IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

# Distributed Denial of Service Attack Prediction Using Novel Ensemble Model

**Nikhil Anand Mahendrakar[1], Sanjay A S[2], Manoj M[3]**

Siemens Healthineers, Bengaluru, India[1-3]

**Abstract**: In the modern world, internet has become a major part in everyone's life. Most of businesses have now gone digital and internet is playing a major role in their success. Internet has enabled businesses to have larger market to which they can sell their products. One important component of the internet is the server. All the devices in the internet are broadly either a server or a client. Server stores the data that are then accessed by the clients. Distributed Denial-of-Service is a network attack on the servers. This attack targets the availability of the server by overwhelming them by a flood of internet traffic. This prevents the server from giving services to the legitimate users. DDoS attack causes significant losses to the organization. In this paper, we discuss a ensemble modelling based approach to mitigate this problem. Ensemble modelling is a process where multiple model which are different fundamentally are combined to make one classifier model.We try to predict the type of DDoS attack by developing a ensemble model by combining Naïve Bayes, Random Forest, Multilayer Perceptron, Stochastic Gradient Descent . The accuracy score of 98.62 percent has been achieved. It can be concluded that this system proves to be an effective deterrent for DDoS attacks.

**Keywords:** DDoS attack,http-flood, udp-flood, smurf , machine learning , ensemble model

## I. INTRODUCTION

Advancement in communication and information technology further strengthens the role of internet in business. Internet is widely used in organization for marketing and promotion of products and services. The concept of the internet (sometimes also referred to as World Wide Web) was conceived by a British computer scientist Tim-Berners Lee in the year 1989. The number of people coming online since then has increased exponentially. It is also important to note here that internet is very recent phenomena spanning just decades ago. We are still in very early stages of the world wide web technology. The internet has indeed changed a lot of things but there is still scope of lot of development left.

As discussed earlier, the increased popularity of the internet has bought various advantages as well as new challenges. One of the important challenge for data stored in the internet is to make the data always available to the user (Availability) , there must be no illegal edits to the data (Integrity) and there should be privacy for the data (Confidentiality). This is also popularly called as the CIA triad.

Distributed Denial of Service attack is caused by various freely available compromised host which create a fake flood of internet traffic. This compromised hosts are also sometimes referred to as "zombies". Distributes Denial of Service is used to violate the foundational security principle of the internet – CIA (Confendiality , Integrity, Availability).

It is usually not possible for an attacker to cause a Denial of Service attack using a single host. Therefore, the attacker attacks the various vulnerable devise present on the internet with malware. Once, this devices are compromised , then the attacker uses those devices for causing a flood of internet traffic to the server. Usually the owners of those devices are not even aware that they have been compromised. The identity of the attacker is also hidden since the attack is taking place through the compromised devices which seem as legitimate clients. The Server is then serving the compromised devices and the legitimate clients are not able to access the server. This Causes huge loss to the organization.

In this paper we try to solve the following problem using machine learning based ensemble modelling technique. Most of the dataset that are available publicly are not equipped with the latest DDoS attacks. The Dataset used in this paper has covered the latest DDoS attacks that are being carried out. The dataset has around 2 million rows and 28 columns. Feature Selection on the attributes has been done to reduce the number of feature for the prediction without affecting the accuracy. Feature Engineering (Hashing based approach discussed in the detail in the Implemantation section) on the values of the attributes has been done to convert from string representation to numerical representation for various computing purposes. A ensemble based modelling has been proposed between Naïve Bayes , Random Forest , Multilayer Perceptron and Stochastic Gradient Descent, for achieving a higher accuracy by training with much less data.

  

