IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 26, No. 3, June, 2021, pp. 303-310
Journal homepage: http://iieta.org/journals/isi

Classification and Analysis of Malicious Traffic with Multi-layer Perceptron Model

Shilpa P. Khedkar[1,2*], Aroul Canessane Ramalingam[1]

[1] Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai 600119, India
[2] Department of Computer Engineering, M.E.S. College of Engineering, S. P. Pune University, Pune 411001, India

Corresponding Author Email: shilpa.khedkar@mescoepune.org

https://doi.org/10.18280/isi.260307

Received: 11 April 2021
Accepted: 13 June 2021

Keywords:
traffic classification, machine learning, deep learning, multilayer perceptron

ABSTRACT

Traffic classification is very important field of computer science as it provides network management information. Classification of traffic become complicated due to emerging technologies and applications. It is used for Quality of Service (QoS) provisioning, security and detecting intrusion in a system. In the past used of port, inspecting packet, and machine learning algorithms have been used widely, but due to the sudden changes in the traffic, their accuracy was diminished. In this paper a Multi-Layer Perceptron model with 2 hidden layers is proposed for traffic classification and target traffic classify into different categories. The experimental results indicate that proposed classifier efficiently classifies traffic and achieves 99.28% accuracy without feature engineering.

1. INTRODUCTION

Due to rapid growth in internet and network topologies network traffic increase exponentially. Traffic classification play very important role in security of networks for detecting intrusion, traffic scheduling, management of network resources and QoS. Recognition of different network classes are also possible using traffic classification. Network operators may take some actions via this technique, such as blocking certain flows and controlling resources [1]. Numerous approaches are developed over a year to satisfy the various and evolving needs of different application scenarios. Major impediments to network classification are evolving around the advances made in communications, which includes port obfuscation and encryption [2]. Different classification methods are shown in Figure 1. There are four methods available for traffic classifications [3], port-based method, inspection of packet method, Analysis of protocol method, and machine learning based method [4, 5]. Research in the area of classification of network's traffic has increased after year 2001. All types of problems in this field were caused by the dynamic changes in characteristics of traffic due to increase in network traffic and improving the backend. These network issues cannot be fully solved by the existing method of network traffic classification.

Due to the growth of big data and computer computing capabilities development, more attention of various is on deep learning. Deep learning models are intelligent and flexible. They automatically extract features during training of the model and become desirable approach for traffic classification. Key advantages of proposed method are enlisted below.

1 Time consuming task of locating and extracting important features has been eliminated.
2 Deep learning algorithm classifies network traffic in multiple classes as periodic, event, query and malicious traffic.
3 The proposed model was compared with the standard algorithm for machine learning like adaboost, Xgboost and SVM. The results indicate that the accuracy of the classification of traffic greatly enhanced.

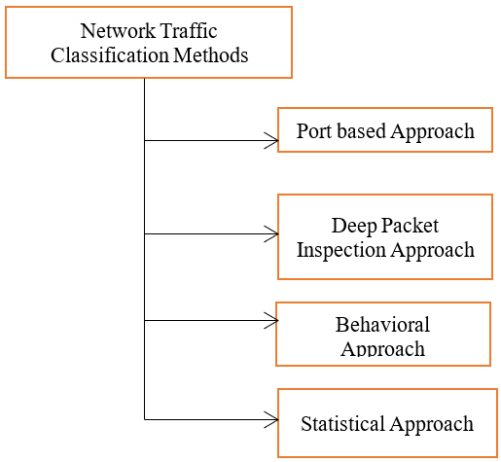

Figure 1. Methods of network traffic classification

2. RELATED WORK

The following 4 subsections present overview of different methods available in network traffic classification.

2.1 Port based method

The classification of traffic using port is primarily used in conventional applications. TCP/IP packet headers have been used to collect port information using this approach. The Internet Assigned Numbers Authority plans equivalent TCP and UDP service ports [6] for each application on the network. It is simplest and fastest method. The drawbacks of this method are (1) addition of open ports, (2) obfuscation of port, embedding of protocol and assignment of random ports

reduced accuracy of this method.

2.2 Deep Packet Inspection method (DPI)

DPI based Traffic classification provides high accuracy. IP, TCP header along with payload is inspected in deep packet inspection. Usually, payloads are used to generate signatures and then these are used to find match in traffic [7]. DPI method also has few limitations: (1) no relevant and transparent specifications [8] for new applications and protocols, (2) this method increase burden on processor which reduce computer speed.

2.3 Behavioral method

In behavioral method, pattern from traffic is observed at transport layer by receiving complete traffic at host. Different applications running on host is classify by this method. Using behavioral method CNN based classifier was developed for malicious traffic classification using image as traffic. Using same method encrypted traffic can be classify using CNN [9]. The main advantage of this method is that there is no need for packet payload access.

2.4 Statistical method

Due to development in machine learning researchers are using these techniques for traffic classification. In this method traffic data packets are captured and statistical information is calculated for the specific application traffic. Supervised and unsupervised learning comes under this method. Supervised algorithm used labeled data to train data. It produced output from previous experiences and optimized performance. Various real time problems are solved using these algorithms. In unsupervised learning algorithm works on its own to learn from experiences and used unlabeled data.

