

Analysis of Ensemble and Hybrid Approaches for Intrusion Detection

*final project report submitted in partial fulfillment of the requirements
for*

Bachelor of Technology
project

by

Abhinav Jaiswal (2016IPG-004)
Mohit Kumar (2016IPG-053)
Surendra Singh Gangwar (2016IPG-107)



विश्वजीवनामृतं ज्ञानम्

**ABV INDIAN INSTITUTE OF INFORMATION
TECHNOLOGY AND MANAGEMENT
GWALIOR-474 010**

2019

CANDIDATES DECLARATION

We hereby guarantee that the work, which is being introduced in the report, entitled as **Analysis of ensemble and hybrid approaches for intrusion detection**, in halfway fulfillment of the prerequisite for the honor of the Degree of **Bachelor of Technology** and our very own genuine record work completed during the period *May 2019* to *September 2019* under the supervision of **Dr. Saumya Bhadauria** and **Dr. Santosh Singh Rathore**. We tend to additionally cited the reference regarding the text(s)/figure(s)/table(s) from wherever they need been taken.

Date:

Signatures of the Candidates

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Date:

Signatures of the Research Supervisors

ABSTRACT

Network security is an important issue these days as the network usage is increasing effectively in multi-dimensions due to increase in the number of users. Our purpose is to analyze the various types of attacks and their parameters and build an efficient algorithm to detect these types of attacks. Our approach is based on the mixture of experts technique which takes the network of weighted classifiers and classify the type of attack based on its parameters. Four major attacks, either known in the literature or new, are studied in detail: Denial of service attack, User to root attack, Remote to local attack and probing attack. Taking parameters of the attack generated on the system, and then using dynamic weighting and appropriate classification algorithms for classification of attacks. Choosing the base classifiers as the input of the algorithms for the proposed model. The proposed system considers all the parameters required for the attack classification. It gives brief idea regarding the possible input values for which the classification of attacks can be performed effectively. The results show how the proposed model using mixture of experts technique helps in classification of attacks efficiently in the network system.

Keywords: Network security, Intrusion detection, Denial of service attack, User to root attack, Remote to local attack and probing attack, wireless ad hoc networks, mixture of experts.

ACKNOWLEDGEMENTS

We are extremely indebted to Dr. Saumya Bhadauria and Dr. Santosh Singh Rathore, and are duty-bound for giving us the autonomy of functioning and experimenting with ideas. We might wish to take this chance to specify our profound feeling to them not just for their educational steering however additionally for his or her personal interest in our project and constant support including confidence boosting and motivating sessions which proved terribly fruitful and were instrumental in infusing assurance and trust within us. The nurturing and maturation of the current work is especially because of their valuable guidance, suggestions, smart judgment, constructive criticism and an eye fixed for perfection. Our mentor continually answered myriad of our doubts with smiling graciousness and prodigious patience, their property feel that we tend to our novices by continually lending an ear to our views, appreciating and rising them and by giving us a freedom in our project. It's solely attributable to their overwhelming interest and useful angle, the present work has earned the stage its.

At long last, we are appreciative to our Institution and partners whose steady consolation served to recharge our soul, refocus our consideration and vitality and helped us in doing this work.

(Abhinav Jaiswal)

(Mohit Kumar)

(Surendra Singh Gangwar)

TABLE OF CONTENTS

ABSTRACT	ii
1 INTRODUCTION AND LITERATURE SURVEY	viii
1.1 Background	viii
1.2 Motivation	ix
1.3 Literature Survey	ix
1.3.1 Analysis of Machine learning approaches for intrusion detection	ix
1.3.2 Ensemble Approaches	ix
1.3.3 Mixture of Experts	x
1.4 Research gaps	x
1.5 Problem statement	x
1.6 Thesis Objective	x
2 DESIGN DETAILS AND IMPLEMENTATION	xi
2.1 Methodology	xi
2.1.1 Base Classifiers	xii
2.1.2 Ensemble Approaches:	xiii
2.1.2.1 Bagging:	xiii
2.1.2.2 Boosting:	xiv
2.2 Mixture of Experts	xvi
2.2.1 Architecture of MoE	xvi
2.3 Experimental Setup	xvii
3 DATASET DESCRIPTION	xviii
3.1 Dataset	xviii
3.1.1 Four categories of Attacks	xviii
3.1.2 Features Categories	xix
3.1.3 Attack Distribution	xix
4 Results and Analysis	xxi
4.1 Base Learner Performance	xxi