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

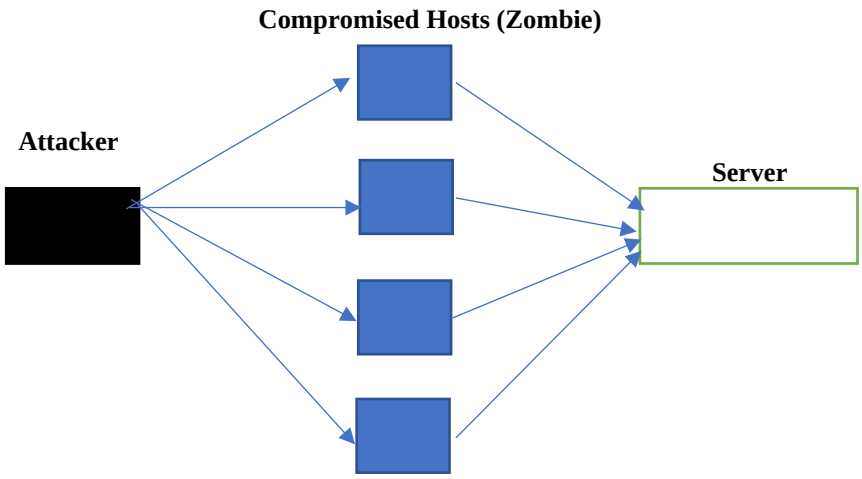


**Fig. 1 Typical Structure of a DDoS attack**

## II. LITERATURE SURVEY

In paper [1] , author has used Random Forest Classifier technique for classifying the packet into normal or anomalous packets. The dataset used is the famous NSL-KDD dataset which contains just 23000 odd rows. The number of features used in the dataset is 42. Similar accuracy can be achieved with much less attributes. More number of attributes increases the time for model building and also increases the complexity. Also, it is a bi-classification problem that means it classifies the packet into two classes normal and anamolous. Our paper presents a method to detect accurately which type of DDoS attack it is.

In paper [2], author has used the combination of both supervised and unsupervised machine learning technique. Agglomerative Clustering is used at the first step to form a cluster of packets which is similar to each other and then they are labelled manually. This labelled data, then acts as an input for building the model. Author has used k-Nearest Neighbors (kNN) , Support Vector Machine (SVM) and Random Forest (RF) for building the model and achieved the accuracy of 95% , 92% and 96.6%. The accuracy achieved in this paper is around 98 percent.

In paper [3] , author used the firewall based approach. A firewall is a system which acts as the first point of entrance for the packet to enter into a server or a computing device. Access control list (ACL) is used to check whether the packet is authentic or not. The IP addresses are used for differentiating anamolous and normal addresses. Firewalls do not provide efficient solution when the attack size increases.

In paper [4] , author has used to classify the ICMPv6 DDoS attack. ICMP is an abbreviation for Internet Control Message Protocol which is used for sending control and error messages. They have used popular classification techniques like Neural Network , Decision Tree , Naïve Bayes , Support Vector Machine and K-Nearest Neighbors and compared the results.

In paper [5], the author uses Ensemble modelling technique , where 4 classifier models are trained and voting is done among the models for the final prediction. The dataset used in this research paper is the famous NSL-KDD dataset. The four classification algorithms used are MLP (NN), SMO (SVM), IBK (KNN) and J48 (DT-C4.5). Hard Voting is used to combine the prediction of different models to get the final outcome. However, Hard Voting ignores the fact that each model has different accuracy and gives equal weightage to each model while giving the final outcome.

In paper [6] , The authors propose a RNN-based model, referred to as DeepDefense. Unlike previously discussed methods, the model does not use machine learning based techniques. Instead, the authors extracted 20 features from packet headers and applied sliding windows to separate continues network traffic into sequences of network packets. For a given sequence of packets, the model classifies the last packet either as a legitimate or attack traffic. Their proposed deep learning model achieves lower error rate than traditional machine learning techniques. However, reliance on packet-by-packet.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

## III. METHODS AND MATERIALS

The following chapter is further divided into multiple sub-chapters where each sub-chapter describes the multiple processes for achieving the final result.

### A. Data

The data is taken from the public dataset [7] . The data consists of the modern DDoS (Denial Distributed of Service) attacks. The attacks include the HTTP Flood, UDP Flood , SIDDOS, Normal, Smurf. These attacks do not cover the entire DDoS attacks but contain the most common ones. The data is present in the text format with the values for each features separated using comma . There are around 2 million rows and 28 features. We believe we have trained the model with all possible scenario. Some of the features which do not contribute in the enhancement of the accuracy have been removed .