Auld et al. [10] built classifier using Bayesian Neural Networks for Classification of traffic related and achieved 99% accuracy for training and 95% for testing. Multi-layer perceptron (MLP) was proposed having input, hidden, and output layer. It allows traffic identification without using any port or device information through Bayesian framework. Xiao et al. [11] proposed artificial neural network (ANN) ensemble method. For that week traffic was captured on a backbone router and then datasets were prepared. Then ANN was trained with Error Correcting Output Codes (ECOC) method used for classification of multi class network traffic. Classification accuracy was increased by 12 to 20%.

Hwang et al. [12] used LSTM model for classification of malicious traffic using packet information. The major advantage of work is that it doesn't required processing of packet into flows which reduced preprocessing time and achieved accuracy of 97%. Yin et al. [13] and Li et al. [14] proposed model using RNN and Restricted Boltzmann Machines (RBM). Using small set of packets micro flow features are extracted and model training is done. They achieved good detection accuracy. Wang et al. [15] created classifier using encrypted traffic. They purposed different deep learning methods. The different existing techniques are discussed according to collection of datasets, input design and architecture of model, etc. In addition, some notable problems and challenges regarding the classification of mobile services traffic using Deep Learning was suggested. Lim et al. [16] presented two different deep learning models for traffic classification using CNN and ResNet. Using preprocessing technique imaged packet data were generated for eight applications. Different DL based model like SAE, LSTM, MLP, and CNN used by Aceto et al. [17] for mobile traffic classification. They achieved 76.37% for FB-FBM dataset and 85.70% accuracy for Android dataset.

Wang et al. [18] developed CNN, SAE, and MLP traffic classifiers for encrypted traffic. The experimental result achieved good accuracy. Lopez-Martin et al. [19] design network classifier using recurrent neural network (RNN), CNN and with a combination. For training of the model high level header data used instead of IP addresses and payload data. CNN was successful in classification problem as a time series data. Also, RNN gives good results with CNN combination. VPN and Non-VPN encrypted traffic used by Miller et al. [20] to build MLP based classifier and got 92% and 93% accuracy respectively. Using semi supervised learning Iliyasu and Huifang [21] created Deep Convolutional Generative Adversarial Network (DCGAN) for encrypted traffic classification. For training purpose small amount of labeled data was used. Shrikantyadav et al. [22] proposed deep learning model using Deep Autoencoder for classification. Wang et al. [9] used CNN for malware classification of traffic. Traffic is converted into 2D images and these images are used for classification. Author achieved 99.41% accuracy. Chen et al. [23] adapted CNN based model for IP traffic classification. Patterns of different applications like facebook and Instagram converted into images and then images are classified. Imbalanced classes of network problem addressed by Lythi et al. [24] using auxiliary classifiers GAN. Deployment of GAN is done to balancing major and minor labeled classes. Lyu and Lu [25] focused media traffic classification using DL models. MLP and CNN used to classify different media traffic like video, audio, text, and image. Table 1 is a summary of four traffic classification approach. Table 2 is a summary of paper reviewed in this section.

Table 1. Traffic classification approach

Approach	Features Used	Merits	Demerits	Granularity	Computational Cost
Port based	Port	Simple and Fast	Hidden port not identify	High	Lightweight
Deep Packet Inspection	Header and payload of Packet	High Accuracy	Encrypted packet cannot be handled. High computational and storage capacity.	High	High
Behavioral Method	Host Pattern Identification	Packet Payload data does not require	Results are not accurate.	Coarse	Lightweight
Statistical method	Packet	User privacy can be preserved, detect unknown application	Too much redundant features	Fine	Lightweight
	Flow			Coarse	Lightweight

Table 2. Summary of network traffic classification work

Algorithm Used for Classification	Data set Used	Key Contributions	Reference
CNN	USTC-TFC 2016	Representation learning approach used for malware detection.	Wang et al. [9]
(Generative Adversarial Network) GAN	NIMS dataset	Imbalanced problem in Network traffic analysis was addressed.	Lythi et al. [24]
RNN+CNN	266160 flows from RedIRIS dataset	For training of the model high level header data used instead of IP addresses and payload data.	Lopez et al. [19]
Deep Convolutional Generative Adversarial Network (DCGAN).	QUIC and ISCX VPN-Non-VPN datasets	Semi supervised approach used for encrypted traffic classification.	Iliyasu and Huifang [21]
MLP	Real network data capture using wire shark	MLP model trained using TCP flow-based features to classify VPN and Non-VPN traffic.	Miller et al. [20]
MLP, CNN, LSTM, SAE	FB-FBM and Android dataset	Different guidelines and challenges discuss in traffic classification using DL methods.	Aceto et al. [17]
MLP, SAE, CNN	Open dataset with 200,000 encrypted data packets from 15 applications	Use of Software define network home gateway to manage smart home network.	Wang et al. [18]
MLP, CNN	Real network traffic	Accurate Video, image, audio and text data classification done	Lyu and Lu [25]
CNN	Real network traffic	Traffic converted into image and then classification done.	Chen et al. [23]

Compared to the previous research on traffic classification there are two major differences in our work: (1) the propose model classifies normal traffic into three different classes; (2) identifies malicious traffic which can be easily block in early stage to avoid congestion in network.

