EVALUATING SHALLOW AND DEEP NEURAL NETWORKS FOR INTRUSION DETECTION SYSTEMS CYBER SECURITY

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY

IN

INFORMATION SYSTEMS

Submitted By:

RAJ KISHORE

(2K18/ISY/09)

Under the supervision of
Ms. ANAMIKA CHAUHAN

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

JULY, 2020

CANDIDATE'S DECLARATION

I, RAJ KISHORE, Roll No. 2K18/ISY/09 student of M.Tech Information Systems, hereby declare that the project Dissertation titled "Evaluating shallow and deep neural networks for network intrusion detection" which is submitted by me to the Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any degree, Diploma Associate ship, Fellowship or other similar title or recognition.

Raj Kishore

Place: Delhi Raj Kishore

Date: 26/08/2020

CERTIFICATE

I hereby certify that the Project Dissertation titled "Evaluating shallow and deep neural networks for Intrusion detection" which is submitted by Raj Kishore , Roll No 2K18/ISY/09 Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Anamika Chauh

Place: Delhi **Ms Anamika Chauhan**

Date: **SUPERVISOR**

ACKNOWLEDGEMENT

I express my gratitude to my major project guide Ms. Anamika Chauhan, Assistant Professor, IT Dept., Delhi Technological University, for the valuable support and guidance she provided in making this major project. It is my pleasure to record my sincere thanks to my respected guide for her constructive criticism and insight without which the project would not have been shaped as it has.

I humbly extend my words of gratitude to other faculty members of this department for providing their valuable help and time whenever it was required.

Raj Kishor

Raj Kishore

Roll No. 2K18/ISY/09

M.Tech (Information Systems)

E-mail: raj.verma5454@gmail.com@gmail.com

ABSTRACT

This project is concerned with intrusion detection systems and several techniques. As we know today's era is of computer networks or of internet of things which can lead to intrusion and can be devastated to our system, so intrusion detection system can help computer administrators to curb such activities and prevent our systems. As the systems are going towards advancement and in day to day life so as the risk of attack is also going to increase.

In this project I am going to discuss about the deep neural networks which is used to calculate the accuracy of the intrusion detection with the learning rate of 0.1 and iteration is 1000. The dataset is used KDD CUP 99 which is a standard set of database which includes large variety of intrusion stimulation in a military network area.

We can shield our ICT (Information and communication technology) systems with anomaly detection systems also but they are not that much efficient. They have some fault/foible or we can say demerit such as we might get difficulties/complexity for defining rules of network detection. We have to define each protocol, analyze and implement it and test for the accuracy. Some of the harmful activities that may cause our system might fall in usual usage range which will lead to not recognized through anomaly based that's why we used IDS which can train and adapt itself after recent novel attacks and becomes indispensable.

So in this project I compared results of several classical machine learning techniques like Adaboost , decision tree , KNN , linear regression , Random forest , SVM linear , SVM rbf also used. Deep neural networks with three layers after 100 iteration gives the better results as compared to classical machine learning techniques with the higher accuracy and better results.

CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

INTRUSION DETECTION SYSTEM

1.1 Introduction

An Intrusion Detection System is software to monitor and protect our network from any kind of intrusions. With the rapid growth in internet nowadays lead to more prone to the cyber intrusions [1]. The area like financial, business, Industries, Health sector, security, WAN , LAN application have progressed. These sites have made the internet an favorable site to get abused which lead to the threat of losing the personal information and data which is very big Susceptibility for the network community [1]. Malevolent users or we can say hackers' hacks the organizations private data to collect information and causes problems in the system or Vulnerabilities like Software failure, temporary failure of administration like set the systems on default settings which can be easily prune [2].

As internet is going into the society, things like malicious software like Worms , Viruses , Phishing are imported into the systems which will be used to crack the passwords , Reading unencrypted text and inject malicious code with some encryption which will encrypt all the data files and lead to data loss which cause vulnerability to the system. Hence some sort of security is needed for us to secure our systems from intruders.

There are several ways from which cyber attacks can be performed for example Man-in-the-middle attack, phishing, Denial of Service (DoS) , Distributed denial of service (DDoS) , SQL injection ,as discussed in [3] . As discussed there are a lot of different techniques from which an attacker can attack into our network to steal our personal information, so in order to prevent the attack different technologies are being used like most common Firewalls which is popular to protect virtual private network form intrusion. IDS are very helpful in the field of banking, health sector , insurance , share market etc. and access control.

According to [4] , In spite of the fact that there are so many different technologies to prevent the intrusion in our system , but still the network is vulnerable to so many new attacks which are hard to detect. So in order to prevent such new malicious attacks IDS are created and still they are being developed with greater rate by new and different technologies.

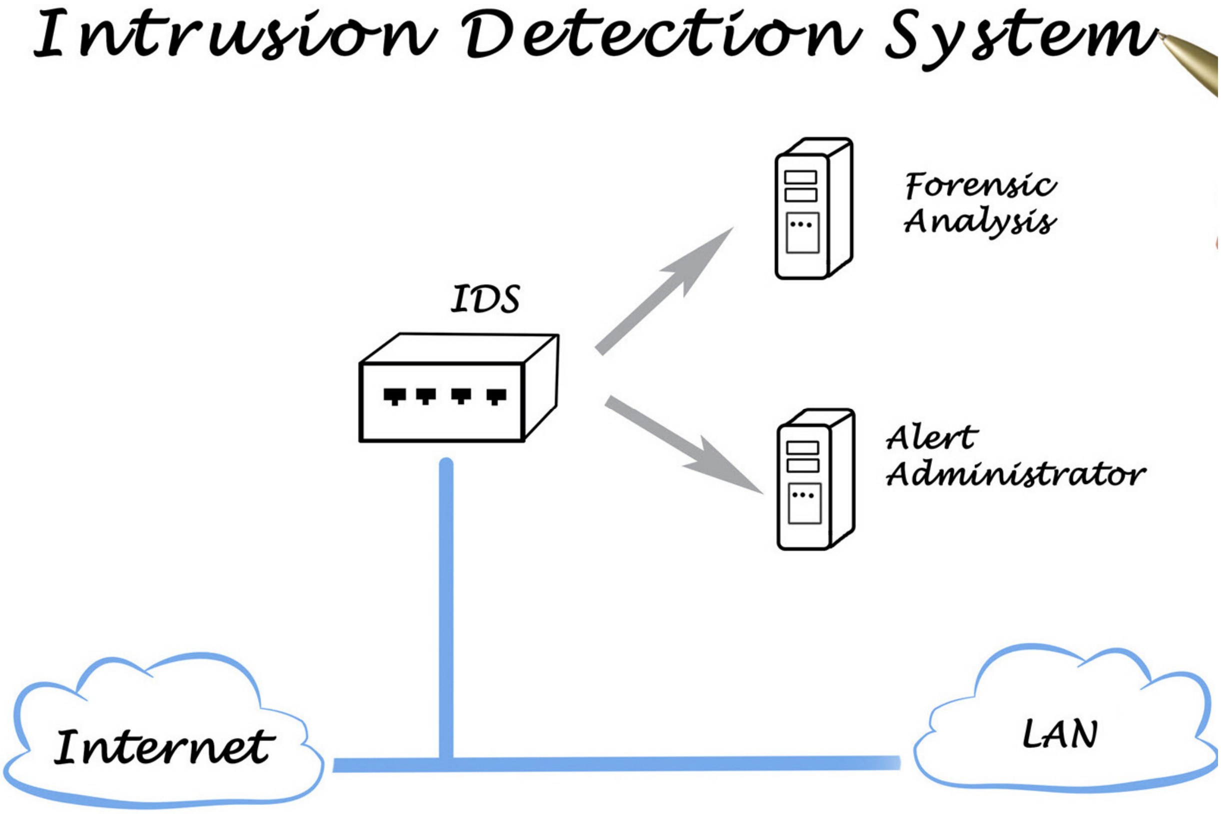


Fig 1.1 Intrusion Detection System

An IDS is known as burglar alarm. Like there is lock system in our house to protect it from theft. If somebody tries to enter into the house after breaking the lock system, then this burglar alarm detects that the locking system has been broken and raise the alarm to alert the owner. However Firewall is doing great job, it filters the incoming traffic from the network to bypass the firewall [4]. Like outside users can connect to internet by installed modem and dialing through modem which is installed in private network (VPN) of the company, this access control cannot be detected by firewall.

An (IPS) Intrusion Prevention system is used to guard network and ensures threat/security prevention technology which is used to audit and monitor network traffic flows to detect/notice and prevent vulnerability utilize. There are two types of Network Prevention system Host (HIPS) , Network (NIPS). These network systems are smart as they automatically watch the network traffics find out the necessary actions to be taken to protect system and networks. Issue with IPS is false negatives and positives. False positive defined as when there is not attack and still produces an alarm. False negative is define as when there is an attack and still no alarm to alert the owner.

1.2 Types of IDS:

- Host Based IDS
- Network Based IDS
- Application Based IDS

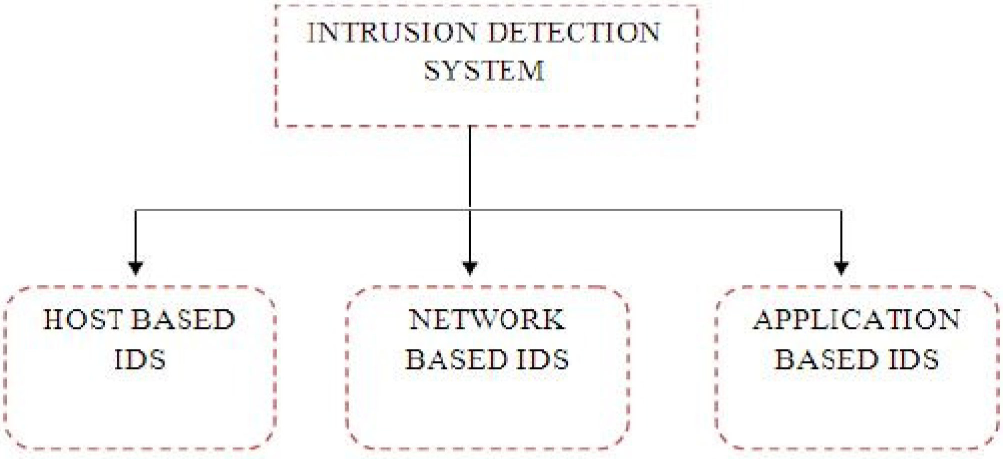


Fig 1.2 Intrusion Detection System –Types

Host based intrusion detection systems (IDS) monitors the sign of malpractices in the local system. For the process of analyzing they use the other information and system's logging. Handler of Host based is basically referred as sensor [12] [13] [14]. Host based sensor collects the data from other sources which includes the log information of the system , log files created by the operating system (OS) processes which also contains content of objects which is not shown logging mechanism of operating system [5]. Host Based IDS blindly trust on the monitored trail. The data allows the IDS (intrusion detection system) to find out the several pattern of malpractices which would not be available at the at high level of abstraction [6] . The basic in IDS including NIDS starts from anomaly HIDS research which is based on Paramount work of Denning's [7]. Host based IDS compared to Network based IDS provides far better results. HIDS sometimes provide amazed results about intrusion like If there is attack then it can tell what type of attack was that, what type of commands were used by the attacker , which files they were accessing . It provides accurate results rather than false accusation and tells whether there was dangerous command which was executed [8]. It is less vulnerable or risky.

ADVANTAGES

- Verifies if the attack was successful or not.
- Always monitors system activity.
- Can detect attacks which are not detected by Network based IDS.
- It does not require any sort of external hardware.
- Near to real time protection.
- Lower entry cost.

Network based IDS systems gathers data from the internet by self rather than collecting data from individual host [9]. The NIDS go through the network malpractice while flows of packets are there in the network. The sensors of the network come with equipped signature of attacks that are rules through which an attack can be done. Most of the NIDS (network based intrusion detection system) allow the owner to define their own signature [9]. Malpractice on the network is based on sensors signature which are from their previous attacks and the process of being audited will also be transparent to the owner and this is very significant [10].

The clarity of the audit decreases the possibility of counseling that can locate it in future and void its potential without the efforts [6]. Network agents of the node are there at every host in the network coverage range which is being protected [11].

ADVANTAGES

- Low cost to own.
- Easy to Deploy.
- Retaining Evidence.

- Detection of failed attacks.
- Detect network based attack.
- Detect real time and provide quick response.

Application Based intrusion detection System (APIDS) will monitor the successful actions and happening of the protocol [11]. The agent or the monitoring system is placed between sack of server and process that audit and analyzes the protocol between devices and the application [11]. Deliberate attacks are malevolent attacks which are done by aggrieved hired hand to cause distress to the organization and unintentional attacks causes' financial distress to the company by vandalizing the most important files of the organization [11]. There are many attacks which can be takes place in the OSI layer.

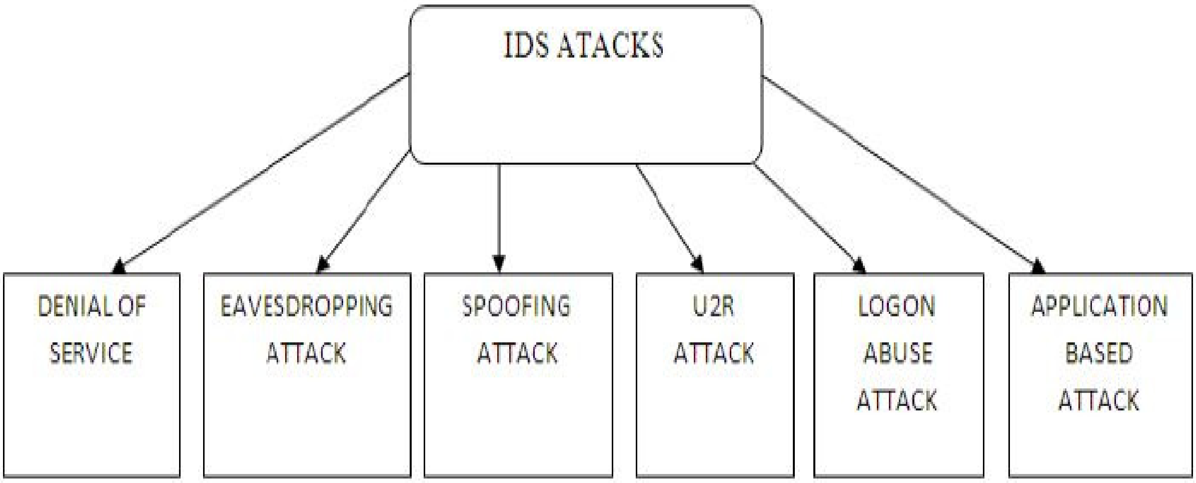


Fig 1.3 Types of attacks.