**Fig. 2 Sample of the Structure of the Data**

```
@attribute SRC_ADD numeric 0
@attribute DES_ADD numeric 1
@attribute PKT_ID numeric 2
@attribute FROM_NODE numeric 3
@attribute TO_NODE numeric 4
@attribute PKT_TYPE {tcp,ack,cbr,ping} 5
@attribute PKT_SIZE numeric 6
@attribute FLAGS {-------,---A---} 7
@attribute FID numeric 8
@attribute SEQ_NUMBER numeric 9
@attribute NUMBER_OF_PKT numeric 10
@attribute NUMBER_OF_BYTE numeric 11
@attribute NODE_NAME_FROM {Switch1,Router,server1,router,clien-4,client-2,Switch2,client-5,clien-9,clien-2,clien-1,clien-14,clien-5,clien-11,clien-13,clien-0,switch1,client-4,clienthttp,clien-7,clien-19,client-14,clien-12,clien-8,clien-15,webserverlistin,client-18,client-1,switch2,clien-6,client-10,client-7,webcache,clien-10,client-15,clien-3,client-17,client-16,clien-17,clien-18,client-12,client-8,client-0,clien-16,client-13,client-11,client-6,client-3,client-9,clien-19,http_client}
@attribute NODE_NAME_TO {Router,server1,Switch2,Switch1,clien-1,clien-5,clien-7,switch1,clien-11,clien-15,clien-13,clien-3,clien-9,clien-6,router,clien-4,clien-14,switch2,clien-8,clienthttp,webcache,clien-10,clien-12,webserverlistin,clien-0,clien-2,http_client,client-13,client-9,client-1,client-19,client-4,client-17,client-7,client-3,client-12,client-2,clien-18,client-16,clien-17,client-0,clien-16,client-18,client-5,client-11,client-14,client-8,client-6,client-10,clien-19,client-15}
@attribute PKT_IN numeric 14
@attribute PKT_OUT numeric 15
@attribute PKT_R numeric 16
@attribute PKT_DELAY_NODE numeric 17
@attribute PKT_RATE numeric 18
@attribute BYTE_RATE numeric 19
@attribute PKT_AVG_SIZE numeric 20
@attribute UTILIZATION numeric 21
@attribute PKT_DELAY numeric 22
@attribute PKT_SEND_TIME numeric 23
@attribute PKT_RESEVED_TIME numeric 24
@attribute FIRST_PKT_SENT numeric 25
@attribute LAST_PKT_RESEVED numeric 26
@attribute PKT_CLASS {Normal,UDP-Flood,Smurf,SIDDOS,HTTP-FLOOD}       target variable(27) .
```

**Fig. 3 Features of the dataset**

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

B. Data Pre-Processing

The data is present in the text format and requires pre-processing. In the file, each row is taken as one whole string and then converted to a string array using split function available in python. A separate list data structure is created for every feature. The value is then put to its respective list.

C. Conversion to Numeric Value

Some of the features had values which are of string datatype and had to be converted to a numeric value. The task here is to represent the each string value uniquely using a numeric value. The reason for this is that each of the string has a unique meaning associated with it. For example, consider the feature node_name_from which says from which type of node the packet is being sent. For this conversion, we use hashing. Hashing is the process of transforming string into a unique key and the value of the hash for a different string is completely different. The value returned by the hash function is also a string which can be then converted to a numeric value using ascii of the characters in the string. We have taken summation of the ascii value of the characters in the hash. This process makes sure that each unique string is assigned unique integer value for computation. However, for the values of the target variable we have assigned unique number as shown the figure 4.

```
for i in range(len(templist)):
    temptext = templist[i]
    hash_object = h.md5(temptext.encode())
    md5_hash = hash_object.hexdigest()
    value = 0
    for j in range(len(md5_hash)):
       value = value + ord(md5_hash[j])
    valuelist[templist[i]] = value
return valuelist
```

**Fig. 3 Python Implementation for Conversion to Numeric Value**

```
{'HTTP-FLOOD\n': 2, 'UDP-Flood\n': 3, 'SIDDOS\n': 4, 'Normal\n': 5, 'Smurf\n': 6}
```

**Fig. 4 Numeric value assigned for the values of the target variable**

D. Scaling of data

Scaling of data is important for getting the maximum accuracy for any classifier. Scaling can be defined as the process of fitting the data on a specific scale. Consider an example of a data which consists of currencies like rupee and dollar. We know that 1 US dollar = 80 rupees . So therefore, for consistent results, the whole data should be converted to either dollars or rupee. If scaling is not done on the data then the classifier considers a difference in price of 1 rupee as important as a difference of 1 US dollar. This would lead to wrong results and reduce the effectiveness of the classifier.

