32*(16). https://doi.org/10.1002/cpe.5402

[19] D. Firdaus, R. Munadi and Y. Purwanto, "DDoS Attack Detection in Software Defined Network using Ensemble K-means++ and Random Forest," 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2020, pp. 164-169, doi: 10.1109/ISRITI51436.2020.9315521.

[20] Y. Chen, J. Hou, Q. Li and H. Long, "DDoS Attack Detection Based on Random Forest," 2020 IEEE International Conference on Progress in Informatics and Computing (PIC), 2020, pp. 328-334, doi: 10.1109/PIC50277.2020.9350788.

[21] Shahraki, A., Abbasi, M., & Haugen, Ø. (2020). Boosting algorithms for network intrusion detection: A comparative evaluation of Real AdaBoost, Gentle AdaBoost and Modest AdaBoost. *Engineering Applications of Artificial Intelligence*, *94*, 103770. https://doi.org/10.1016/J.ENGAPPAI.2020.103770

[22] Alamri, H. A., & Thayananthan, V. (2020). Bandwidth control mechanism and extreme gradient boosting algorithm for protecting software-defined networks against DDoS attacks. *IEEE Access*, *8*, 194269–194288.

https://doi.org/10.1109/ACCESS.2020.3033942

[23] Z. Chen, F. Jiang, Y. Cheng, X. Gu, W. Liu and J. Peng, "XGBoost Classifier for DDoS Attack Detection and Analysis in SDN-Based Cloud," 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), 2018, pp. 251-256, doi: 10.1109/BigComp.2018.00044.

[24] 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), 2016, pp. 104-107, doi: 10.1109/TSP.2016.7760838.

[25] X. Liang and T. Znati, "An empirical study of intelligent approaches to DDoS detection in large scale networks," 2019 International Conference on Computing, Networking and Communications (ICNC), 2019, pp. 821-827, doi: 10.1109/ICCNC.2019.8685519.

[26] Mekala, S., Padmaja Rani Supervisor, B., & Padmaja Rani, B. (2020). Article ID: IJARET_11_11_121 Kernel PCA Based Dimensionality Reduction Techniques for Preprocessing of Telugu Text Documents for Cluster Analysis. *International Journal of Advanced Research in Engineering and Technology*, *11*(11), 1337–1352. https://doi.org/10.34218/IJARET.11.11.2020.121

## Authors' Profiles

**Kishore Babu Dasari** born in Vaivaka, on 26th April, 1983, Pursuing Ph.D in Acharya Nagarjuna university, Andhrapradesh, India in Computer Science and Engineering. Studied M.Tech (Software Engineering) in Avanthi Institute of Technology & Science and B. Tech (CSE) in Gudlavalleru Engineering College.

He is working at the Keshav Memorial Institute of Technology as an Assistant Professor in the CSE Department. He has total 13 years of teaching experience.

He is the member of the CSI professional Society. He qualified UGC NET. He awarded NPTEL translator for translating one course into Indian regional language Telegu.

**Dr. Nagaraju Devarakonda** is working as an Associate Professor at VIT-AP University in the School of Computer Science and Engineering. He has 20 years teaching experience. His total number of research publications is 68. His research domains are data mining, machine learning, soft computing and optimization techniques.

**How to cite this paper:** Kishore Babu Dasari, Nagaraju Devarakonda, "Detection of DDoS Attacks Using Machine Learning Classification Algorithms", International Journal of Computer Network and Information Security(IJCNIS), Vol.14, No.6, pp.89-97, 2022. DOI:10.5815/ijcnis.2022.06.07