- **DENIAL OF SERVICE (DOS ATTACK)**

A denial of service attack is a malpractice attempt which makes the server and network unavailable to the users. In this attack hacker sends plenty of requests to the server and makes it chock. For every server there is a limit to reply no of queries made by it, if the queries exceeds the limit server gets crashed [11]. Denial of service attack can we prevented by seeing plenty of queries are coming from the same IP address (Internet Protocol) and the owner can block that particular IP address. So to overcome this hackers use DDOS

(Distributed denial of service attack). In this attack hacker attacks the system after giving requests from many individual systems. If request is coming from different sources, it is difficult for the user to identify the DDOS attacks, as there may be different users giving request. This technique can easily lead to the server failure and make it chock for legitimate users. DDOS is hard to detect by the users and very common attack in hacking world.

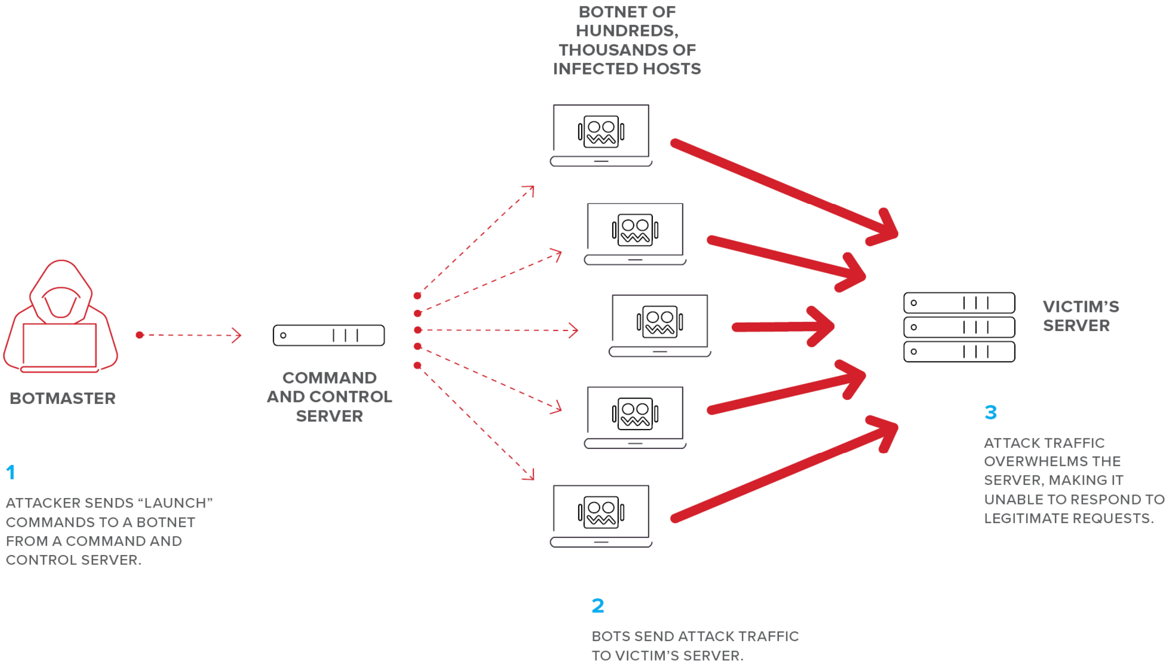

BOTNET OF HUNDREDS, THOUSANDS OF INFECTED HOSTS
BOTMASTER
COMMAND AND CONTROL SERVER
VICTIM'S SERVER
1
ATTACKER SENDS "LAUNCH" COMMANDS TO A BOTNET FROM A COMMAND AND CONTROL SERVER.
2
BOTS SEND ATTACK TRAFFIC TO VICTIM'S SERVER.
3
ATTACK TRAFFIC OVERWHELMS THE SERVER, MAKING IT UNABLE TO RESPOND TO LEGITIMATE REQUESTS.

Fig 1.4 Denial Of Service

a. SYNC ATTACKS

As cleared by its name SYNC attack means synchronous attack in which flood of request is being send from the hackers side to server , which is beyond the capacity of the server to handle all the request simultaneously and lead to system crash.

b. PING OF DEATH

In this the hacker sends targeted system a ping request which is way bigger than 65,536 bytes which leads to system crash. The normal size must be 56 or 84 bytes if we consider internet protocol header [11]

- **EAVESDROPPING ATTACKS**

In this attack hacker mainly affect the communication process. This attack mainly done on telephone lines, electronic mails and several others communication channel. When two users are connected through some channel for communication, it may be internet, telephone line etc. then hacker manipulates the communication channel , one user is sending the data to legitimate user in the encrypted form and gives the user a private key to decrypt the file. Hacker uses that key to encrypt the data , and send the manipulated data to the user [11].

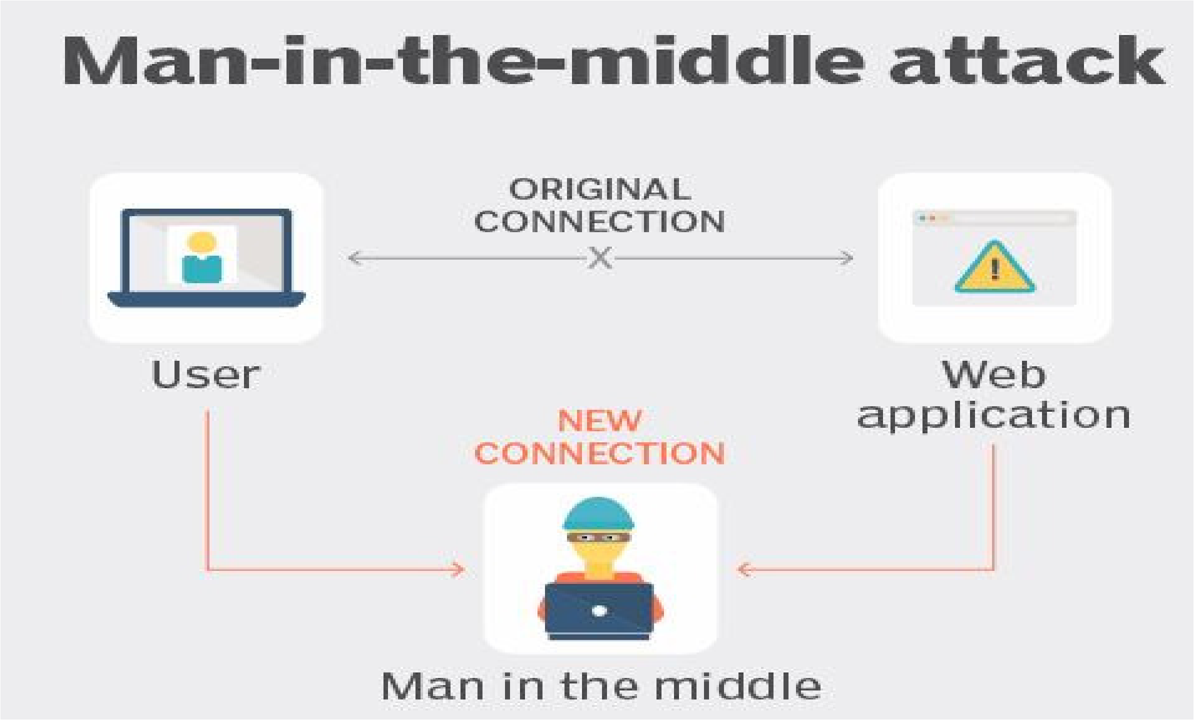


Fig 1.5 Man in Middle attack.

- **SPOOFING ATTACKS**

In this attack hacker portrays itself as another legitimate user to steal , vandalize the data and sometimes takes advantage of the illegal events on the network protocol [11] [12]. It gain access to the system after changing its own IP address to the legitimate user . Systems get easily fooled by the trick , it always think the communication which has been established is a legitimate user but in reality it's the attacker. System despite of real fact gives the access to the attacker into the system.

- **USER TO ROOT ATTACKS (U2R)**

In this attack attacks tries to gain access into the system through internet after data leak in the system. This data leak occurs when server receives more data than its capacity and not being programmed to handle that much amount of data, in this condition buffer overflow happens. Buffer overflow also happens if the program to handle the server is not good, like buffer overflow can occurs in the program of the system. Like string overflow and integer overflow if we try to store more data into the memory than also there is a leak in the memory which can be used by the hacker to gain access into the system through that memory leak, they try to inject malicious code into that leaked memory. The malicious code compiled itself and makes .exe file which automatically get executed and can provide the access to its owner it works like someone is serving for its owner [11] [12].

- **LOGON ABUSE ATTACK**

It is also called brute force attack. In this type of attack attacker does not care about the authentication, they directly access the control management of the system than the system grant an attacker with more advantages [11].

- **APPLICATION LEVEL ATTACKS.**

In this type of attack hacker targets the weaker part of the application layer. For example there is weakness in the security of the server side or there are certain fault controls towards the server end [11].

1.3 FUNCTIONS OF THE INTRUSION DETECTION SYSTEMS (IDS)

The IDS mainly have four major key functions which are followed below.

- Data collection.
- Feature selection.
- Analysis.
- Action.

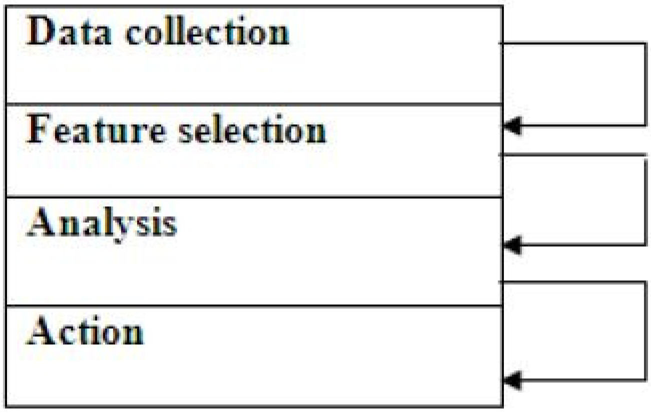


Fig 1.6 Functionality Of IDS

- DATA COLLECTION

This functionality of IDS passes the data to IDS as input. This mechanism is very useful factor for the IDS which can also affect the accuracy of the intrusion detection system (IDS). In data collection we don’t apply hard sampling technique of statistics which can lead to bulkier system with more time to react, we simply use simple random sampling technique for the procedure of data collection which can provide new data collection data frame for IDS [16]. IDS records the data into a single file then analyze it. As network based IDS (NIDS) gather and adulterate the data packets where as host based IDS (HIDS) gathers the information of physical system like process of the system, usage of the disk. As proved the by experimental results data model can progress in the effectiveness of the data collection and improve the initial processing of IDS.

- FEATURE SELECTION

This is a preprocessing feature of IDS which can lead to solve IDS queries smartly by selecting relevant data and removing redundancy with the irrelevant features [15]. To select the feature data which is available in the network , then this data will be use to evaluate any intrusion in the system. Like there are several ways which can be used as key for intrusion follow as IP address of targeted system, IP address of source, length of head, protocol type. Relevant data may have useful information of the classes , which can be essential for correct procedure of the classifier.

- ANALYSIS

In this functionality of IDS, data which has been selected after the data selection and feature selection thoroughly analyzed to find whether it is correct or not. There is further division of IDS based on rule based and signature based. In rule based IDS it analyzed the data on the other hand predefined signature or patter will be checked against incoming traffic [15] [16]. There is another IDS known as anomaly based IDS in which Mathematical model are implemented and study the system behavior [16].

- ACTION

In this functionality it has been defined about the reaction time of systems on attacks which are carried by hackers [17]. It can be of different type like it may alert the owner through mailing system or through alarm or it can play smartly in the system and start dropping the packets which are being sent by the intruder so that it cannot gain access through the system or even it can block the IP , if hackers id using DDOS then it may block the entire port and stop the communication [16].

1.4 IDS LIFE CYCLE

As there is vast increase in the network areas which enhance the fear of getting intrude. So this is also increases the production IDS , organizations frequently releases IDS with aggressive attitude to compete with the latest intrusion which is in the market [18]. Estimation about the new systems is not very efficient. Steps like hiring and training the workers with the administrator security , network security, benefits of IDS and even also about IDS is challenging task [18]. As the IT industry is growing faster and faster in which new technologies come and go. With the new technologies new intrusion also comes into the market and making an IDS for long term is a very challenging task.

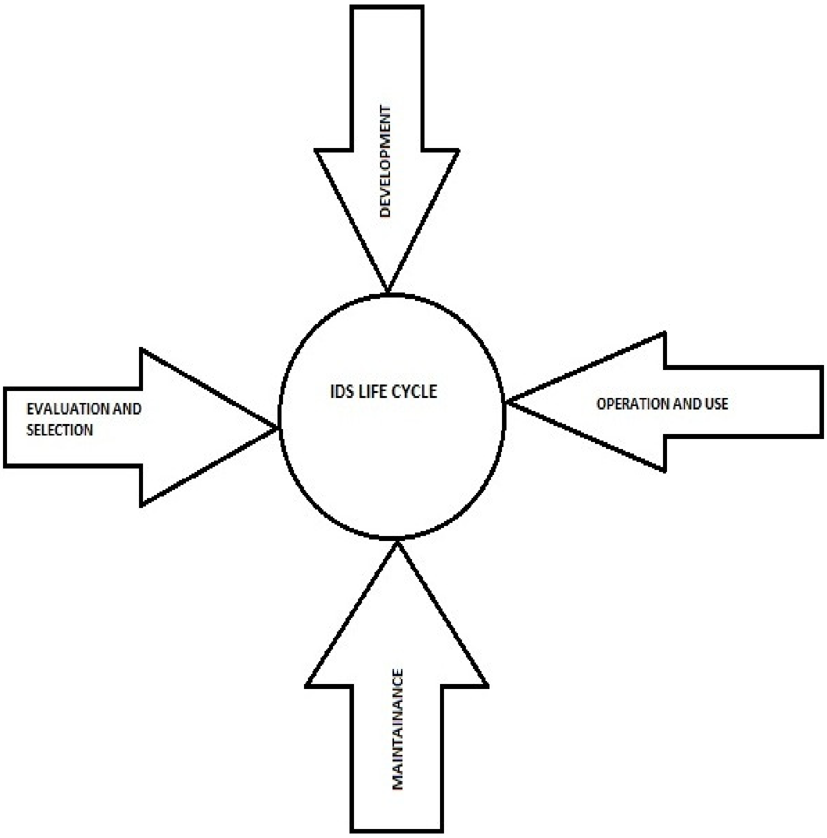


Fig 1.7 IDS Life Cycle.