E. Ensemble Learning

In this paper we implement a ensemble model which is combination of four classifiers which are Stochastic Gradient Classifier , Multilayer Perceptron (also known as Artificial Neural Network) , Random Forest Classifier and Naïve Bayes Classifier. Brief working of each of the classifier has been given for completeness. Ensemble model is created to combine the results and improve the accuracy of the model .We combine the results of these models by using a mathematical function which is discussed below.

In the first step, we assign weights for each of the model. The weights signify how much the classifier shall contribute in giving the final prediction. Weights are the accuracy of the classifier when run seperately on the dataset. Consider the Table I for weights assigned to the classifiers.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

Table I

| Serial No. | Classifier | Weights (Accuracy) |
|---|---|---|
| 1 | Stochastic Gradient Descent | 98.50 % |
| 2 | Multilayer perceptron | 98.62 % |
| 3 | Random Forest | 97.23 % |
| 4 | Naïve Bayes | 96.97 % |

Figure 5 shows the 2 dimensional list containing the class probabilities of each data point for stochastic gradient descent..Class probabilites are nothing but the probability of the values of the target-variable given the values of the feature variable have already occurred. Now , in order to predict the final result we multiply the weight of the each classifier with the class probability. Once it is computed for all the classes (or values) of the target variable and the one which has the maximum value is chosen as the final prediction . Consider figure 6 which is the python implementation of the above idea . We design iteration on the values (class) of the target variable (type of distributed denial of service) and calculate according the equation shown figure 6. Once the calculation is completed for all the values of the target variable , then simply the one with maximum number is chosen as the predicted value.

```
[[0.00000826 0.01097373 0.00027594 0.98526049 0.00348158]
 [0.00000262 0.01066955 0.00022107 0.98515328 0.00395347]
 [0.00000122 0.01039839 0.00015766 0.98524586 0.00419688]
 ...
 [0.00000171 0.01053096 0.00019098 0.98517701 0.00409933]
 [0.00002233 0.01083333 0.00016394 0.98444659 0.0045338 ]
 [0.0000014  0.01058564 0.0002341  0.98541981 0.00375906]]
0.9850214790307719
```

**Fig. 5 two dimensional list of class probabilities of each data point for stochastic gradient descent**

```
for i in range(len(X_test)):
    listofprobabilities.clear()
    dict_listofprobabilities.clear()
    j=0
    for key in dict_integer_target_variable:
        dict_listofprobabilities[dict_integer_target_variable[key]] = weightdecisiontree*y_probdecisiontree[i][j] + weightgnb*y_probgnb[i][j]+weightmultilayerperceptron*y_probmultilayerperceptron[i][j]+weightsgd*y_probsgd[i][j]
        listofprobabilities.append(weightdecisiontree*y_probdecisiontree[i][j] + weightgnb*y_probgnb[i][j]+weightmultilayerperceptron*y_probmultilayerperceptron[i][j]+weightsgd*y_probsgd[i][j])
        j=j+1
    listofprobabilities.sort()
    searchvalue = listofprobabilities[len(listofprobabilities)-1]
    counter = counter + 1
    for keys in dict_listofprobabilities:
        if float(dict_listofprobabilities[keys])==float(searchvalue):
            y_test_pred.append(keys)
print(counter)
```

**Fig. 6 Python implementation of the ensemble model**

For better understanding , we predict the value of the target variable for the first data point using our ensemble learning apporach. The class probabilities of the first data point for naïve baiyes figure 7 , stochastic gradient descent figure 5, multilayer perceptron figure 8, random forest figure 9 is shown.

```
[[0.         0.         0.         0.99999998 0.00000002]
```

**Fig. 7 Naïve bayes class probabilities for first data point**

```
[[0.         0.00866969 0.00013761 0.98697545 0.00421724]
```

**Fig. 8 Multillayer perceptron class probabilities for first data point**

  