3. BACKGROUND ON DEEP NEURAL NETWORK

Deep learning (DL) is a class of machine learning. Most DL models are based on Neural Networks (NNs). NNs used processing units which are highly interconnected. NNs process information based on input and give output [26]. Usually, using large number of building blocks called neurons these networks are formed. Neuron linked to each other by certain connections. Such connections are called links, and a weight value is added to each of them. During the training of the NN large number of data samples needs to feed and to get the desired performance from the NN, learning algorithm adjusts the weights (called back propagation). Deep learning frame work can be regarded as a special form of NN with several (hidden) layers. Training deep NNs has become more plausible today with the quick growth of computing resources and graphical processing units (GPUs) availability. Researchers [27, 28] also explore the use of DL systems for scientific study. Following section present Multi-Layer Perceptron used in method proposed for classification of network traffic.

Multi-Layer Perceptron:

A MLP is a feed forward neural network, which maps number of inputs to proper outputs. In a directed graph, MLP has several layers of nodes, each of which is completely linked to the next layer. If input is x vector then output y is calculated as shown in Eq. (1)

$$y = \sigma(W.x + b) \quad (1)$$

In this equation learning weight are W, b is bias neuron and σ(.) is activation function. An MLP consists of three-layer type: input, output and one or more hidden layers [29]. The architecture can be perceived as deep if there is more than one hidden layer. Input layer received input to be processed task of prediction and classification is being done by output layer. Hidden layer placed between inputs and output layer act as computational engine for MLP. In feed forward network data flows in forward direction from input to output [30]. Neurons trained using back propagation algorithm. The main uses of MLPs are in the field of classification, prediction, and recognition.

Different hyper parameters can be tuned by the multilayer perceptron, which are listed below.

1. Hidden layers: The depth of the algorithm will define by number of hidden layers and, accordingly, indicate how difficult relations can the MLP model process.
2. Neurons (Perceptron) per layer: The network width and latent space was told by the number of neurons per layer. Of course, there needs to be a number of neurons in that single hidden layer for this to hold, but it is not straightforward to know how many neurons are required.
3. Activation functions: In deep learning, activation functions (AF) are a key component. It determines the output of an input or collection of inputs provided by that node. The function is attached to each neuron to determine whether it can activate or not. It introduces non-linearity in to output

Types of Activation Functions:

1. Sigmoid Function: It is nonlinear function used in feed forward neural network.
2. Hyperbolic Tangent Function (Tanh): It is zero centered function supporting back propagation. Mostly used in NLP and speech recognitions.

3. Soft max Function: Another form of AF used in NN to compute real vector's probability distribution is the soft max function. An output varying between 0 and 1 and equal to 1 with the sum of the probabilities.
4. Soft sign Function: It is another AF used as alternative for Tanh in NN. Main use of this function in regression computation problems, it is now also used in text to speech applications based on DL.
5. Rectified Linear Unit Function (RelU): The ReLU function is most common AFs in DL models. Due to simple formula, it takes less computational expenses as compared to sigmoid and tanh.
6. Exponential Linear Units Function (ELU): The ELUs speed up NN training just as ReLU function. Removal of vanishing gradient problem is the main advantage of the ELU function.

4. ARCHITECTURE OF PROPOSED MODEL

This section presents detailing of used dataset for proposed work and DL model. Figure 2 shows high level overview of proposed framework used for the network traffic classification.