4.2	Ensemble Approaches	xxii
4.2.1	AUC-ROC curve of Bagging	xxiii
4.2.2	AUC-ROC curve of Adaptive Boosting	xxiv
4.2.3	AUC-ROC curve of Gradient Boosting	xxiv
4.3	Mixture of Experts	xxv
4.3.1	AUC-ROC curve of MOE	xxvi
4.4	Web interface for Intrusion Detection	xxvii
5	CONCLUSION AND FUTURE WORK	xxix
5.1	Conclusion	xxix
5.2	Future Activities	xxix
	REFERENCES	xxix

LIST OF FIGURES

2.1	Methodology	xi
2.2	Bagging , adopted from [8]	xiv
2.3	Boosting , adopted from [8]	xv
2.4	Mixture of Expert Architecture	xvii
3.1	Attack Distribution	xix
4.1	Comparison plot of Base Learners	xxii
4.2	Comparison plot of Ensembles	xxiii
4.3	AUC-ROC of Bagging	xxiii
4.4	AUC-ROC of Adaptive Boosting	xxiv
4.5	AUC-ROC of Gradient Boosting	xxiv
4.6	Comparison plot of MOE	xxv
4.7	AUC-ROC curve of MOE	xxvi
4.8	Home Page	xxvii
4.9	Dataset Page	xxvii
4.10	Prediction Page	xxviii

ABBREVIATIONS

ML	Machine learning
IDS	Intrusion Detection Systems
SVM	Support Vector Machine
LR	Logistic Regression
KNN	K-Nearest Neighbors
SOM	Smallest of Maximum
ANN	Artificial Neural Network
MoE	Mixture of Experts
DoS	Denial of Service
R2L	Remote to Local
U2R	User to Root
ROC	Receiver Operating Characteristic Curve

CHAPTER 1

INTRODUCTION AND LITERATURE SURVEY

1.1 Background

In recent years, there has been a revolutionary change in the field of networking, use of the internet is growing day by day and so is the risk of intrusion. Nowadays, it is very crucial to have a secure network because of increasing dependability on the internet and to achieve high-level security [1]. One of the possible ways is through the intrusion detection. Intrusion detection is an efficient technique, which aims to identify various network attacks and further classify these network attacks into several categories. Presently, the web is broadly utilized in various organizations and the business prerequisites have caused enterprises to convey their very own data frameworks on the web [2]. Due to large dependability on an information system, it became a major bottleneck, and thus the primary target for the intruders.

Nowadays, with the evolution of new advances like cloud computing, Big data, Internet of things, block chain, etc., which generates the enormous amount of network traffic, it becomes almost impossible to analyze the data and detect the intrusion in time. With the advancement in the machine learning (ML) domain, ML techniques enabled an intrusion detection system (IDS) results in less error rate and more accurate to classify various attacks in less possible time by eliminating the data which are of no use [3].

In this work, we investigate and evaluate two different hybrid approaches, ensemble methods and mixture of experts, which take advantage of various machine learners with the aim to give higher and accurate prediction performance that can not be achieved by using a single classifier .

1.2 Motivation

The methods used for intrusion detection have proven to be advantageous but classification of different intrusion attack type efficiently is still a major concern. Previous work performed by Chia-Ying Lin in year 2009 [1] includes classification of intrusions using different machine learning algorithms.

The ability of different machine learning algorithms to accurately detect a particular attack type, drives a motivation to use them to build a hybrid model which is capable of predicting all attack type precisely with low bias and low variance.

The methodology proposed by S. Reza [3] in 2014 describes how Mixture-of-experts technique can be used with various experts that can out-stand many single machine learning algorithm for multiclass classification.