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747


**Fig. 9 Random forest class probabilities for first data point**

Now, to predict the value of the target variable we mulitply the weight and the class probability. The f(x) represents the function of the ensemble learning and x takes all the values of the target variable. Therefore,the numeric value for the first class is f(x1) = 0.0 * 0.9697 + 0.0 * 0.9862 + 0.0*0.9723 + 0.00000826 * 0.985 and f(x1) = 0 . f (x2) = 0.0 * 0.9697 + 0.00866969 * 0.9862 + 0.0*0.9723 + 0.010973 * 0985 and f(x2) 0.0193591723. Similarly, f(x3) = 0.0040 , f(x4) = 3.88593 , f(x5) = 0.00758839. Clearly, f(x4) has the maximum value and x4 is predicted as the ouput for the target variable. x4 represents Normal packet. So , the first datapoint corresponds to a normal packet.

This approach takes into account that not all classifiers have same accuracy and classifiers which have better accuracy should have more contribution in the prediction. This method has increased the accuracy of each of the classifer and given a more robust , effective and more accurate model.

F. Naïve Bayes

Naive Bayes Classifier is a probability based classifier. The main mathematical concept behind the usage of this classifier is bayes theorem.Bayes theorem is used to calculate the probability of an event A happening given that the event B has already occurred. Therefore, it can be said that the event B is the evidence and event A is the hypothesis. It is important to understand here that Naive Bayes assumes the features are independent. The presence of one particular feature does not affect the other. Since , broadly the classification problem can be further divided into either multiclass prediction or biclassification (yes or no). The mathematical formula differs for both the types. Consider the Figure 10 for bi-classification problems and Figure 11 for multiclass problem [8].

$$P(y|x_1,...,x_n) \propto P(y) \prod_{i=1}^{n} P(x_i|y)$$

**Fig. 10.Bayes Theorem for bi-classification**

$$P(y|x_1,...,x_n) \propto P(y) \prod_{i=1}^{n} P(x_i|y)$$

**Fig. 11 Bayes Theorem for multiclass classification**

G. Stochastic Gradient Descent

Stochastic Gradient Descent Classifier has two main mathematical concepts. The first concept is the Gradient Descent Algorithm and the other part being Stochastic. Stochastic in plain term means random. Gradient Descent is an optimization algorithm. It is used to find local manima of a differential function. Consider an example of a parabola as shown in the figure 12. In order to find the local minima dx/dy should be equal to zero. Gradient Descent takes larger step size when the value is far the from the local minima and small step sizes when the value is near the local minima. As mentioned earlier stochastic means random. However the dataset consists of multiple rows (or samples) and each row having multiple features and each contributing to the prediction of the target variable (type of Distributed Denial of Service attack) .Stochastic Gradient Descent randomly takes one feature randomly instead of iterating over all features in the samples of the dataset. This reduces the computation enormously and makes it suitable as a classifier.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

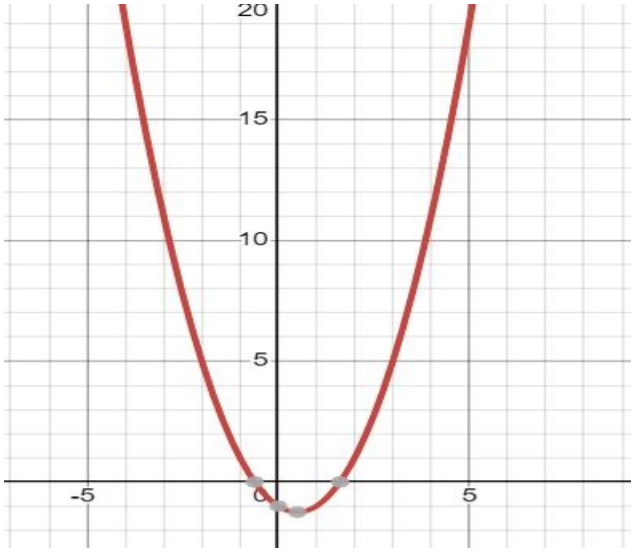


**Fig. 12 Diagram of a parabola**

H. Multilayer Perceptron

Multilayer Perceptron or sometimes also referred as the Neural Network or the Artificial Neural Network is based on the simple model of the brain. The builiding block of the Artificial Neural Network is the neurons. Each neuron takes weighted input signals and computes the output signal based on the activation function. There are mainly three layers in the multilayer perceptron – input layer, hidden layer, output layer. Usually , there is only one input layer,one output layer and many hidden layers. The activation function is the sigmoid function. The sigmoid activation function takes real numbers as input and converts it into value between 0 and 1 .