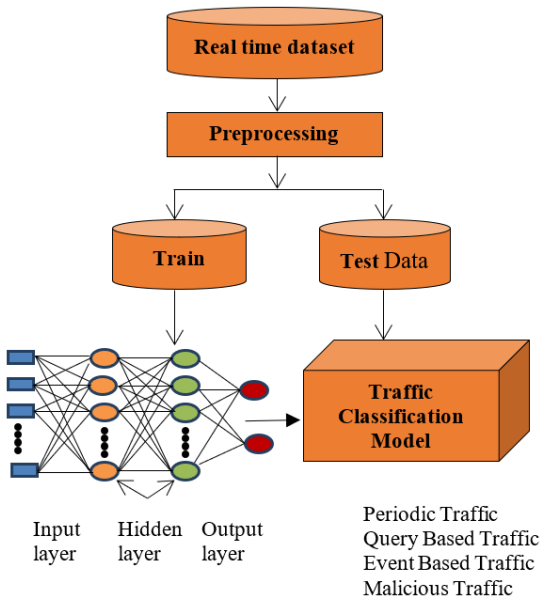

Figure 2. MLP based network traffic classification

4.1 Real time dataset

Used of real data set with 14000 records is done for training model and it is available at the repository of university of California, Irvine. Traffic Classification used one or more categories of features such as time series, header, payload data, statistical information, etc. Dataset have 18 features as shown in Table 3. The distribution of data is shown using Kernal density estimation (KDE) plot. In KDE plot, every point in the data set is represent using box, triangle, gaussian curve, etc. The plot is produced by drawing kernel (small curve) for points along an axis. The advantage of KDE is that it produces smooth estimation. The bandwidth or Kernal width controls smoothness of curve. Bandwidth should be chosen such way that it highlights all important features while maintain smoothness of curve. Plot explains whether the data has normal distribution or any kind of skewness. The KDE plots for packet received, Packet Received Rate, packet lost, packet transmitted, received bytes, transmitted bytes, and utilised bandwidth rate are shown in Figures 3 to 9. In Figure 3 KDE plot for packet received is shown. The X axis is the range of packet received in data set and the Y axis is the probability density function. Curve has many peaks indicates that distribution is not normal. From remaining figures it is observed that most of the column do not have normal distribution and mostly data having skewness. Hence, there is a need of normally distributed data while applying machine learning algorithms. Hence, there is need to apply preprocessing methods for our dataset.

Table 3. Features used for classification

S. No.	Name of feature
1	Node
2	Utilised Bandwith Rate
3	Packet Drop Rate
4	Full_Bandwidth
5	Average Delay Time Per Sec
6	Percentage of Lost Packet Rate
7	Percentage of Lost Byte Rate
8	Packet Received Rate
9	Used Bandwidth
10	Lost Bandwidth
11	Packet Transmitted
12	Packet Received.
13	Packet lost
14	Transmitted Byte
15	Received Byte
16	Node Status
17	Flood Status
18	Target: {0,1,2,3} {'Periodic':0, 'Event':1, 'Query':2, 'Malicious':3}