- EVALUATION AND SELECTION

If there is need of an IDS system into an organization and planning to get, it should definitely inspect the assets for the system working and maintenance [18]. Lifecycle of the product can be increased by productive IDS. There is also option of third party evaluation for the IDS and their outcomes are generally available openly [18]. This certain process can tell us about how to find an intruder and how much amount of work needed for maintaining the system. Selection process in the system defines as approach of IDS , Identification of the characters, effectiveness and the accuracy of the IDS.

- DEPLOYMENT

In this phase it ensures working of sensor to increase the protection of the important files for the organization by configuring intrusion detection system with new attack policy after installing signatures [18]. Users of the IDS must design some new rules for monitoring the alerts and to synchronize with the other systems of the organization. The working group of Intrusion detection system (IETF) is creating a familiar attentive format that will use the IDS to alert different systems which will reportedly have a common display screen or console [18].

- OPERATION AND USE

Organization conducts the IDS to audit the host for responding on the report as alert. It builds roles and authority for auditing and monitoring results of manual and automotive responses [18]. Smart hackers have plenty of knowledge about the IDS systems, if they realize that in certain network there is IDS which has already been deployed on internet attack done by them, then they force IDS to provide false report to the owner of the organization [18].

- MAINTENANCE

Maintenance is very important factor of the IDS system because intrusion is increasing day by day; new attacks are there in the network field so we should install new signatures and upgrade the IDS for better security of the organization. We should check periodically the sensor of the system whether it ensures network changes [18]. For an organization it must attract qualified employees, train them and retain the qualified employees for better maintenance of IDS technologies [18].

1.5 IDS TECHNIQUES

1. *ANOMALY BASED INTRUSION DETECTION SYSTEM.*

Anomaly based intrusion detection system is a system used for detecting or identifying the intrusion in network as well as in computers, if any misuse is there by auditing the activity of the systems and classify it whether normal or anomalous with the help of predefined classical classifiers. The categorization is based on predefined rules or heuristics, instead of old pattern and signatures [19]. It will try to detect any type of malicious activity which does not fall in normal activity of system operations. This type of systems is good in respect of signature based intrusion detection because they can only detect the attack for which the signatures have already been defined [11].

In order to make our system feasible and identify the attacks with greater accuracy in network traffic attack, then the system must know about the normal activity of the systems and can able to distinguish between normal and abnormal activity. There are two major phases of anomaly based intrusion detection system which are training phase (where instructions are built for normal behavior) and testing phase (where current network packets are comparing with the defined behavior in the training phase) [20]. There are several ways for anomaly detection in the systems most common is Artificial Intelligent Techniques. Systems which are using neural networks have always given the best result. Techniques of deep learning using auto encoders [21], or generative models also give remarkable results. Another method is to define the normal usage activities of the system using strict mathematical models or Boolean flag. If any deviation occurs according to the defined model then it is an attack. This method is known as strict anomaly detection [22].

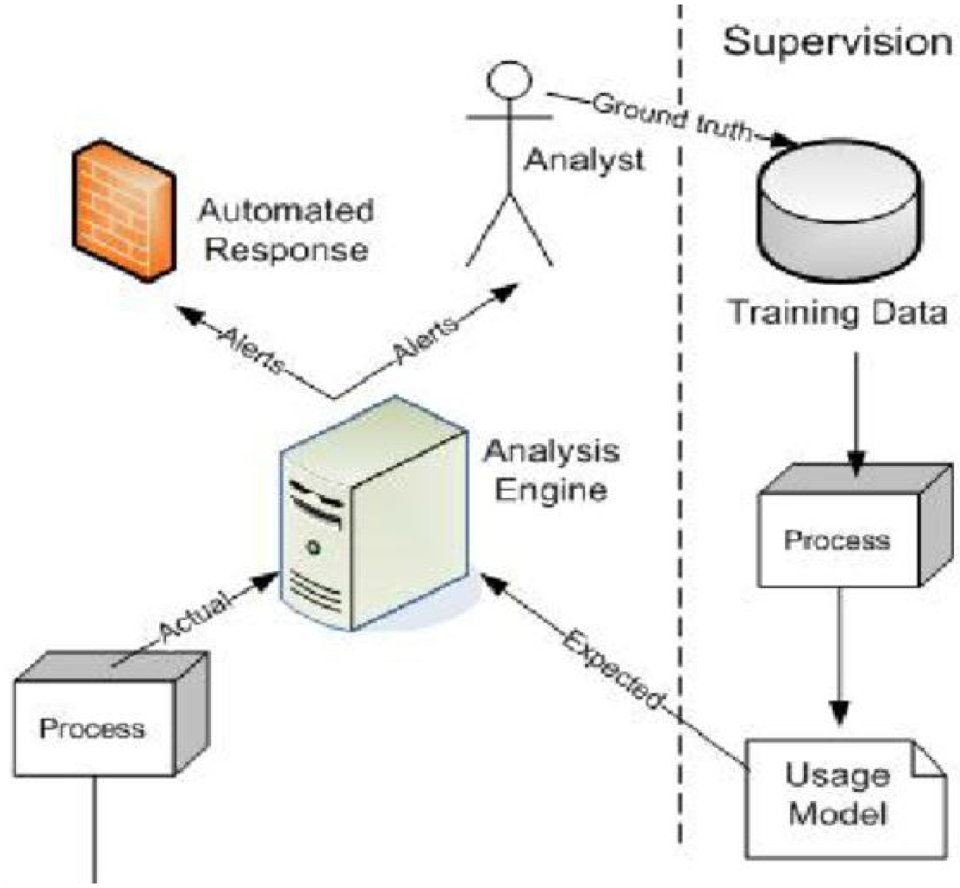


Fig 1.8 Anomaly Based Detection

Anomaly based provides good results for finding attacks like buffer overflow or memory leak, Denial of service (DoS) , Distributed Denial of service(DDoS) ,application overflow anomaly.

- ***TECHNIQUES USED IN ANOMALY BASED INTRUSION DETECTION***

There are several techniques defined for anomaly based IDS which have been already implemented.

- STATISTICAL MODEL

a. Operational model :- The action that happens over time which can regulate the alarming system. This can be evoked with WIN2K lock. Intruders after ‘z’ ineffectual access attempts can regulate the alarm. With the defined lower limit as 0 and upper limit as ‘z’.

b. Markovian process or stochastic process: - A stochastic process is collection of random variables with rest to time, so the system is inspected at fixed time and read the behavior[24]. If the behavior is anomaly then the probability of the current state is low as compared to the previously defined in the training set.

- COGNITION MODELS

a. Finite state machine: - This machine is also known as finite automata which is used to capture model of the behavior of states, transition and actions. We define state as the past information. An action contains the elaboration of the activity which is to be performed lately [24] there are certain actions entry, exit transition action [24].

b. Description Scripts: - Scripting languages like python can easily characterize the intrusion on computers networks. All scripting languages are self sufficient for examination of the specific event [24].

- COGNITION BASED DETECTION TECHNIQUES

CBDT works on the audited data which uses the predefined rules of classification for classes and tributes [24].

a. Boosted decision tree: - It uses ADA boost to create decision tree classifier based on the predefined data in the process of training in IDS [24].

b. Support vector machine SVM: - It is used for binary classification. If we merge decision tree and svm then it will provide efficient results.

2. ***Signature based Intrusion detection system.***

Signature based is used for intrusion detection. In this large dataset of instances is there and every instance of the dataset must be labeled with the normal or intrusion. Machine learning techniques have been used to train the data according to the labels. This method itself generates the signature to detect the intrusion. Misuse technique will get created automatically by the user, but work is harder and accurate compared to manually created [24].

- Techniques used in misuse detection

a. Express matching: This technique is the simplest to implement in misuse detection. In this it will look a sack of like events ex- log entries for finding the exact pattern.

b. State transition analysis: This model intrudes the state or transition of network. Every occurrence of event in internet network will be applied to finite automata machine instances which will results in transition. It is always believed that an attack will be occurred in the final state of machine.

3. ***Target monitoring***

In this technique IDS report the user if there is any modification happen in any location in the system. This is usually done through cryptographic algorithms, in which it calculates crypto checksum for each targeted file in the system [17]. If any changes occurred it will immediately informed to IDS. The name of checker which checks is any changes occurred or not is Tripwire checksum [25].

4. **Stealth probes**

This technique is used to gather and associates the data. It can detect the intrusion which generally takes long time to react. Hackers waits for over months to see the change in the system and also wait for another several months to generate attack in the system and take wide area of samples.

1.7 TOOLS FOR INTRUSION DETECTION

There are several tools available for intrusion detection for the organization to increase the security.

a. Snort: - Snort is open source software which is light and easy to understand. It uses flexible rule-based language to give information about data packets. Snort can detect vulnerability in the system, worm's suspicious behavior.

b. OSSEC-HIDS: - This is also a open source software which uses client/server architecture. It is recommended to run this software on major operating systems. Provide authentication logs, ISP, HIDS etc.

c. Fragroute:- This is a fragmenting router. Ip packets are sent to fragroute from attacker then fragment it and transformed to the party.

d. HoneyD: - It is a tool to create a virtual server, where attackers try to intrude it and our actual database is safe from the attacks. IDS analyze the type of attack and make our database even more invincible.

e. Kismet: It is basically a tour guide for the use of (wireless intrusion detection system). It finds the burglar point.

CHAPTER 2

RELATED WORK

This chapter is regarding the work done in the field of intrusion detection system. I have studied several technologies relating to this field which are efficient and up to date.

Machine learning is doing great job in intrusion detection systems. Classical Classification techniques of machine can be used for IDS but in today's life neural network is popular and attraction of the source. First we will discuss about the classical machine learning approaches, then about different approaches of IDS like IDS in data mining, fuzzy based IDS, Location based IDS (by Google).

2.1 There are four main processes for machine learning techniques to give results.

- **Data Preprocessing.**
- **Model Generation.**
- **Evaluation.**
- **Deployment in real network**.

2.2.1 *Data Preprocessing* : This is the very first step in every model generation process. In this step we actually excision the raw data and make it ready for model generation. These raw datasets contains numerous amounts of data with greater dimensionality which makes it bulkier to work with. There is also large amount of redundant data which has to be removed for further processes. Some feature of the datasets does not contain any value or does not have supported values then in this step we initialize these values with some constant value or make the void entry. Normalization is used to reduce the redundancy in the data set [26]. Extracting features from the data is a part of different machine learning algorithms.

2.1.2 *Model Generation*: Model building for the machine learning approach is crucial part. We have to deal with the high dimensionality of the datasets. We use dimensionality reduction or we only work on some percentage of the data sets. This is the core of IDS because on the basis of the model IDS detects intrusion. We generally divide the data set into training set, test set and validation set. Training set is use to train IDS which learning and training phase. Validation set validates the training phase as well as the training data. Final is the test set in which compare the output result with the test data to find the accuracy. We apply algorithm for many iterations in order to achieve high accuracy.

2.1.3 *Evaluation:* This steps consist of final output of the IDS compare with the actual result for evaluation of the IDS . We evaluate accuracy of the approach, time taken, easy to implement or not and make desirable changes to make out system more invincible.

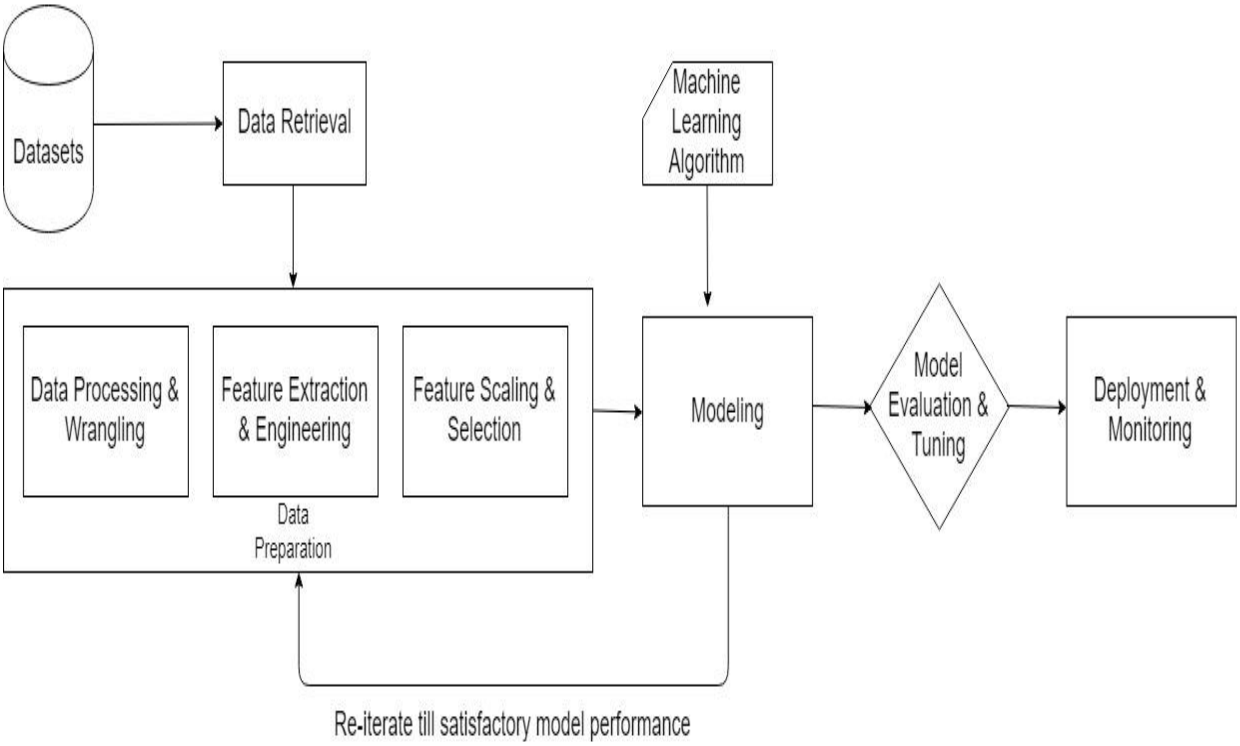

Datasets
Data Retrieval
Machine Learning Algorithm
Data Processing & Wrangling
Feature Extraction & Engineering
Feature Scaling & Selection
Data Preparation
Modeling
Model Evaluation & Tuning
Deployment & Monitoring
Re-iterate till satisfactory model performance

Fig 2.1 Machine Learning Training phase

2.1.4 *Deployment and monitoring :* After evaluation and making full analysis of the factors of IDS it is implemented into real network, where it will work as IDS. But here our work does not stop we monitor the process of IDS in the real time to make sure of its right working.

2.2 **Different machine learning techniques.**

2.2.1 *Linear regression*:

Linear regression is used to find relationship between input and targeted variables which are dependent and independent variables. Logistic model are only used for predicting the continuous values such as predicting the price, predicting the height of family members etc.