I. Random Forest

Random forest Classifier applies decision tree algorithm on various subsets of the dataset and combines the results for prediction of the target variable. Decision tree constructs a flow-chart like tree structure where the internal node denotes the features , the branch denotes the values that the feature holds and the leaf node consists of the values held by the target variable.

## IV. RESULTS AND DISCUSSION

The ensemble model is based on the mathematical equation as discussed in the implementation section. The ensemble model has an accuracy of 98.621 percent which is higher than all the individual classifier models. Though the difference in accuracy is in the range of 0.5 percent to 1.5 percent as compared to the famous classifiers used in this paper , however 1 % increase in accuracy on a dataset containing 2 million dataset is predicting 20,000 more rows correctly which is a very good increase. Any machine learning model is judged based on the f1-score, precision and recall score.

Precision score quantifies the number of positive class predictions that actually belong to the positive class [9] .The figure 13 is the classification report .Recall quantifies the number of positive class predictions made out of all positive examples in the dataset. F-Measure provides a single score that balances both the concerns of precision and recall in one number.Confusion matrix as shown in the figure 14.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11747

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 2 | 0.98 | 0.94 | 0.96 | 1352 |
| 3 | 1.00 | 0.90 | 0.95 | 66274 |
| 4 | 0.88 | 0.95 | 0.91 | 2130 |
| 5 | 0.99 | 1.00 | 0.99 | 639048 |
| 6 | 0.96 | 0.32 | 0.48 | 4217 |
| accuracy | | | 0.99 | 713021 |
| macro avg | 0.96 | 0.82 | 0.86 | 713021 |
| weighted avg | 0.99 | 0.99 | 0.99 | 713021 |