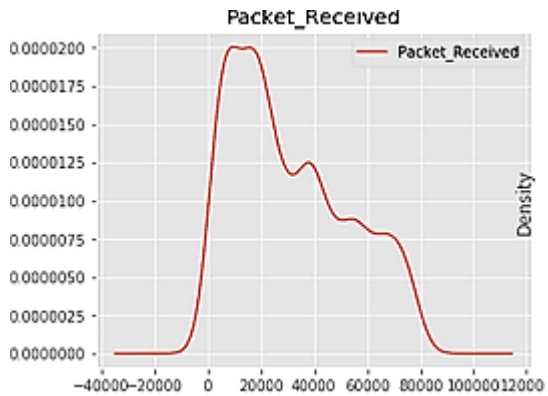

Figure 3. KDE plot for packet received

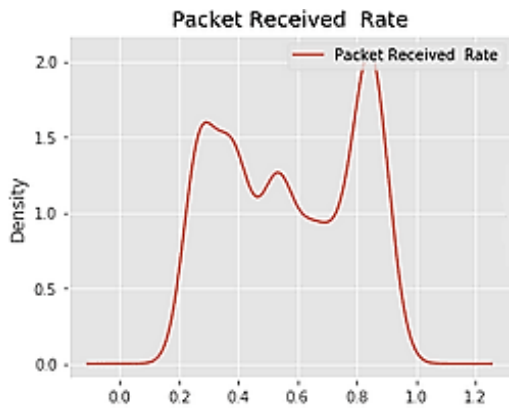

Figure 4. KDE plot for packet received rate

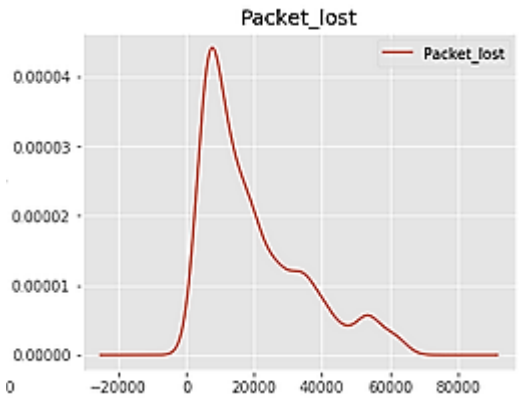

Figure 5. KDE plot for packet lost

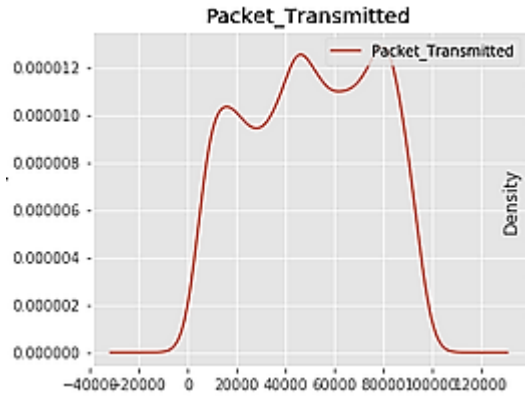

Figure 6. KDE plot for packet transmitted

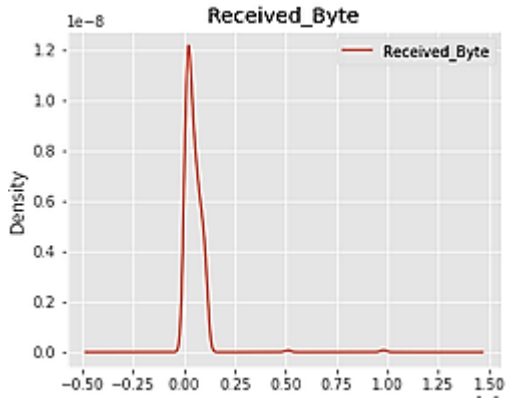

Figure 7. KDE plot for received byte

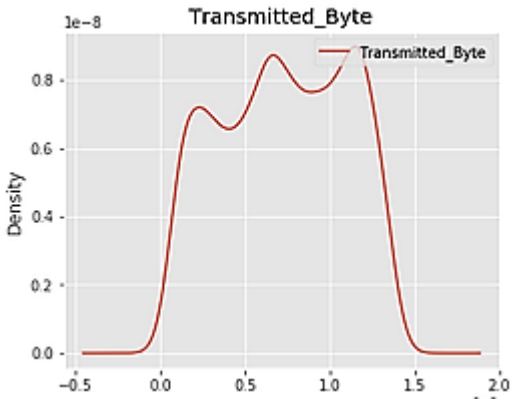

Figure 8. KDE plot for transmitted byte

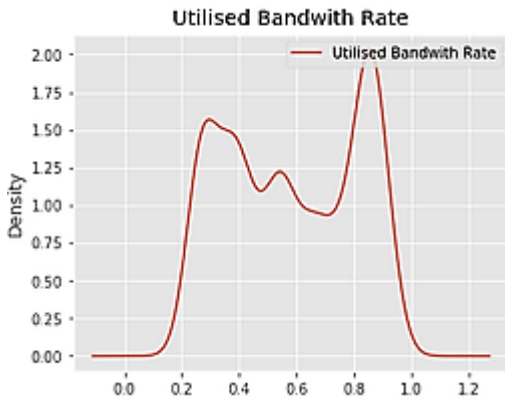

Figure 9. KDE plot for utilised bandwidth rate

4.2 Preprocessing

Preprocessing of dataset is required to clean and normalize dataset. This is considered as an important step before training classifiers using dataset. Dataset cleaning and removing of duplicate instances is done in this step so that training and processing time will be reduced during evaluation of performance. Dataset consist of some missing value instances which lead to reduced performance of traffic classification. In such case use of this invalid values classification of classifier may go wrong. To avoid it Min-Max normalization technique is used to normalize dataset. Data-set also consist of numeric values with huge range gap which need to be converted into some specific range. This is required to reduce classifier's training time. If wide range of values is used to train classifier it requires double of the time required with reduced range of values. Cleaning of the data-set is done for removing Nan, infinity and duplicate values and duplicate column. Standardization method is used for preprocessing of data. Standardization transforms data as mean equal to 0 and standard deviation as 1. The Eq. (2) used for standardization.

$$f(x) = \frac{x - \mu}{\sigma} \tag{2}$$

where, x is the actual vector of the feature, μ is mean of x, σ standard deviation of x.

4.3 Traffic classification model

We are proposing MLP model in this section. Due to high learning rate on non linear data MLP is used. Architecture is as shown in Figure 10 which used 4 layers out of this one input, two hidden and one output layer. As proposed model is multiclass classification output layer used softmax activation function.

The performance of DL algorithms is depending on initial hyper parameters such as architecture, optimizers, parameters used for regularization, etc. First layer is called as Input layer which consists of 18 neurons as we have 18 features used for classification. Data is provided through input layer. Using initial weight data is process and passes to hidden layer.

Next layer is first hidden layer with 36 neuron which is double of input layer neurons. All neurons of previous layer are connected to neurons of next layer using dot product as in Eq. (1). The next layer is second hidden layer with 72 neurons doubled neurons form previous layer neurons. Activation functions are applied to generate new outputs. Relu activation

function is used for input and hidden layers of model as shown in Eq. (3).

$$F(x) = \max(0, x) \tag{3}$$

where, x is the input to Relu.

The last layer is output layer with 4 neuron due to 4 classes. Softmax layer added after output layer as the loss function. This loss is a strong indicator of how easily two discrete probability distributions can be distinguished from one another.

While training, the epochs are the number of times the whole training data is presented to the network. Epoch is considered as 100. The count of sub samples set to the network after which parameter updates occur is known as the batch size which is taken as 10. Choice of optimizer is very important for DL algorithms. Adam optimizer is used instead of classical stochastic gradient descent to update weights iteratively. Updated weights are propagated through Adam optimizer by using back propagation algorithm. This optimizer compute individual learning rate for different parameters.