We cannot use linear regression for classification use, so we use logistic regression. It predicts the probability such as true/false, head/tail, which can easily classify the datasets into classes.

$$Y_i \ = B_0 + B_1 X_i + e_i$$

Linear regression equation

Y_i = Dependent variable.

B_0 = Y intercepts.

B_1 = Slope coefficient.

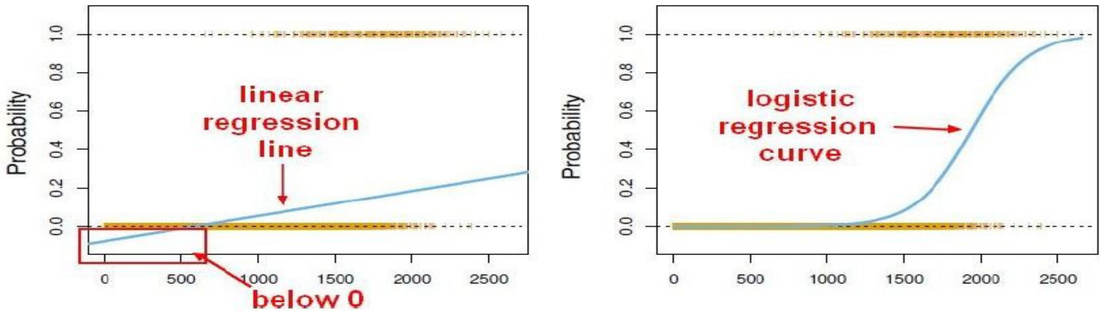
X_i = Independent variable.

e_i = Random Error.

Logistic Regression method for classification (Binary classification) :

- Spam vs Ham.
- Loan Default.
- Disease diagnosis.

So far we have predicted the continuous values, now we will predict the discrete values. We cannot use linear model in binary groups, it won't give better results. So we will transform our linear model into logistic regression curve.



Sigmoid function is used for logistic regression in which it gives output in the range (0,1) for any input. Let's set cutoff 0.5. If value is smaller than 0.5 it falls in class 0 otherwise in class 1.

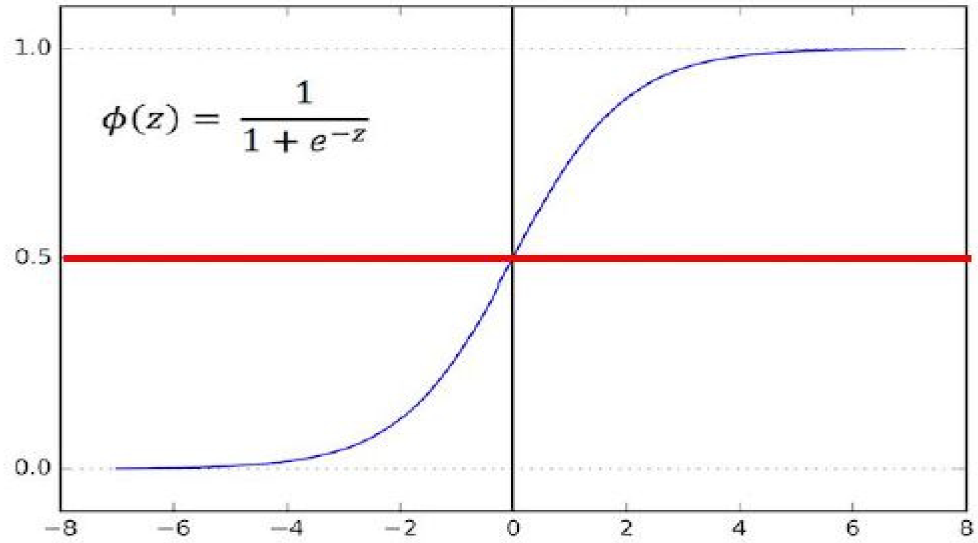


Fig 2.2 Sigmoid Function cutoff.

After training we evaluate the result on different data sets. We will find confusion matrix for classification.

Basic terminology

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

N=165	Predicted No	Predicted Yes
Actual No	50	10
Actual Yes	5	100

	Predicted No	Predicted Yes	
Actual No	TN=50	FP=10	60
Actual Yes	FN=5	TP=100	105
	55	110	

Accuracy = (TP+TN)/Total = 150/165 =0.91

Misclassification Rate = (FP+FN)/Total= 15/165 = 0.09

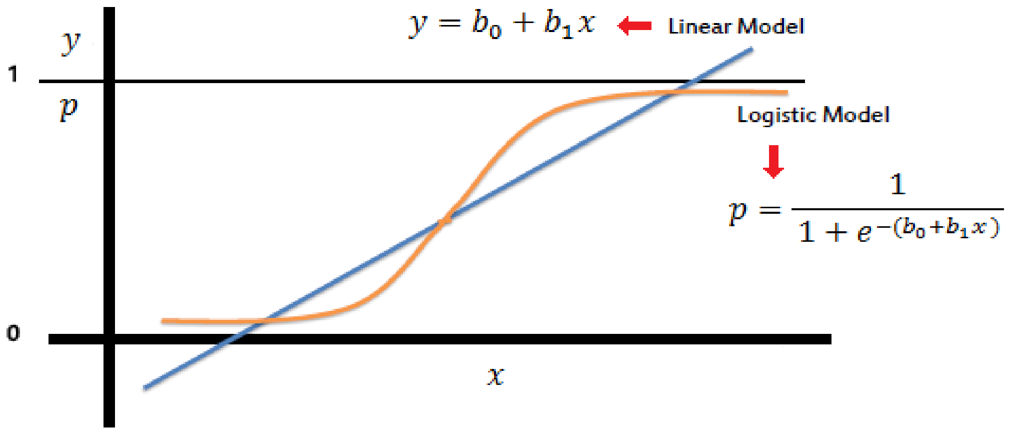


Fig 2.3 Linear and Logistic regression.

2.2.2 *Naïve Bayes Classifier :* This is a simple technique in probability and statistics used to classify which is based on simple bayes theorem.

$$\Pr\,(A|B) = \frac{\Pr(B|A)\,.\Pr\,(A)}{P(B)}$$

Bayes Theorem.

A,B = Events

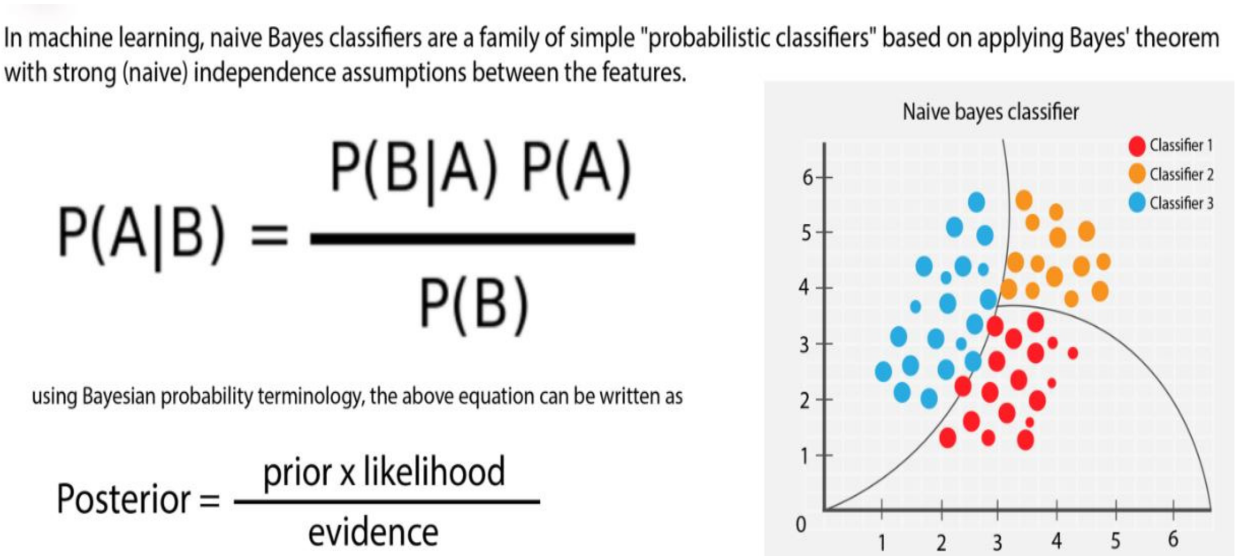
Pr(A|B) = Probability of A when B is given.

Pr(B|A) =Probability of B when A is given.

P(A) = Probability of A.

P(B) = Probability of B.

It is assume that features are mutually exclusive or independent of each other. These classifiers are simple to implement but can produce high accuracy [27] when combine with kernel density estimation function.



In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$

using Bayesian probability terminology, the above equation can be written as

$$\text{Posterior} = \frac{\text{prior x likelihood}}{\text{evidence}}$$

Fig 2.4 Naïve Bayes

This technique is highly scalable which require maximum of linear variables for prediction/feature extraction in learning phase. It takes linear time for evaluating rather than expensive iteration in most of other classifiers.

2.2.3 *K nearest neighbors KKN:* This is a classification algorithm which is very simple to implement. It is difficult of predict the value of K , but it is recommended k=sqrt(n) which represent the no of classes in classification. Output is in the form of class membership of data variables.

Let us suppose we have two imaginary datasets of horse and dogs associated with heights.

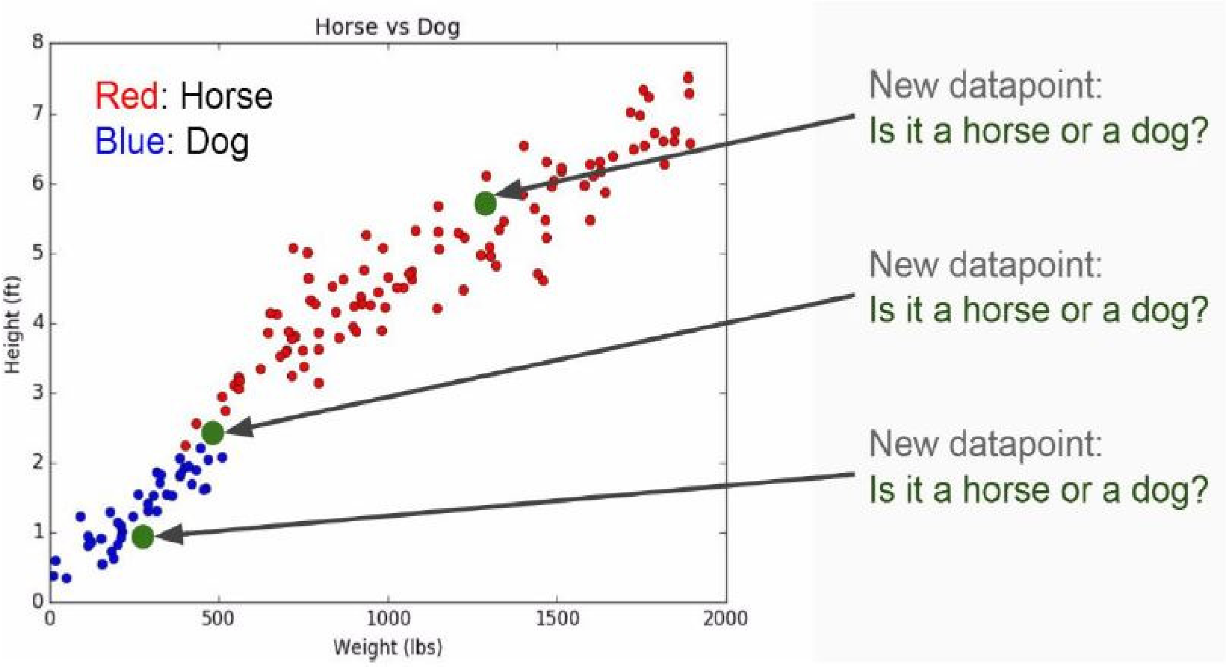


Fig 2.5 KNN Algorithm.

- Train algorithm by storing the data.
- We will calculate the distance distance of all points from Y.
- Majority label will be predicted of 'K' point.

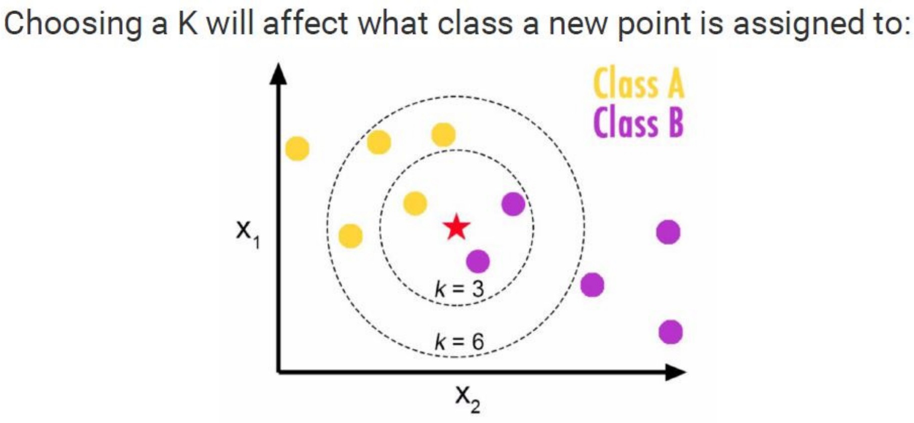
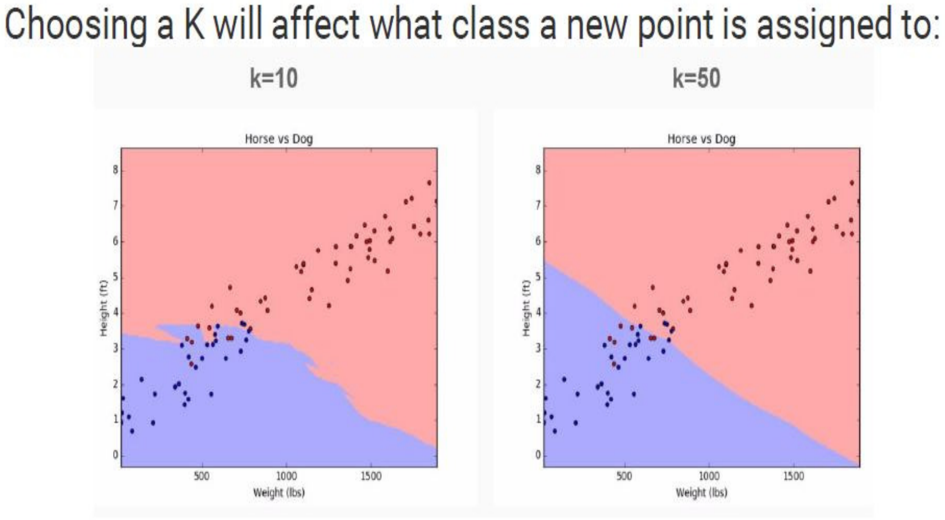

Choosing a K will affect what class a new point is assigned to:
Class A
Class B
X1
X2
k = 3
k = 6

Fig 2.6 Choosing K.