**Fig. 13 Classification Report**

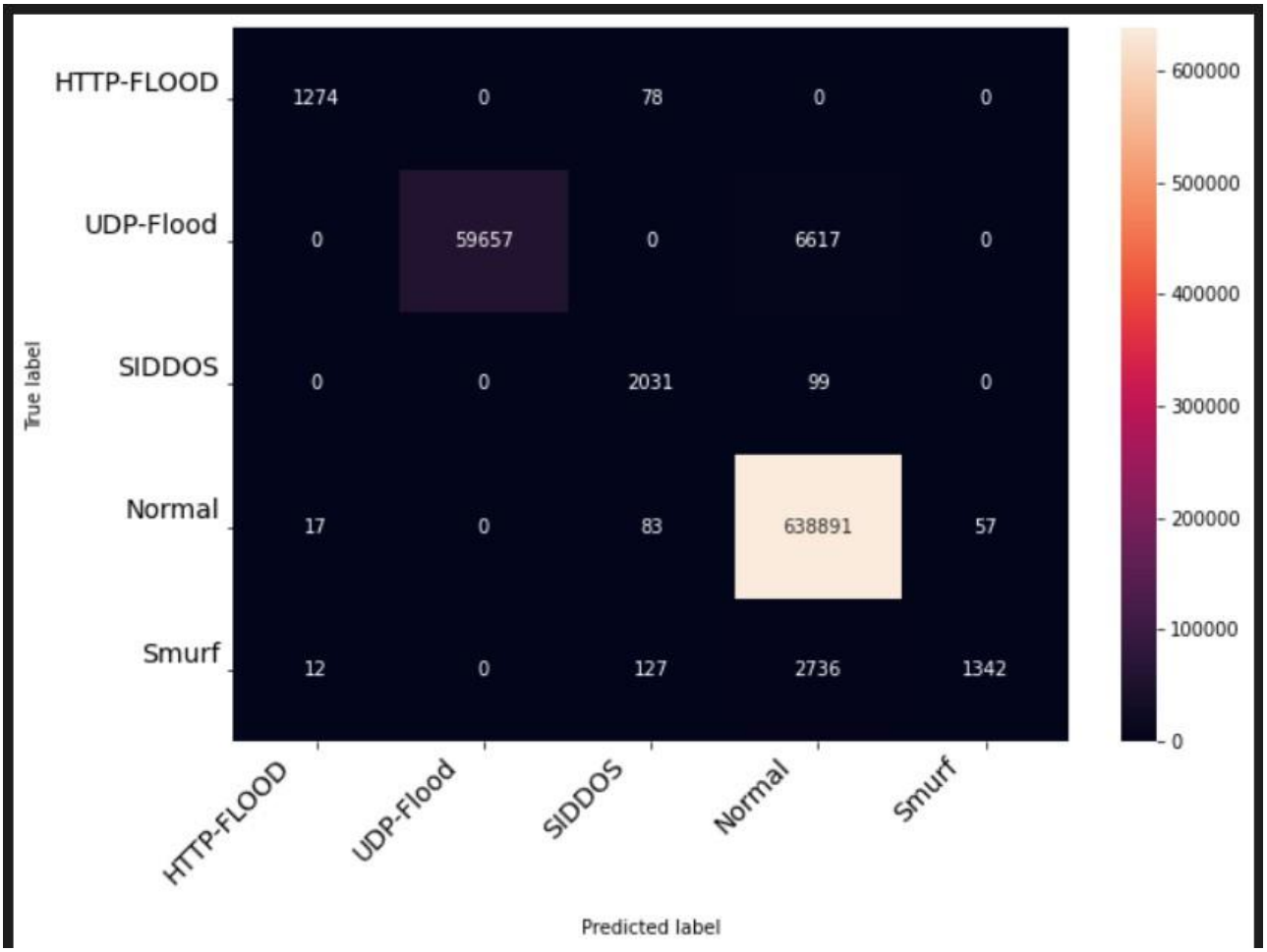


**Fig. 14 Confusion matrix**

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11747

## V. CONCLUSION

Distributed Denial of Service (DDoS) attacks are a serious threat to the network security of various servers. With the ever expanding business activities on the internet, DDoS attack can cause billions of dollars in revenue for the organizations. They attacks have the capability of denying the services for the legitimate user by keeping the server busy by sending a flood of packets. In this paper, we have given a possible solution to help mitigate this problem. The novel ensemble model designed in this paper has an acuracy of 98.62%. The ensemble model has been developed using the four famous classifiers that is naïve bayes , stochastic gradient descent , multilayer perceptron and random forest. The data contains of total 2 million dataset and 28 features.

## REFERENCES

[1]. Pande, S., Khamparia, A., Gupta, D., Thanh, D.N.H. (2021),“ DDOS Detection Using Machine Learning Technique. In: Khanna, A., Singh, A.K., Swaroop, A. (eds) Recent Studies on Computational Intelligence“, Studies in Computational Intelligence, vol 921. Springer, Singapore.
[2]. NG, B. A., & Selvakumar, S. (2019), “Deep radial intelligence with cumulative incarnation approach for detecting denial of service attacks”, Neurocomputing.
[3]. M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in Proc. ECOC’00, 2000, paper 11.3.4, p. 109.
[4]. Elejla, Omar E., Bahari Belaton, Mohammed Anbar, Basim Alabsi, and Ahmed K. Al-Ani, "Comparison of classification algorithms on ICMPv6-based DDoS attacks detection." In Computational Science and Technology, pp. 347-357,Springer, Singapore, 2019.
[5]. S. Das, A. M. Mahfouz, D. Venugopal and S. Shiva, "DDoS Intrusion Detection Through Machine Learning Ensemble," 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C), 2019, pp. 471-477.
[6]. X. Yuan, C. Li and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," 2017 IEEE International Conference on Smart Computing (SMARTCOMP), 2017, pp. 1-8.
[7]. Dataset used in this paper which contains 2 million rows and 28 features.[Online]. Available: https://www.kaggle.com/datasets/jacobvs/ddos-attack-network-logs?resource=download.
[8]. Geeksforgeeks article describing naïve bayes classifier. [Online]. Available: https://www.geeksforgeeks.org/naive-bayes-classifiers.
[9]. Definition of precision score, f1 socre and recall score. [Online]. Available: https://machinelearningmastery.com/precision-recall-and-f-measure-for-imbalanced-classification.

### Author Information

**Nikhil Anand Mahendrakar, Sanjay A S, Manoj M**
Siemens Healthineers, Bengaluru, India

  