



# A REVIEW PAPER IN MACHINE LEARNING CATEGORIZATION OF NETWORK CONGESTION

<sup>1</sup>ER.RIYA SHARMA, <sup>2</sup>MS.RANVIR KAUR

<sup>1</sup>ASSISTANT PROFESSOR, <sup>2</sup>ASSISTANT PROFESSOR  
RAYAT BAHRA UNIVERSITY, MOHALI

*Abstract :- Network means connecting computers and servers across the world using routers, switches and telephone lines, other communication devices and channels. Thousands of people are using the same network simultaneously because of that a large number of data packets are gathered in a network that cause network congestion. Some Packets on the network are malicious while others are non-malicious. Malignant data packets are sent by an attacker to cause harm to the network or to hack the network. In this research SVM Classifier and KNN Classifier is used for categorization of data packets on networks.*

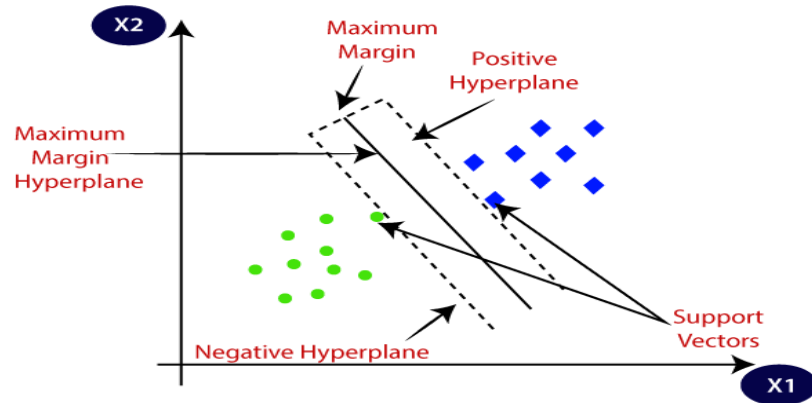
**INTRODUCTION:** When two or more devices are linked together with the purpose of attaining a discrete benefit, a network is established. Once a connection is established, protocols — PSTN and VoIP, IP Protocol are used for communication.

- PSTN (Public Switched Telephone Network):-** This system came into existence in the late 1800s. In PSTN System we made communication through the copper wires that are placed underground. It is one of the reliable means to communicate with anyone around the world for generations' PSTN is a combination of switches at common points on a network that enables the communication of two points.
- Voice over Internet Protocol (VoIP):** The connection of VoIP is much smoother. VoIP is a technology that allows user to make voice calls over a broadband Internet connection in inspite of phone line. We can make a call to telephone numbers that are National ,local, long-distances. With the advancement in VoIP now video call is also possible over the network.
- Internet protocol(IP):-**Initially we have IPV4 version now IPV6 is developed. Internet protocol is a type of protocol that is used for sending the packets from one place to other place i.e destination. The Major task of IP is to deliver the packets based on the IP address from source to destination.

## Methods used for compartmentalization of Network Traffic

- **Support Vector Machine (SVM):** SVM is one of the most common Supervised Learning algorithms. It is basically used for Classing as well as Regression problems. The Use of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can predict the class of new data and put it into the category to which it belongs. This good decision boundary is called a hyperplane.

The Following diagram represents two different categories that are classified using a hyperplane or decision boundary.



We have two types of data packets in the network malevolent and non-malevolent. The SVM Classifier has data packets into a hyperplane and it puts the packets into the category to which they belong. Class of the upcoming data packets is determined by seeing the behavior of that packet.