Choosing a K will affect what class a new point is assigned to:

Fig 2.7 K effect.

Pros

- Very simple to implement.
- Training is insignificant.
- Can be work with any no of classes.
- More data can be added easily.
- Parameters
  - K
  - Distance metric.

Cons

- High Cost.
- Not efficient with high dimensional data.
- Classification doesn't work well.

2.2.4 *Decision Tree:* Decision tree is tree like structure with several Childs for their decision and consequences. Each node represents a decision whether yes or no, also represents chance, outcome, costs, utility. This can predict the behavior like my friend plays badminton with me but sometime he won't come due to certain factors like bad weather, temperature, humidity so I have tracked all the features of him to predict whether he comes or not.

For detecting whether he shows up or not the institutive way is decision tree.

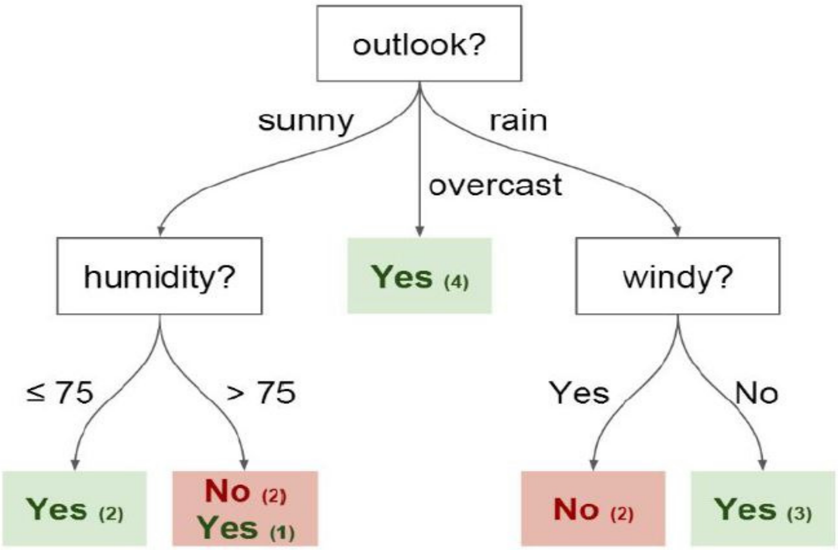


Fig 2.8 Decision Tree.

- Nodes: Split for attributes.
- Edge: Outcome.
- Root: Perform split.
- Leaves: Have no Childs which predict the output.

$$Entropy$$

$$G(S) = -\sum_{i}^{n} p_i\,(S)log_{2p_i(S)}$$

$$Information\ gain$$

$$IH(S,A) = G(S) - \sum_{v\,\in \text{Values}(A)} \frac{|S_v|}{S}\,G(S_v)$$

To improve the accuracy of the decision tree we can actually choose many other trees for splitted feature from random sample. This technique is known as Random forest which is RF. New sample will be chosen for every tree and for splitted nodes where it is recommended to choose m as sqrt of p.

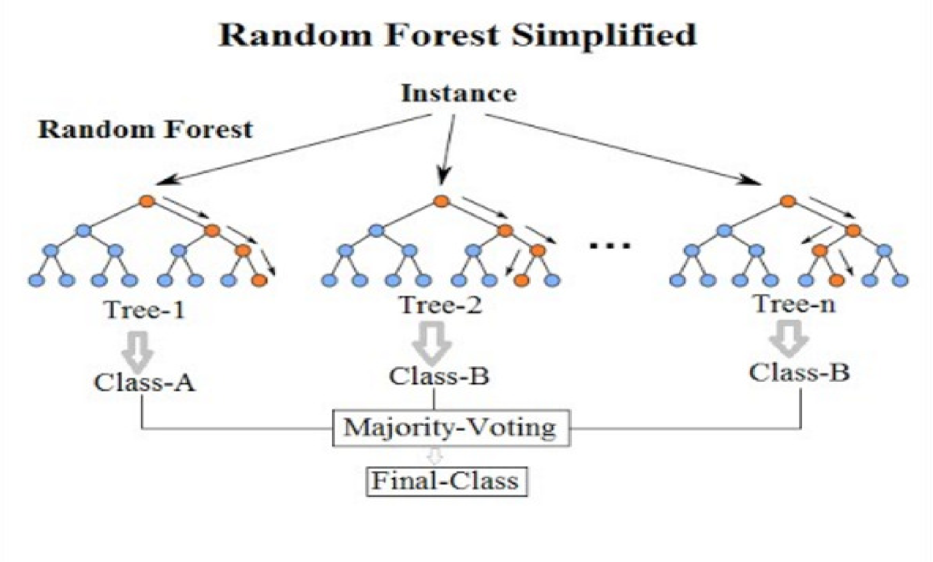


Fig 2.9 Random Forest

It is a classifier based on ensemble learning method like we have a strong feature dataset, we use bagged tree, so in this most tree will choose that to split highlighting ensemble of similar trees which are highly correlated. Average of highly correlated feature does not lead to reduction of variance. Random forest decorrelate the tree by choosing feature from each split randomly so this can results in reduction of variance.

2.2.5 Adaboost: It is also same as decision which is based on ensemble learning , which was then created for increasing the efficiency of classifiers which can be binary classifier. This is short and adaptive boosting algorithm. It works on iterative method which iterates over and over for better results. Learning from its initial mistake and making the weaker classifiers into stronger one.

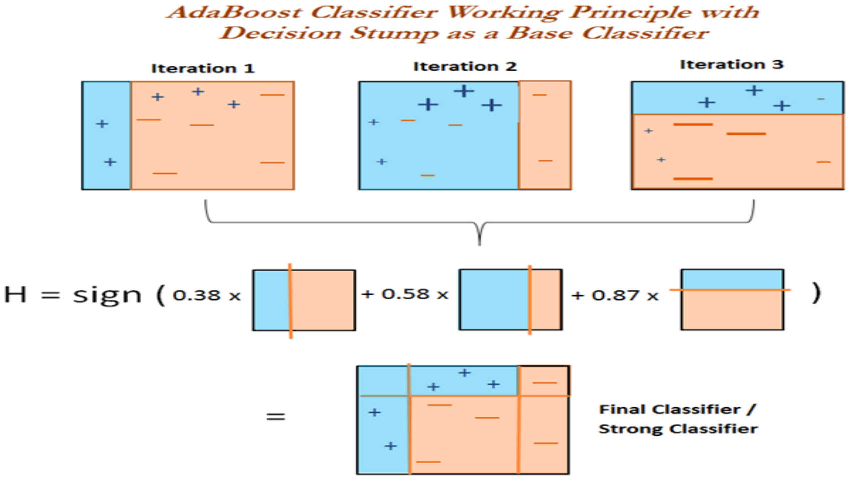

AdaBoost Classifier Working Principle with
Decision Stump as a Base Classifier
Iteration 1
Iteration 2
Iteration 3
H = sign (0.38 × + 0.58 × + 0.87 ×)
=
Final Classifier /
Strong Classifier

Fig 2.10 Adaboost.

2.2.6 *Support vector machine* : Support vector machine is classical classifier used in machine learning techniques which is a supervised learning.

- Supervised learning: This is a machine learning task which is purely based on input and output pairs of datasets. It predicts the function on the basis of the input given and expected output.

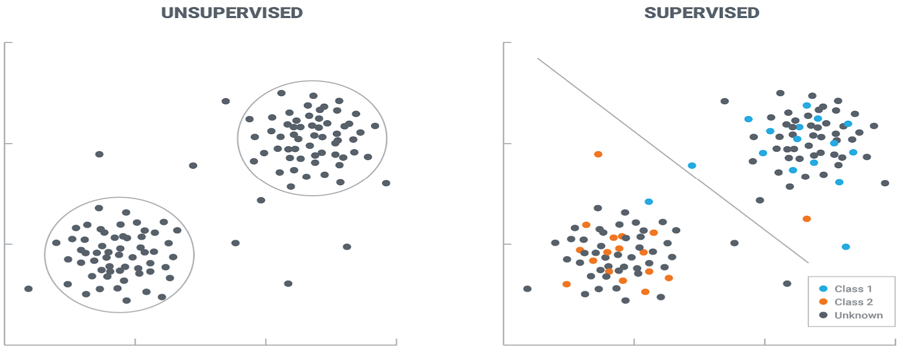


Fig 2.11 Unsupervised & Supervised

- Unsupervised learning : This will help in determining the pattern in the datasets in which dataset is not associated with the outputs, which always look for previously undetected patterns.

SVM is very popular in machine leaning field for classification and regression problems. SVM model is like points in space, so we can easily divide the categories by gap which is good if with increase of gap for classification in which a line has been drawn to classify the attributes.

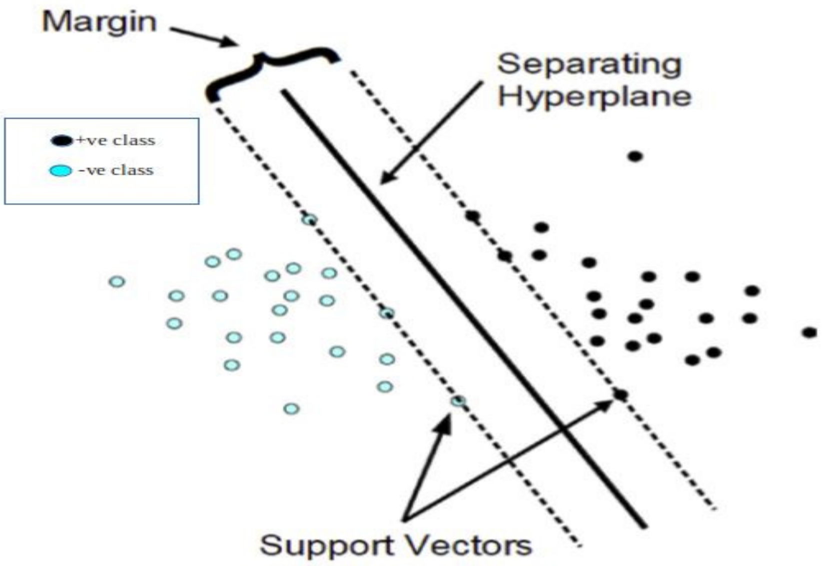


Fig 2.12 Support vector Machine.

The research on IDS is increasing more with the increase of network demand. As the demand increases we are switching on machine learning techniques (ML) to create aggregate ID system which is becoming common day by day. Datasets are limited in the real world which caps the ability of getting trained. DARPA dataset is all-embracing dataset available for public use so we can learn about the nature of real time intrusion. In 1998 this datasets gets clean and a contest was organized in 1999 named as KDDCUP which was named as fifth international conference on Knowledge Decision and data mining. Data was already preprocessed into normal traffics in the following intrusions DoS intrusion, Probing, U2R, R2L.

The initial work was done using MADAMID framework in the completion [29]. Top three places team used various kinds of decision trees which showed marginal difference. Top 17 submissions considered to be good and summarize in [30]. Major part of the submissions were tested on 10% of KDDCUP dataset few researcher also used custom datasets [31].

The main reason ML- based approach for IDS is emerging because of its potential to caps the threat from newly emerging intrusion which is complex, diverse and also deliver acceptable false positivity rate with low cost of computation. In past [32] used for PNrule derived from P and N rule which is used to find existence and non existence of class. This can detect other types of rate with good efficiency except U2R category attacks.

Artificial Neural Networks: It is a computing system inspired by biological nerve system of animals. Working of ANN is similar to human brain [38], it learn from earlier events and perform certain tasks. Like we can give certain datasets of dog images and based on these earlier events it can predict the image of Dog. They are just being trained by the image datasets of dogs not any knowledge like they have fur, or they have 4 legs etc. They can automatically generate the classifying units among them.

Instead of structure like human brain as we can see in fig 2.13 artificial neural network have

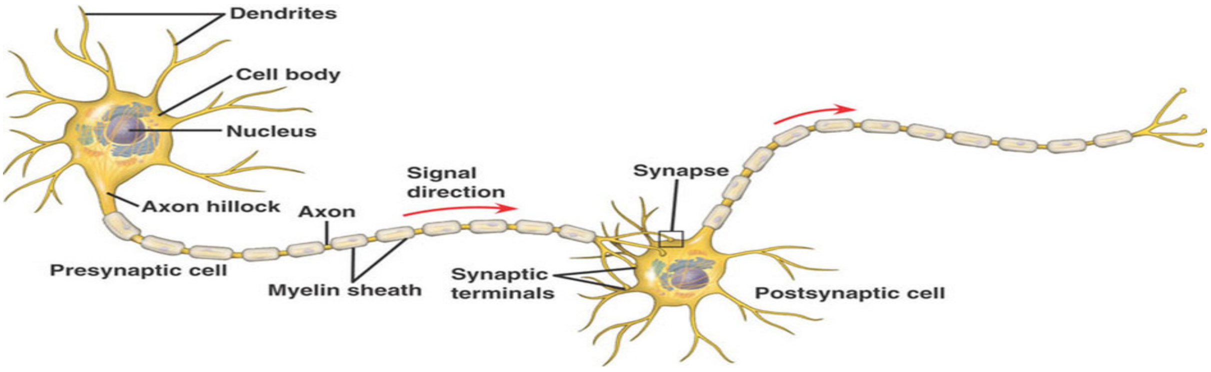


Fig 2.13 Nervous system

Different structure as shown in fig 2.14.

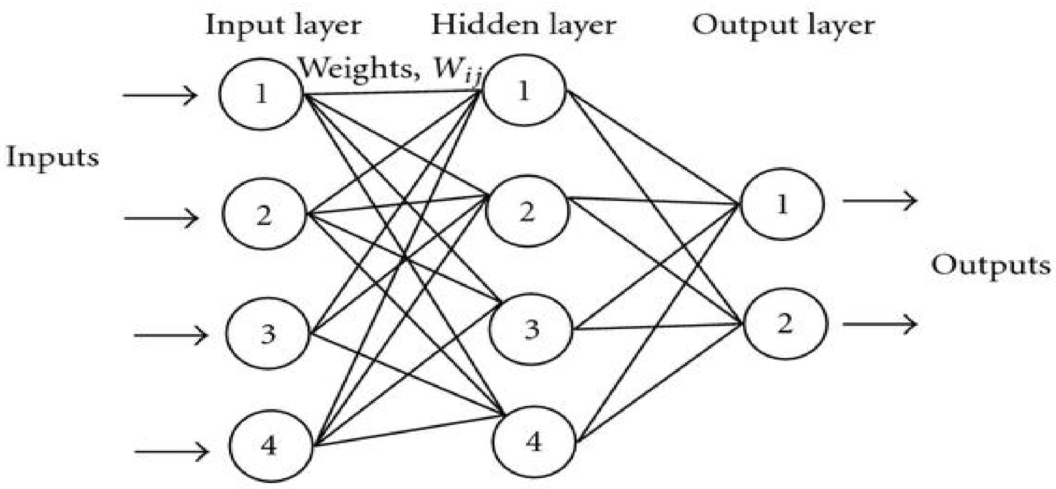


Fig 2.14 Artificial neural network.