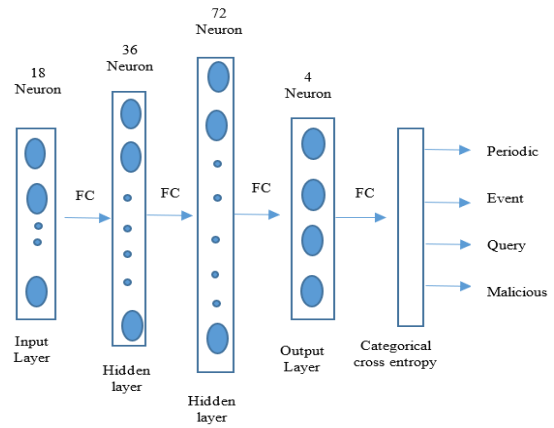

Figure 10. Proposed multilayer perceptron model (FC is fully connected)

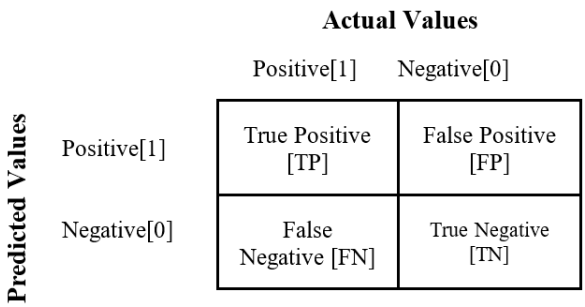

Figure 11. Confusion matrix

The evaluation matrices used for MLP were Confusion Matrix, precision, recall, F1 Score and Receiver Operating Curve (ROC). Confusion matrix used to describe the instances of predictions generated by the classifier. The confusion matrix represented in the form of instances as shown in Figure 11 which is combination of actual and predicted values.TP, TN, FP, FN terms of confusion matrix used to check accuracy of MLP classifier. Where TP Predicted as positive and it is true, FN Predicted as negative and it is false, TN Predicted as negative and it is true, FP Predicted as positive and it is false. The precision is the proportion of the relevant results. It is used to show how many of the positive classes that we correctly predicted are actually positive. Ideally precision should be 1 for good classifier. Recall is referred to as sensitivity or true positive rate. F1 Score is a measure that takes both precision and recall into account. Precision, recall and F1 score stated as Eq. (4), Eq. (5), Eq. (6).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{6}$$

ROC curve is a method for comparing several classifier models and visualizing the results in the form of curves. This tool is primarily used to examine radar images. As a result, it's a useful tool to visualize a classifier's output and deciding on an appropriate operating point, or decision threshold. When comparing a variety of different classification systems, however, it's always preferable to have a single figure to use as a measure of the classifier's results.

5. RESULT AND DISCUSSIONS

The MLP classifier is implemented using keras and tensorflow then it was evaluated on dataset available. For training of the model 67% of the dataset used and remaining 33% for testing. Model is executed on google colab and achieved training accuracy is 99.1 and testing accuracy is 99.28. Figure 12 is confusion matrix for MLP classifier.

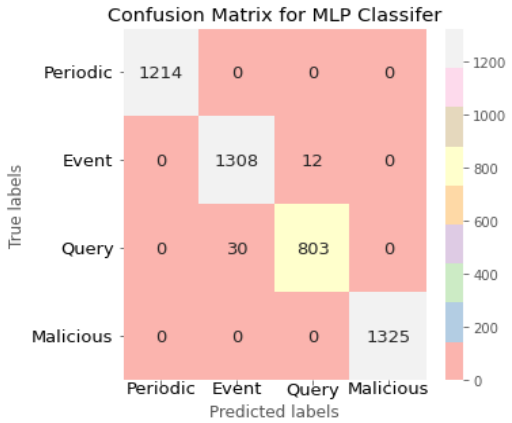

Figure 12. Confusion matrix for MLP

Comparison based-on training and testing Accuracy of MLP classifier shown in Figure 13 with other algorithms. Comparison of MLP model with SVM, adaboost and Xgboost based on F1 score, precision score and recall score shown in Figure 14. From Figure 13 it is observed that MLP classifier has good training and testing accuracy compared to others classifiers. From Figure 14 it is observed that MLP has F1 Score - 0.99, Precision Score - 0.99, Recall Score - 0.99. Adaboost has F1 Score - 0.93, Precision Score - 0.96, Recall Score - 0.91. XGboost classifier has F1 Score - 0.9, Precision Score - 0.89, Recall Score - 0.92. For SVM F1 Score - 0.89, Precision Score - 0.81, Recall Score - 0.99. Figure 15 shows a

ROC for MLP classifier which is closer to left side border so indicate good accuracy.