#### ● K-NEAREST NEIGHBOR (KNN) ALGORITHM FOR MACHINE LEARNING

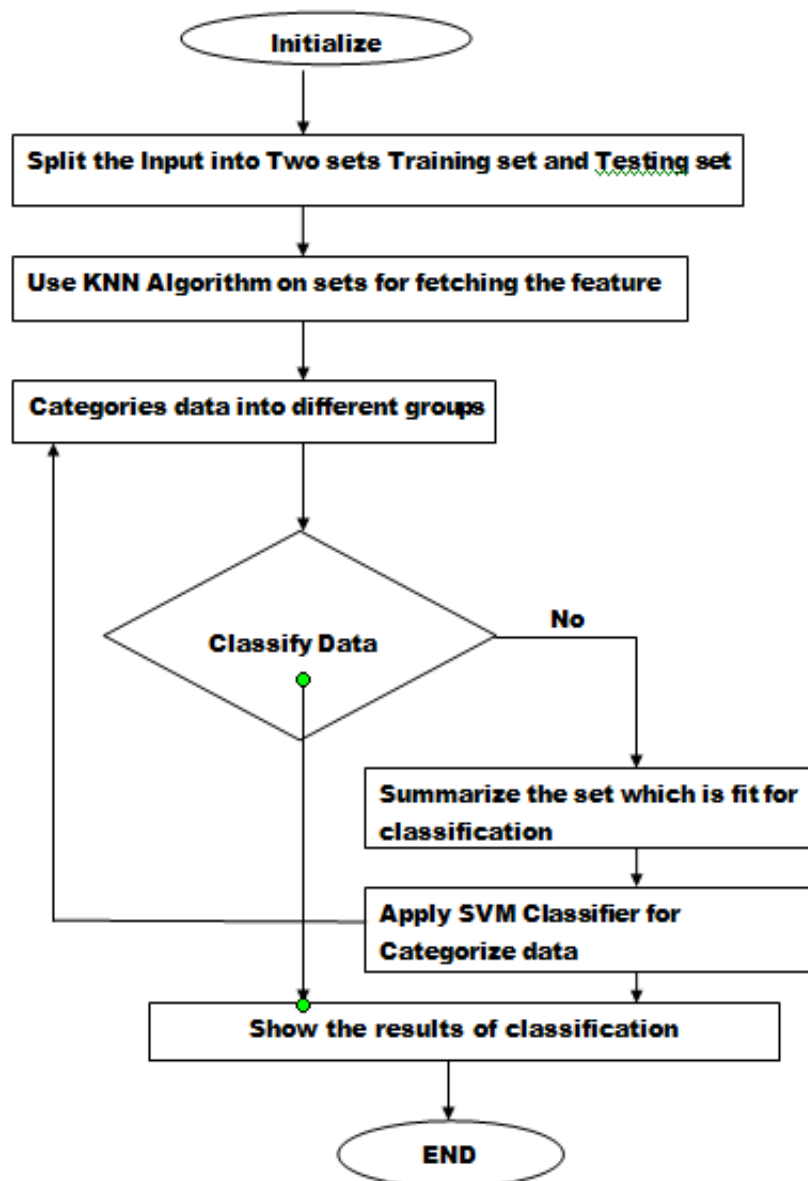
- K-Nearest Neighbor is one of the simple and Supervised Learning technique. It is based on Machine Learning.
- It assumes the similarity between the new case/data and available cases and lay the new case into the group that is most similar to the available pigeonhole.
- K-NN is a type of non-parametric algorithm, which means it does not make any guess on underlying data.
- This algorithm is also called a lazy learner algorithm. This algorithm does not learn from the training set . It is a type of algorithm that stores the dataset and at the time of categorization , it performs an action on the dataset.

#### Steps followed in Research

- ➔ First of all we use the K-mean Clustering technique to cluster the same data at one side and dissimilar data on another side.
- ➔ Then to purify the data set in the form of input we then eliminate redundant and misplaced data.
- ➔ After that we calculate central point of network and Euclidean distance then we divide the similar data packets in one cluster and dissimilar in another.
- ➔ The last step is to use SVM Classifier because it improves efficiency and execution time.

#### Proposed Flowchart

International Research Journal  
**IJNRD**  
 Research Through Innovation



## CONCLUSION

Categorization of data is a vital step in machine learning. These steps also gives support in the prediction of unlabeled occurrence. Huge quantities of data categorizing models are made because of the presence of a large number of applications. Many Algorithms are deployed in this work. Computations such as numbers of neurons, hidden layers of ANN, value of k in KNN classification will be performed.

## REFERENCES

- [1] Necati DEMIR, Gokhan DALKILIC, “Modified stacking ensemble approach to detect etwork intrusion”, 2018, Turkish Journal of Electrical Engineering & Computer Sciences, 26:418-433.
- [2] M. Mazhar, U. Rathore, “Threshold-based generic scheme for encrypted and tunneled Voice Flows Detection over IP Networks” in 2015.
- [3] Muhammad Shafiq,, “Network Traffic Classification Techniques and Comparative Analysis UsingMachine Learning Algorithms”, 2016 2nd IEEE International Conference on Computer andCommunications.
- [4] JaiswalRupeshChandrakant, LokhandeShashikant. D., “Machine Learning Based Internet Traffic Recognition with Statistical Approach”, Annual IEEE India Conference,2013.
- [5] RiyadAlshammari, A. NurZincir-Heywood, “Identification of KDD encrypted traffic using a machine learning approach”, in 2015.