Types of neural network:

a. Feed forward ANN: The flow of information in feed forward neural network in unidirectional. Weights cannot update if there is error.
b. Feedback neural network: This network can flow information in both backward and forward directions. The main fundamental feature is they can update the weight via back propagation and improve its accuracy for finding results.

The assumption to long establish FFN inspired from biological events is basically a network in early stages which is known as Convolution neural network. It was used for preprocessing of images. They use several types of method for preprocessing like they use 2D layers, connected layers or they also use pooling 2D layers. In [32] in it described about Convolution neural networks for Intrusion detection systems and KDDCUP “99” dataset was used, further they compare the results with other algorithms. After getting the results and with some research they have brought the conclusion that CNN has performed better as compared other algorithms. Then a great classifier with the ability of seeing past and can assume certain records named (LSTM) Long Short-Term [33] memory was used with KDDCUP “99” dataset which has given great results in Intrusion detection.

Deep Neural Networks: These networks are artificial neural network in which input and output layer are involved with several multi layer. They can easily model complex relationships of non linear datasets and represents/produce arithmetic model in which objects represents as the combination of layers of primitive.

Classical machine learning classifier algorithms are linear in nature whereas deep neural network consist with complexity and hierarchy with level of abstraction. Each layer in deep neural network (DNN) applies a function while taking input such as non linear transformation and gives output such as statistical model from whatever it learns from its training phase. Input is accepted by the input layer then it is passed on to hidden layer. All the hidden layers perform arithmetic operations on our inputs. The main challenging task in neural networks is to decide the no of hidden layers and the neurons count in each layer. Neurons in the neural networks comprises with activation function to standardize the output. The meaning of "Deep" in deep neural networks or deep learning is which have more than one hidden layer. In neural networks output is given by the output layer, and until the output is not accepted in terms of accuracy the iteration continues.

Now we will see about ReLU activation function which is Rectified linear units. ReLu has its own different kind of capabilities which makes it more efficient and capacity of increasing the process of neural networks altogether [34]. Generally Artificial networks (AAN) use sigmoid activation function or tangent function but these activation functions can lead to a problem of gradient vanishing [35]. Disappearance of gradient when in DNN its lower layers have gradient value of nearly null, this is because of higher layers of DNN saturated at the asymptotes which consist of the function known as hyperbolic tangent. ReLU gives an approach to sigmoid non –linear function which can mitigate the discussed issues [36].

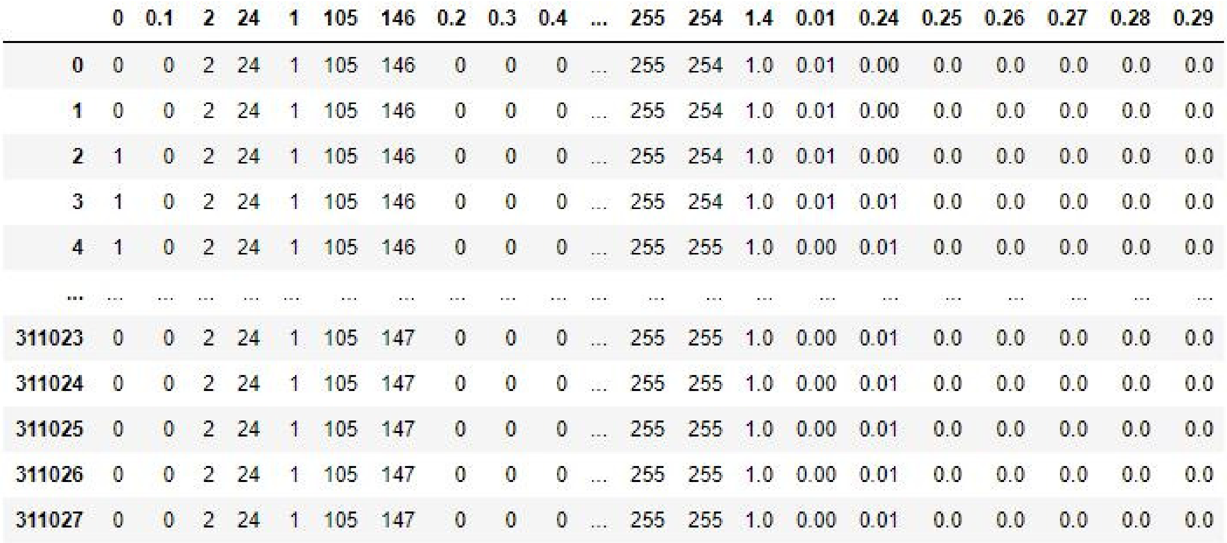
CHAPTER 3

EXPERIMENTAL APPROACH

In this project Keras [37] is used on the top of tensor flow [38] for increasing the processing of data exponentially in deep learning architectures.

DATASET DESCRIPTION

In 1998 Lincoln Labs of MIT started a DARPA's program in which evaluation of ID of 1998 was directed and created. The main focus of the program was to inspect and conduct experiments and research on ID. Then a dataset was prepared after certain steps of cleaning and preprocessing which have several types of attacks emulated in a military environment and was made standard dataset which is publicly available. The dataset of KDD'99 was well refined and systematic version [39].



	0	0.1	2	24	1	105	146	0.2	0.3	0.4	...	255	254	1.4	0.01	0.24	0.25	0.26	0.27	0.28	0.29
0	0	0	2	24	1	105	146	0	0	0	...	255	254	1.0	0.01	0.00	0.0	0.0	0.0	0.0	0.0
1	0	0	2	24	1	105	146	0	0	0	...	255	254	1.0	0.01	0.00	0.0	0.0	0.0	0.0	0.0
2	1	0	2	24	1	105	146	0	0	0	...	255	254	1.0	0.01	0.00	0.0	0.0	0.0	0.0	0.0
3	1	0	2	24	1	105	146	0	0	0	...	255	254	1.0	0.01	0.01	0.0	0.0	0.0	0.0	0.0
4	1	0	2	24	1	105	146	0	0	0	...	255	255	1.0	0.00	0.01	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
311023	0	0	2	24	1	105	147	0	0	0	...	255	255	1.0	0.00	0.01	0.0	0.0	0.0	0.0	0.0
311024	0	0	2	24	1	105	147	0	0	0	...	255	255	1.0	0.00	0.01	0.0	0.0	0.0	0.0	0.0
311025	0	0	2	24	1	105	147	0	0	0	...	255	255	1.0	0.00	0.01	0.0	0.0	0.0	0.0	0.0
311026	0	0	2	24	1	105	147	0	0	0	...	255	255	1.0	0.00	0.01	0.0	0.0	0.0	0.0	0.0
311027	0	0	2	24	1	105	147	0	0	0	...	255	255	1.0	0.00	0.01	0.0	0.0	0.0	0.0	0.0

Fig 3.1 Dataset

Shortcoming of KDD'99 dataset

ReLu activation function gives more efficient results and shortcomings of available standard datasets such as KDD'98 and KDD'99 has been discussed with detailed report in [40]. The main criticism made to KDDCup-'99 was that they failed to validate the dataset real world intrusion network profile, still the KDDcup-'99 is mostly used worldwide for research process worldwide and several algorithms was proposed for intrusion detection.

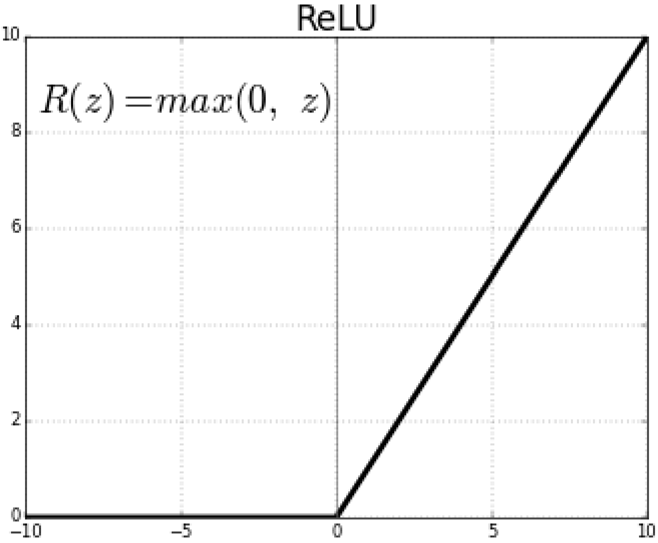


Fig 3.2 Relu Activation Function

$$f(x) = \max(0, x)$$

Reason behind the classical machine learning classifiers shows less capability to detect the attacks such as R2L, U2R in the KDDCUP dataset briefly discussed in [40]. They have discussed for ML algorithms it is impossible to achieve desired detection rate, if somehow we want high rate of detection of intrusions in our systems, in most of the cases we have to produce dataset by combining test and train data, output dataset must be refined and augmented but the approach is not been discussed.

The DARPA'S KDDCUP dataset did not give prominent results with traditional intrusion detection systems which lead to criticisms. The problem was eradicated via using snort IDS. However the results were not so good enough and false positive rates were high and also failed to detect dos and probing category attacks but coincidently performed good for R2L and U2R. KDDCUP is still most widely used dataset for IDS evaluation.

So as far we have KDDCUP but the efforts are being taken to mitigate the problems with current KDDCUP'99, then a redefined new version of dataset was created which was named as NSL-KDD [41]. In this dataset record of connection redundancy was removed from both train and test data and invalid records were also removed. These certain measures can limit the classifier to go in the direction of most frequent records or we say it caps the biased nature of dataset. Even though this dataset failed to give promising results for the problem discussed in [42, 43]. So new dataset was created named was UNSW-NB15.

DARPA / KDDCUP-'99' DATASET

The evaluation group of DARPA collected data based on networks IDS by the process of air force base local area network (LAN) with UNIX nodes of more than 1000s and continuously monitoring for over 9 weeks. Hundreds of users was gathered in Lincoln Lab and then created into groups of 7 and 2 weeks of testing and training and raw TCP data was extracted. In MIT lab the Windows, UNIX operating system was used for like almost all the incoming intrusion. The dataset consists of 32 distinct attacks with seven different scenarios and total attacks 300 were simulated. KDD-'99' dataset grouped with forty nine lakhs 49,00,000 different connections which has feature count of forty one (XVI) . These attacks were classified briefly as given below.

- *Denial of Service Attack (Dos)* = A denial of service attack is a malpractice attempt which makes the server and network unavailable to the users. In this attack hacker sends plenty of requests to the server and makes it chock. For every server there is a limit to reply no of queries made by it, if the queries exceeds the limit server gets crashed.

- *User to root Attack (U2R)* = In this attack attacks tries to gain access into the system through internet after data leak in the system. This data leak occurs when server receives more data than its capacity and not being programmed to handle that much amount of data, in this condition buffer overflow happens. Buffer overflow also happens if the program to handle the server is not good, like buffer overflow can occurs in the program of the system.

- *Remote to local Attack (R2L)* = In this attack attacker send packets to the targeted system and gain access into it as a legitimate user to vandalize or steal the data.

KDD-CUP'99 dataset divided into 3 groups: Primary features: Attributes which is collected from a collection of TCP/IP is belonging to this group. In this feature they show delaying in detection. Traffic feature: Feature calculated w.r.t. time which is further categorized into two groups.

- *Same Host Feature:* The connections has similar host as the connection which are under consideration for at least two seconds comes in this category. This also serves as calculating the behavior of protocols and represents it into some statistics.

- *Same Service Feature*: The connection which has only similar services as present connections until two seconds comes in this category.
- *Content Features:* In general attacks like probing and Dos (denial of service) have some common sequential patterns of intrusion dissimilar to R2L and U2R attacks. The main reason is that they have many connections to isolated set of host which is also only for short period of time while the two others intrusion attacks are combined with data partitions packet which is generally associated with only one connection. For detecting these attacks we required few extra notable features by which we can easily search for the irregular behavior. This is called content features.

Identifying networks parameters

To get the optimal set from the given sets to achieve the output by doing hyper-tuning of given parameters is itself an entirely different field with lots of further scope and research work. In defined approach the rate of learning is always remains constant which is at 0.01 and several other parameters are improved. The count of neuron in a layer was drastically changed between the ranges of 2 to 1024. The neuron count was increased ahead to another level to the value of 1280 but it didn't show the promising results in terms of better accuracy so the neuron count was adjusted to 1024.

Identifying network structures

Truly speaking increasing the number of counts of neuron in layers doesn't yield to promising results but increasing the counts of yield into good results. So we have used the following network topologies after closely monitored and culminate the optimal structure of network for our datasets.

- DNN (deep neural network) with the one, two, three, four, five layers

For the above network topologies, 100 iterations (epochs) were executed then the outputs was thoroughly observed. The results shown by the DNN with 3 layers was very promising and the best among all others. To elaborate the exploration for better outcomes different type's classical ML algorithms were used and the outcomes of the algorithm were compared to result of DNN 3 layer's outcome which still provides the best results among them.

Proposed architecture

This is an over-view of DNN architecture for all the instances and algorithms as shown below.