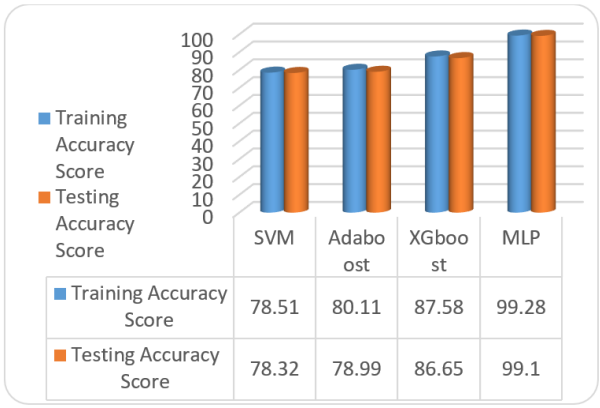

Figure 13. Comparison of training and testing accuracy

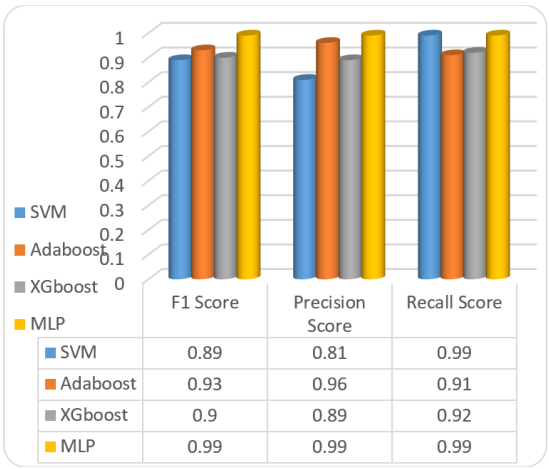

Figure 14. Comparison of F1 score, precision and recall score

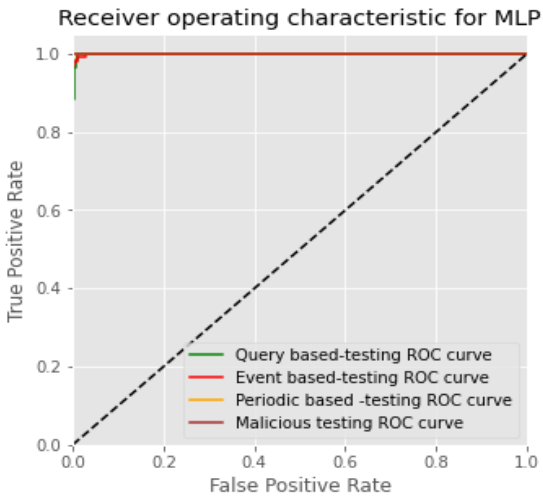

Figure 15. ROC curve for MLP classifier

6. CONCLUSION

Due to advancement in technology traffic transmission speed is improved. But malicious traffic degrades the performance by consuming bandwidth and other resources. In this paper, deep learning method is presented to classify network traffic in different classes. Due to this segregation malicious is traffic block in early stage to improve throughput of network. The performance of proposed model is evaluated using accuracy, F1 score, precision score and recall score. MLP classifier was trained using dataset created using 21 features. Preprocessing was done to remove noise and to clean data. Feature selection was done automatically using deep learning algorithm. Classifier is compared with SVM, adaboost, and Xgboost. Result show that proposed approach improved accuracy by 13%. In future work traffic prediction will be done to detect congestion in network along with allocating proper channels.

REFERENCES

[1] Shafiq, M., Yu, X., Laghari, A.A., Yao, L., Karn, N.K., Abdessamia, F. (2016). Network traffic classification techniques and comparative analysis using machine learning algorithms. In 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, pp. 2451-2455. https://doi.org/10.1109/CompComm.2016.7925139

[2] Shahbaz, R., Xin, L. (2019). Deep learning for encrypted traffic classification: An overview. IEEE Communications Magazine, 57(5): 76-81. https://doi.org/10.1109/MCOM.2019.1800819

[3] Neeraj, N., Shikha, A., Sanjay, S. (2015). Recent advancement in machine learning based internet traffic classification. Procedia Computer Science, 60: 784-791. https://doi.org/10.1016/j.procs.2015.08.238

[4] Tongaonkar, A., Torres, R., Iliofotou, M., Keralapura, R., Nucci, A. (2015). Towards self adaptive network traffic classification. Computer Communications, 56: 35-46. https://doi.org/10.1016/j.comcom.2014.03.026

[5] Wang, Y., Xiang, Y., Zhou, W., Yu, S. (2012). Generating regular expression signatures for network traffic classification in trusted network management. Journal of Network and Computer Application, 35(3): 992-1000. https://doi.org/10.1016/j.jnca.2011.03.017

[6] Shi, H., Li, H., Zhang, D., Cheng, C., Wu, W. (2017). Efficient and robust feature extraction and selection for traffic classification. Computer Networks, 119: 1-16. https://doi.org/10.1016/j.comnet.2017.03.011

[7] Hubballi, N., Swarnkar, M. (2018). $BitCoding$: Network traffic classification through encoded bit level signatures. IEEE/ACM Transactions on Networking, 26(5): 2334-2346. https://doi.org/10.1109/TNET.2018.2868816

[8] Zhang, J., Chen, C., Xiang, Y., Zhou, W., Vasilakos, A.V. (2013). An effective network traffic classification method with unknown flow detection. IEEE Transactions on Network and Service Management, 10(2): 133-147. https://doi.org/10.1109/TNSM.2013.022713.120250