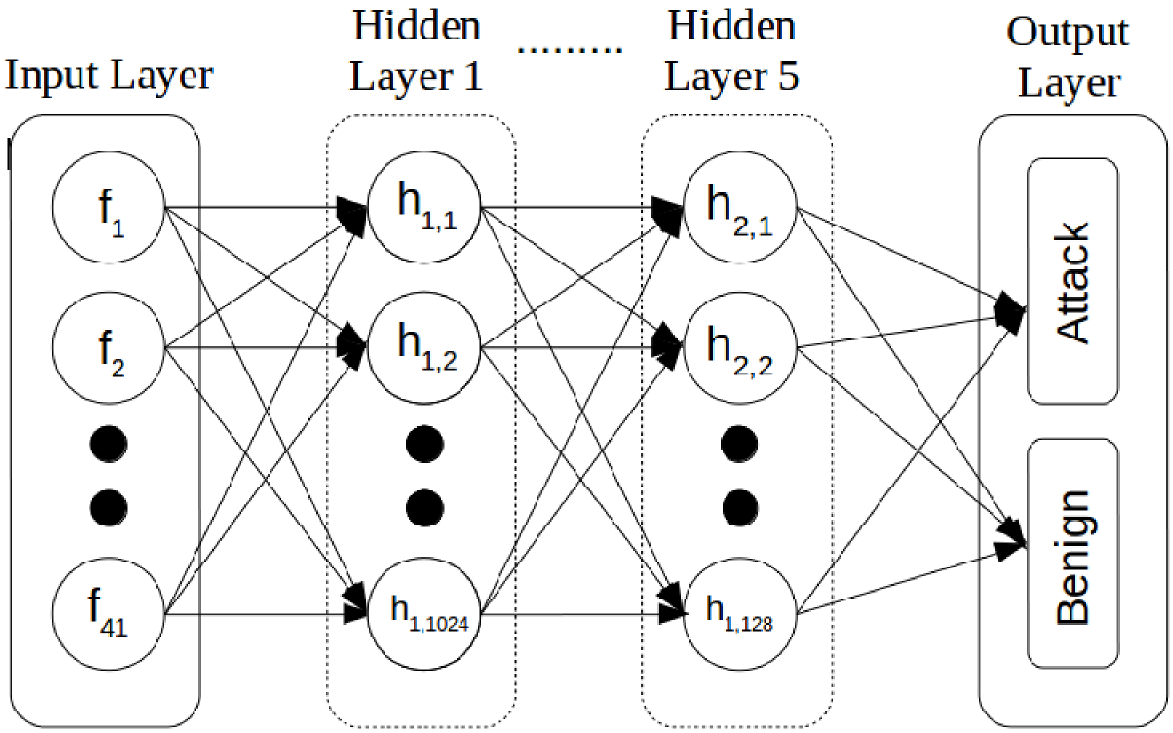


Fig 3.3 Proposed architecture

This architecture comprises of 5 hidden layers with one output layer. As shown in fig Input layer of the architecture comprises of 41 neurons. Neurons in this architecture are fully connected from input to hidden-layer and also they are fully connected from hidden to output layer. Mechanism was used back propagation to train the Deep Neural Networks (DNN) networks. This architecture also comprises of fully connected bias layers, drop out layers to make the architecture as more powerful network.

Input layers and hidden layers: In this layer there the neuron count is 41 which are then integrated into the architecture's hidden layers. These layers uses an activation function called as ReLu (Rectified activation) function as a non uniform activation function. Then several counts of weights are integrated into the layers to perform the next operation of feeding them forward for next hidden layers. Then count of neuron for every hidden layer will be reduced slowly from the start makes result more accurate and also the computational cost was also reduced.

Regularization: For making the whole system efficient, accurate, faster to process, Dropout (0.001). The dropout is for unplugging the neurons in randomizes way, which makes the system robust and also protects our system to over-fit out datasets.

Output layer and classification: As we have discussed output layers only consists of two neurons named as Attack and Benign. Our hidden layers consist of 1024 neurons so we have convert this to 2 neurons layers for this sigmoid activation function is used. As we are aware about the sigmoid activation function which gives binary (two) outputs which favors the binary categorization and also the results of this approach.

CHAPTER 4

RESULTS

The results of this approach, the KDDCup-'99' datasets was used for traditional machine learning algorithm and also for the DNN (Deep neural networks) for different hidden layers was also used. The training of the systems, several techniques, models; algorithms and there results was compared for F majors like (f1-score, precision, recall, accuracy).

First we will see the results with classical machine learning algorithms.

1) Linear Regression.

Accuracy=0.848, Precision= 0.989, Recall= 0.821, f1-Score =0.897

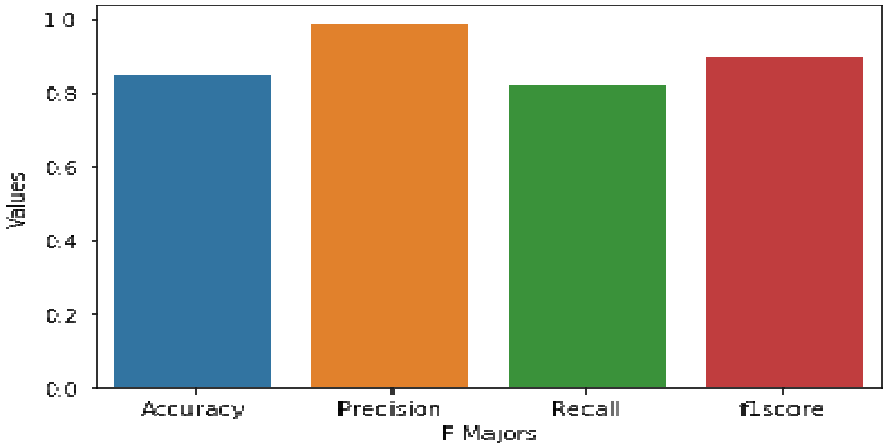


Fig 4.1

2) Naïve Bayes

Accuracy=0.929, Precision= 0.988, Recall= 0.923, f1-Score =0.955

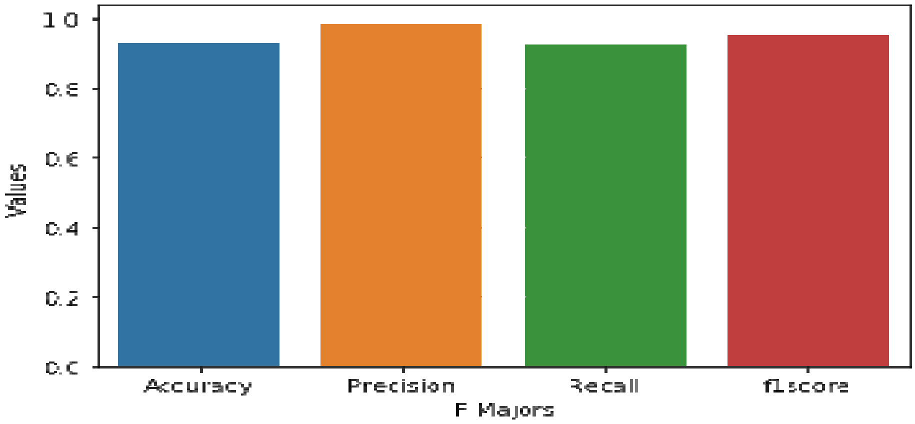


Fig 4.2

3) K nearest neighbor

Accuracy=0.929, Precision= 0.988, Recall= 0.913, f1-Score =0.954

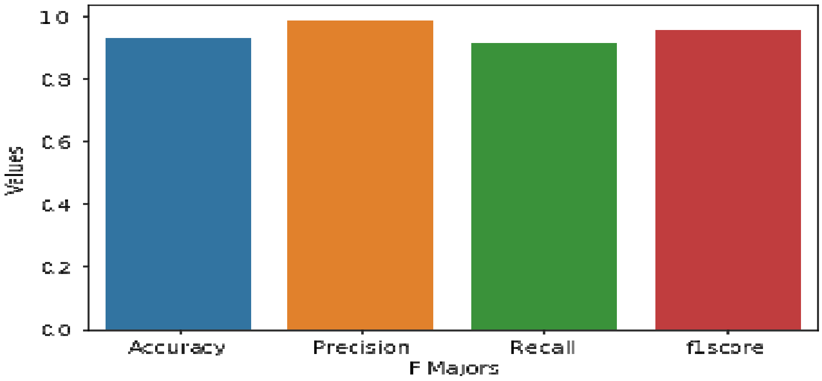


Fig 4.3

4) Decision Tree

Accuracy=0.928, Precision= 0.999, Recall= 0.912, f1-Score =0.953

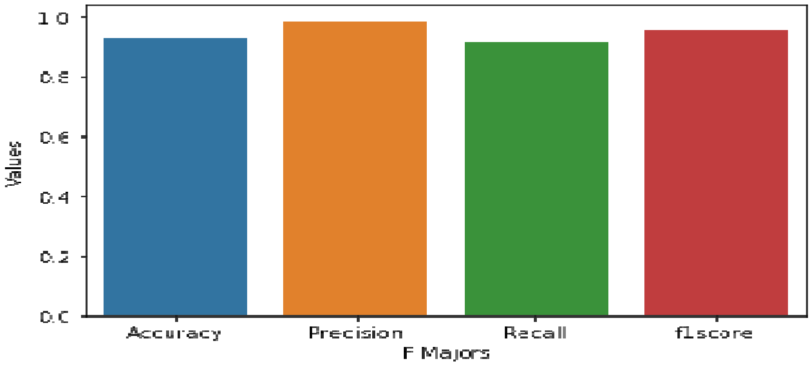


Fig 4.4

5) Adaboost

Accuracy=0.926, Precision= 0.996, Recall= 0.912, f1-Score =0.950

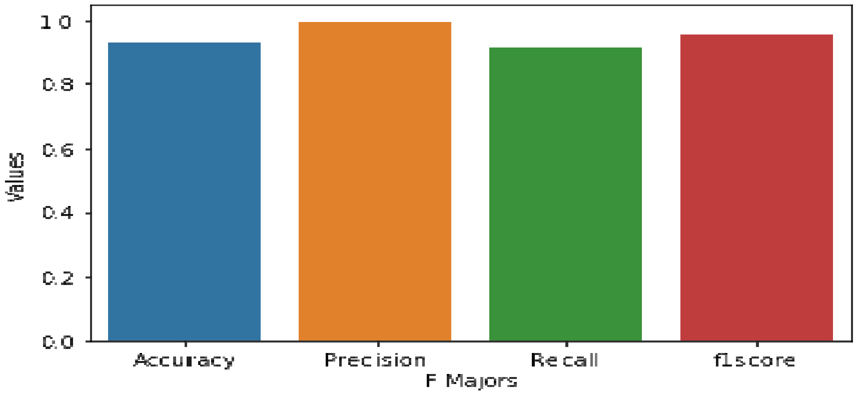


Fig 4.5

6) Random Forest

Accuracy=0.927, Precision= 0.998, Recall= 0.911, f1-Score =0.951

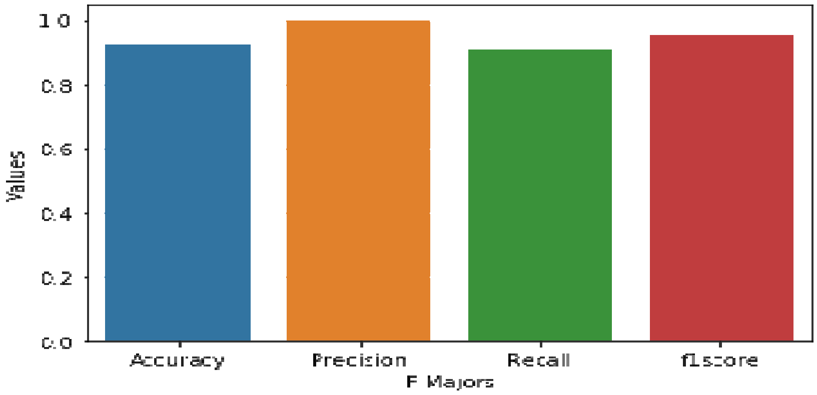


Fig 4.6

7) SVM rbf

Accuracy=0.811, Precision= 0.992, Recall= 0.772, f1-Score =0.868

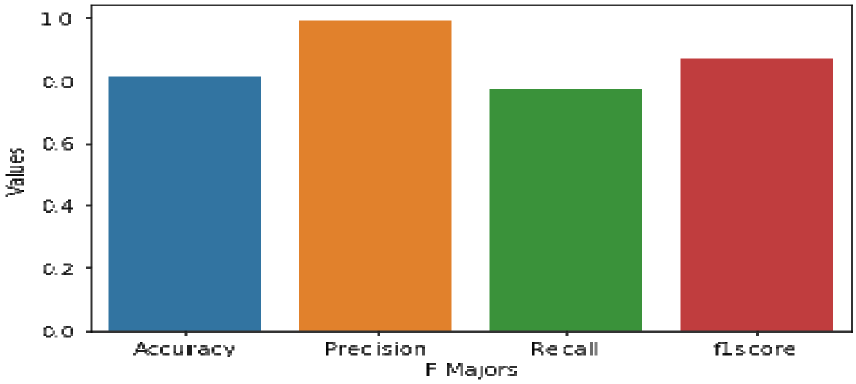


Fig 4.7

8) SVM Linear

Accuracy=0.811, Precision= 0.994, Recall= 0.770, f1-Score =0.868

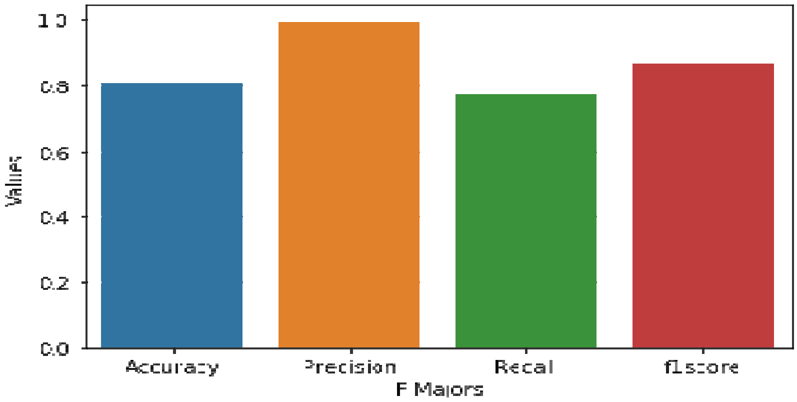


Fig 4.8

9) DNN-1

Accuracy=0.929, Precision= 0.998, Recall= 0.915, f1-Score =0.954

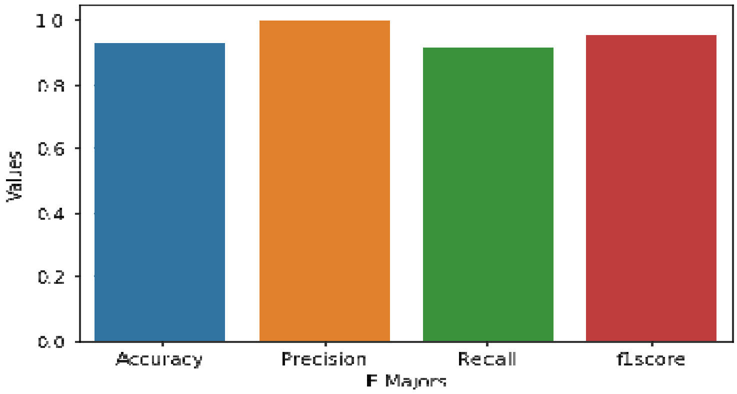


Fig 4.9

10) DNN-2

Accuracy=0.929, Precision= 0.998, Recall= 0.914, f1-Score =0.954

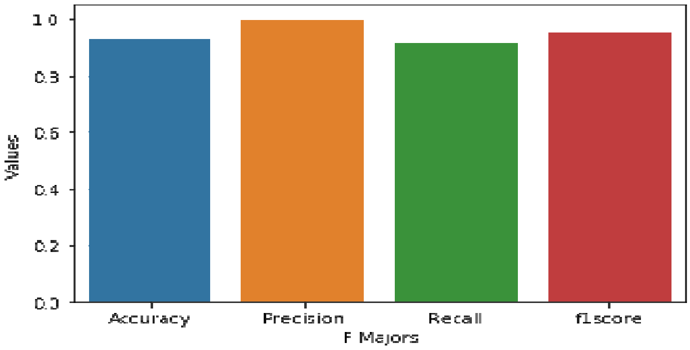


Fig 4.10

11) DNN-3

Accuracy=0.930, Precision= 0.997, Recall= 0.915, f1-Score =0.955

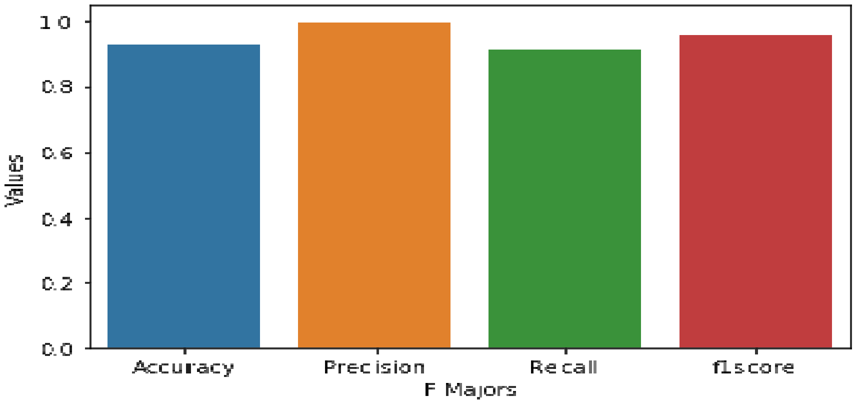


Fig 4.11

12) DNN-4

Accuracy=0.929, Precision= 0.999, Recall= 0.913, f1-Score =0.954

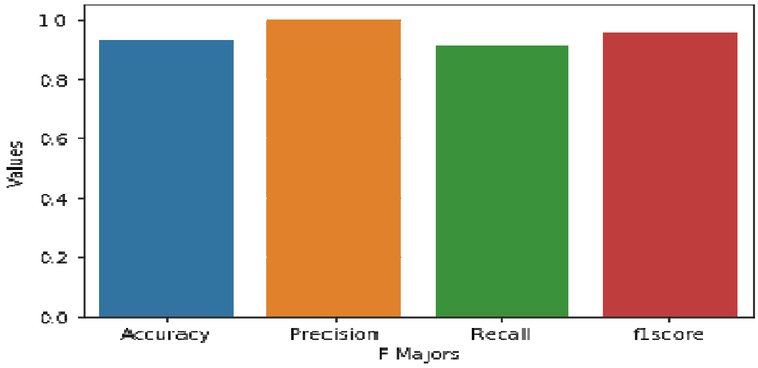


Fig 4.12

13) DNN-4

Accuracy=0.927, Precision= 0.998, Recall= 0.911, f1-Score =0.953

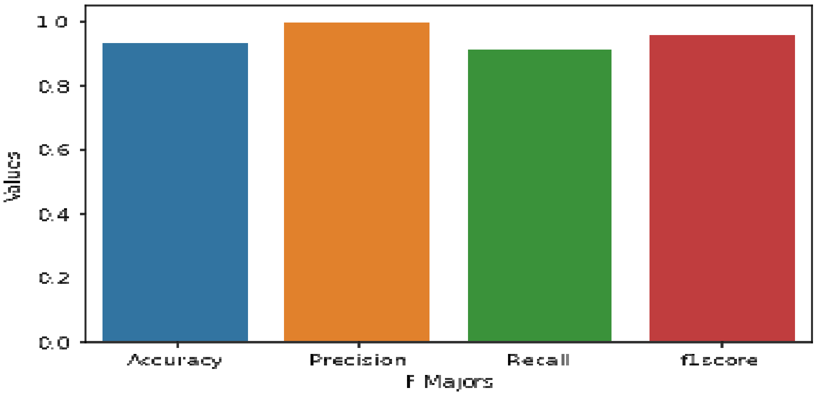


Fig 4.13

As we have seen the accuracy of all the algorithms in which Dnn-3 provide the better results. Now we will see the table wise comparison

Algorithms	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Linear Regression	*0.849*	*0.988*	*0.822*	*0.896*
Naïve Bayes	*0.928*	*0.987*	*0.924*	*0.953*
K Nearest Neighbor	*0.928*	*0.997*	*0.914*	*0.953*
Decision Tree	*0.927*	*0.998*	*0.911*	*0.952*
Adaboost	*0.926*	*0.996*	*0.912*	*0.950*
Random Forest	*0.927*	*0.998*	*0.911*	*0.951*
SVM rbf	*0.811*	*0.992*	*0.772*	*0.868*
SVM linear	*0.813*	*0.993*	*0.771*	*0.867*

DNN-1	*0.928*	*0.997*	*0.914*	*0.953*
DNN-2	*0.929*	*0.996*	*0.913*	*0.955*
DNN-3	*0.930*	*0.996*	*0.916*	*0.956*
DNN-4	*0.928*	*0.998*	*0.914*	*0.953*
DNN-5	*0.927*	*0.999*	*0.912*	*0.954*

Table 4.1

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

So far we have seen broadly the usefulness of DNN in intrusion detection systems. For confirming the results of DNN is the best, several classical machine learning algorithms were taken into account for comparing the results of DNN with them. The KDDCUP-‘99’ dataset which is publically available has been used as primary source of study and benchmarking, which can lead to show the superiority of Deep neural networks as compared to other classical machine learning algorithm. For further clarity the approach goes in deeper of in terms of hidden layers count which concluded that DNN with 3 layers gives best accuracy.

As it can be seen from the results, we can DNN technique is good and favorable path for security and intrusion tasks. The performance on artificial datasets is commendable. But we have to see the performance of the same on more complex and recent types of attacks on network traffic which is real time traffic. In addition of this there can be field like flexibility of DNN in opposed environments is required. So increase in approaches of deep learning techniques also makes curiosity to evaluate the overall results of these algorithms in order to achieve more accuracy towards IDs. So this can be a direction in which IDS research can travel which can be a future work.

REFERENCES

[1] PeymanKabiri and Ali A.Ghorbani-“Research on Intrusion Detection and Response Survey”- International Journal of Network Security, Vol.1, No.2, PP.84–102, Sep. 2005.

[2] Christopher Low –“Understanding Wireless attacks &detection “-GIAC Security Essentials Certification (GSEC) Practical Assignment 13 April 2005 -SANS Institute InfoSec Reading Room.

[3] M. Latah and L. Toker, "Towards an efficient anomaly-based intrusion detection for software-defined networks," in *IET Networks*, vol. 7, no. 6, pp. 453-459, 11 2018.

[4] A. Borkar, A. Donode and A. Kumari, "A survey on Intrusion Detection System (IDS) and Internal Intrusion Detection and protection system (IIDPS)," *2017 International Conference on Inventive Computing and Informatics (ICICI)*, Coimbatore, 2017, pp. 949-953.

[5] Bace, Rebecca-“An Introduction to Intrusion Detection &Assessment”- Infidel, Inc. for ICSA, Inc.

[6] Rebecca Gurley Bace-“Intrusion Detection”- Macmillan Technical Publishing, 2000.

[7] Denning, Dorothy E. – “An Intrusion Detection Model”- Proceedings of the Seventh IEEE Symposium on Security and Privacy May 1986.

[8] F. Sabahi and A. Movaghar, "Intrusion Detection: A Survey," *2008 Third International Conference on Systems and Networks Communications*, Sliema, 2008, pp. 23-26.

[9] “Global Information Assurance Certification Paper”- Copyright SANS Institute Copyright SANS Institute Author Retains Full Rights” 2014.

[10] “SANS penetration testing copyright by SANS”-Copyright SANS Institute Author Retains Full Rights 2014.

[11] Karthikeyan .K.R and A. Indra- “Intrusion Detection Tools and Techniques a Survey.

[12] F. Sabahi and A. Movaghar, "Intrusion Detection: A Survey," *2008 Third International Conference on Systems and Networks Communications*, Sliema, 2008, pp. 23-26.

[13] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," in *IEEE Access*, vol. 7, pp. 41525-41550, 2019.

[14] A. K. Saxena, S. Sinha and P. Shukla, "General study of intrusion detection system and survey of agent based intrusion detection system," *2017 International Conference on Computing, Communication and Automation (ICCCA)*, Greater Noida, 2017, pp. 471-421.

[15] A. Gül and E. Adalı, "A feature selection algorithm for IDS," *2017 International Conference on Computer Science and Engineering (UBMK)*, Antalya, 2017, pp. 816-820, doi: 10.1109/UBMK.2017.8093538.

[16] K. Zhao, M. Zhang, K. Yang and L. Hu, "Data Collection for Intrusion Detection System Based on Stratified Random Sampling," *2007 IEEE International Conference on Networking, Sensing and Control*, London, 2007, pp. 852-855, doi: 10.1109/ICNSC.2007.372892.

[17] Sriram Sundar Rajan, Vijaya Krishna Cherukuri-"An Overview of Intrusion Detection Systems 2012.

[18] John McHugh, Alan Christie, and Julia Allen- "The Role of Intrusion Detection Systems"- Software Engineering Institute, CERT Coordination Center, pp. 29041-29053, 2018.

[19] Wang, Ke (2004). "Anomalous Payload-Based Network Intrusion Detection" . *Recent Advances in Intrusion Detection*. Lecture Notes in Computer Science. Springer Berlin. 3224: 203–222. doi:10.1007/978-3-540-30143-1_11. ISBN 978-3-540-23123-3. Archived from the original .on 2010-06-22. Retrieved 2011-04-22.

[20] Khalkhali, I; Azmi, R; Azimpour-Kivi, M; Khansari, M. "Host-based web anomaly intrusion detection system, an artificial immune system approach". *ProQuest* .2011

[21] S. Zavrak and M. İskefiyeli, "Anomaly-Based Intrusion Detection From Network Flow Features Using Variational Autoencoder," in *IEEE Access*, vol. 8, pp. 108346-108358, 2020, doi: 10.1109/ACCESS.2020.3001350.

[22] A. Taylor, N. Japkowicz and S. Leblanc, "Frequency-based anomaly detection for the automotive CAN bus," *2015 World Congress on Industrial Control Systems Security (WCICSS)*, London, 2015, pp. 45-49, doi: 10.1109/WCICSS.2015.7420322

[23] Anita K. Jones and Robert S. Sielken –"Computer System Intrusion Detection A Survey "International journal of Computer Theory and Engineering, Vol.2, No.6, December, 2010.

[24] Sriram Sundar Rajan, Vijaya Krishna Cherukuri-"An Overview of Intrusion Detection Systems 2012.

[25] Carl Endorf, Eugene Schultz, Jim Mellander "Intrusion detection & prevention" by Written-published by McGraw-Hill.

[26] A. L. Blum and P. Langley, ``Selection of relevant features and examples in machine learning," in *Proc. AAAI Fall Symp. Relevance*, 1994, pp. 140_144.

[27] Piryonesi S. Madeh; El-Diraby Tamer E. (2020-06-01). "Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems". *Journal of Transportation Engineering, Part B: Pavements*. 146 (2): 04020022.

[28] R. Lippmann, J. Haines, D. Fried, J. Korba and K. Das. "The 1999 DARPA off-line intrusion detection evaluation". Computer networks, vol. 34, no. 4, pp. 579 595, 2000. DOI http://dx.doi.org/10.1016/S1389- 1286(00)00139-0.

[29] W. Lee and S. Stolfo. "A framework for constructing features and models for intrusion detectionsystems". ACM transactions on information and system security, vol. 3, no. 4, pp. 227261, 2000. DOI http://dx.doi. Org/10.1145/382912.382914.

[30] R. Agarwal and M. Joshi. "PNrule: A new framework for learning classier models in data mining". Tech. Rep. 00-015, Department of Computer Science, University of Minnesota, 2000.

[31] H. Kayacik, A. Zincir-Heywood and M. Heywood. "Selecting features for intrusion detection: A feature relevance analysison KDD 99 intrusion detection datasets". In Proceedings of the third annual conference on privacy, security and trust (PST). 2005.

[32] Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017, September). Applying convolutional neural network for network intrusion detection. In Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on (pp. 1222-1228).IEEE.

[33] Staudemeyer, R. C. (2015). Applying long short-term memory recurrent neural networks to intrusion detection. South African Computer Journal, 56(1), 136-154.

[34] Bengio, Y., Simard, P. and Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. IEEE transactions on neural networks, 5(2), pp.157-166.

[35] Maas, A.L., Hannun, A.Y. and Ng, A.Y., 2013, June. Rectifier nonlinearities improve neural network acoustic models. In Proc. icml (Vol. 30, No. 1, p. 3).

[36] F. Chollet, ”Keras (2015),” URL http://keras. Io, 2017.

[37] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard et al., ”Tensorflow: A system for large-scale machine learning.” in OSDI, vol. 16, 2016, pp. 265283.

[38] Stolfo, S., Fan, W. and Lee, W., KDD-CUP-99 Task Description. 1999- 10-28)[2009-05-08]. http://KDD. ics. uci. du/databases/kddcup99/task, html.

[39] J. McHugh. ”Testing intrusion detection systems: A critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory”. ACM transactions on information and system security, vol. 3, no. 4, pp. 262294, 2000. DOI http://dx.doi.org/10.1145/382912.382923.

[40] Sabhnani, Maheshkumar, and Gursel Serpen. ”Why machine learning algorithms fail in misuse detection on KDD intrusion detection data set.” Intelligent Data Analysis 8, no. 4 (2004): 403-415.

[41] Tavallaee, Mahbod, Ebrahim Bagheri, Wei Lu, and Ali-A. Ghorbani. ”A detailed analysis of the KDD CUP 99 data set.” In Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009. 2009.

[42] Moustafa, Nour, and Jill Slay. ”The evaluation of Network Anomaly Detection Systems: Statistical analysis of the U[8] H. Kayacik, A. Zincir- Heywood and M. Heywood. ”Selecting features for intrusion detection: A feature relevance analysison KDD 99 intrusion detection datasets”. In Proceedings of the third annual conference on privacy, security and trust (PST). 2005.

[43] Moustafa, Nour, and Jill Slay. ”UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data).”Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015.

LIST OF PUBLCATIONS OF THE CANDDATE’S WORK

1. Raj Kishore, and Anamika Chauhan. "Intrusion Detection System a need" In 2020 IEEE International Conference for Innovation in Technology, INOCON2020.
2. Raj Kishore and Anamika Chauhan. “Intrusion Detection using Fuzzy and deep neural Networks” In 2020 IEEE International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020. (Waiting for the acceptance).