[9] Wang, W., Zhu, M., Zeng, X., Ye, X., Sheng, Y. (2017). Malware traffic classification using convolutional neural network for representation learning. International Conference on Information Networking (ICOIN), Da Nang, Vietnam, pp. 712-717. https://doi.org/10.1109/ICOIN.2017.7899588

[10] Auld, T., Moore, A.W., Gull, S.F. (2007). Bayesian neural networks for internet traffic classification. IEEE

Transactions on Neural Networks, 18(1): 223-239. https://doi.org/10.1109/TNN.2006.883010
[11] Xiao, X., Yang, B., Chen, Y., Wang, L., Chen, Z. (2009). Network traffic classification based on error-correcting output codes and NN ensemble. 6th International Conference Fuzzy System Knowledge Discovery, Tianjin, China, pp. 475-479. https://doi.org/10.1109/FSKD.2009.694
[12] Hwang, R.H., Peng, M.C., Nguyen, V.L., Chang, Y.L. (2019). An LSTM-based deep learning approach for classifying malicious traffic at the packet level. Applied Sciences, 9(16): 3414. https://doi.org/10.3390/app9163414
[13] Yin, C., Zhu, Y., Fei, J., He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access, 5: 21954-21961. https://doi.org/10.1109/ACCESS.2017.2762418
[14] Li, C., Wang, J., Ye, X. (2018). Using a recurrent neural network and restricted Boltzmann machines for malicious traffic detection. NeuroQuantology, 16(5): 21954-21961. https://doi.org/10.14704/nq.2018.16.5.1391
[15] Wang, P., Chen, X., Ye, F., Sun, Z. (2019). A survey of techniques for mobile service encrypted traffic classification using deep learning. IEEE Access, 7: 54024-54033. https://doi.org/10.1109/ACCESS.2019.2912896
[16] Lim, H.K., Kim, J.B., Heo, J.S., Kim, K., Hong, Y.G., Han, Y.H. (2019). Packet-based network traffic classification using deep learning. In 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Okinawa, Japan, pp. 46-51. https://doi.org/10.1109/ICAIIC.2019.8669045
[17] Aceto, G., Ciuonzo, D., Montieri, A., Pescapé, A. (2019). Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges. IEEE Transactions on Network and Service Management, 16(2): 445-458. https://doi.org/10.1109/TNSM.2019.2899085
[18] Wang, P., Feng, Y., Xuejiao, C., Yi, Q. (2018). Datanet: Deep learning based encrypted network traffic classification in SDN home gateway. IEEE Access, 6: 55380-55391. https://doi.org/10.1109/ACCESS.2018.2872430
[19] Lopez-Martin, M., Carro, B., Sanchez-Esguevillas, A., Lloret, J. (2017). Network traffic classifier with convolutional and recurrent neural networks for Internet of Things. IEEE Access, 5: 18042-18050. https://doi.org/10.1109/ACCESS.2017.2747560
[20] Miller, S., Curran, K., Lunney, T. (2018). Multilayer perceptron neural network for detection of encrypted VPN network traffic. In 2018 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA), pp. 1-8. https://doi.org/10.1109/CyberSA.2018.8551395
[21] Iliyasu, A., Huifang, D. (2019). Semi-supervised encrypted traffic classification with deep convolutional generative adversarial networks. IEEE Access, 8: 118-126. https://doi.org/10.1109/ACCESS.2019.2962106
[22] Srikanthyadav, M., Gayatri, K., Padmaja, R. (2020). A Deep learning approach to network intrusion detection using deep autoencoder. Revue d'Intelligence Artificielle, 34(4): 457-463. https://doi.org/10.18280/ria.340410
[23] Chen, Z., He, K., Li, J., Geng, Y. (2017). Seq2img: A sequence-to-image based approach towards IP traffic classification using convolutional neural networks. In 2017 IEEE International Conference on Big Data (Big Data), pp. 1271-1276. https://doi.org/10.1109/BigData.2017.8258054
[24] Lythi, V., Cong, B., Quang, U. (2017). A deep learning based method for handling imbalanced problem in network traffic classification. Proceedings of the Eighth International Symposium on Information and Communication Technology, pp. 333-339. https://doi.org/10.1145/3155133.3155175
[25] Lyu, Q., Lu, X.J. (2019). Effective media traffic classification using deep learning.3rd international conference on compute and data analysis. Kahului, Hawaii, USA, pp. 139-146. https://doi.org/10.1145/3314545.3316278
[26] Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.R., Jaitly, N. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Processing Magazine, 29(6): 82-97. https://doi.org/10.1109/MSP.2012.2205597
[27] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
[28] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 1631-1642.
[29] Collobert, R., Bengio, S. (2004). Links between perceptrons, MLPs and SVMs. In Proceedings of the Twenty-First International Conference on Machine Learning, p. 23.
[30] Goodfellow, I., Bengio, Y. Courville, A. (2016). Deep Learning. MIT Press, https://doi.org/10.4258/hir.2016.22.4.351