1.3 Literature Survey

1.3.1 Analysis of Machine learning approaches for intrusion detection

Many different types of machine learning approaches has been implemented to classify the intrusion, some has used single learning techniques such as Support Vector Machine(SVM) ,Logistic Regression(LR) , K-nearest Neighbors(KNN), Artificial Neural Network (ANN) and some uses multilevel techniques in which they first use some nature inspired algorithm and genetic algorithm to do feature selection [11] and then used single classifier to predict the result, enormous works has been done in the area of classification of intrusion detection but most of them are only able to binary classify the attacks into normal or abnormal. Only few work has been done to classify the attack type with less accuracy rate. It is very important to classify the attack type so that proper measure can be taken to built a defence against it. It is important to quickly identify the intrusion at runtime which can only be possible if we identify the most important features of dataset and reduce it's dimensions to quickly train and predict from your model [13].

1.3.2 Ensemble Approaches

Ensemble learning is an effective technique to increase classification and prediction accuracy. It combines several machine learning algorithms (stacking) or uses many algorithms of same type (boosting , bagging) into one model in order to reduce loss and variance. Author Jie Gu [3] proposed a framework for intrusion detection which uses SVM ensemble with feature enhancement that gives the robust performance with high

precision and accuracy than any existing model only for binary classification.

1.3.3 Mixture of Experts

Mixture of experts works on the principle that every expert is specialized in a particular domain of a input data space by imitating a gating network which is liable for learning the combined weight of the specialized experts for a particular input [11] ,by this method input space is dynamically divided and conquered by the gating network and the experts.

1.4 Research gaps

The past work done in the field of intrusion detection predominantly centers around binary classification, detecting a specific intrusion type effectively is as yet a note worthy concern[16]. Additionally, there is a need of consistent dataset so that there would not be any biasness in order to predict any attack type. The proposed work uses different machine learning algorithm to detect a particular attack type in which they give the best result, which can be used as hybrid method using mixture of expert, and also using different feature selection algorithm to reduce the features which are of less significance for the predicating the different attack type can be useful for increasing the efficiency of our model.

1.5 Problem statement

Comparative Analysis of previously built intrusion detection algorithms with hybrid Mixture of expert algorithm which combines the expertise of different machine learning models to built a meta-model that is able to provide a result which would out-perform many intrusion algorithms previously developed.

1.6 Thesis Objective

1. Evaluating the performance of many different machine learning algorithms for the intrusion detection.
2. Implementation of mixture of experts technique for intrusion detection.
3. To perform a comparative analysis of ensemble approaches(Bagging,Boosting) with hybrid approach(Mixture of experts).

CHAPTER 2

DESIGN DETAILS AND IMPLEMENTATION

2.1 Methodology

Figure 2.1 describes the methodology comprising of subjecting input dataset to data preprocessing, feature selection, selection of base learners for finding the experts and finally analyzing the output.

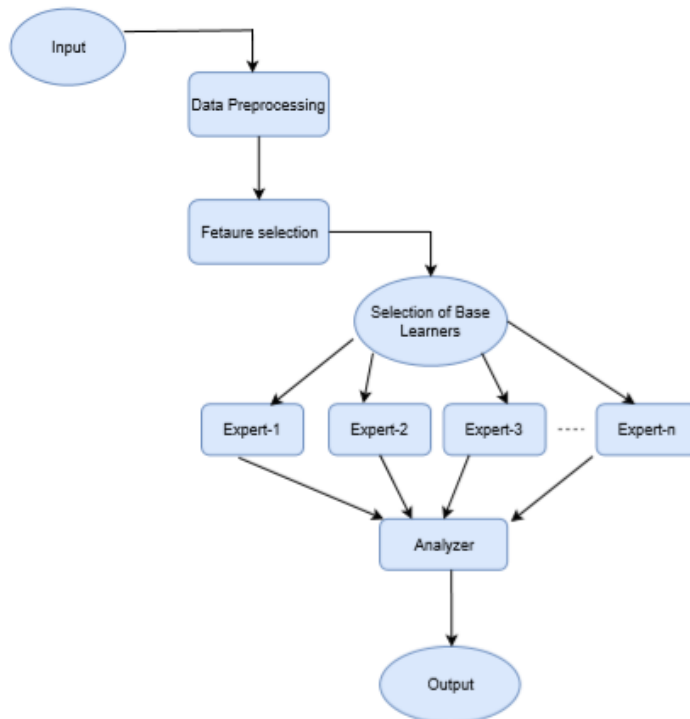


Figure 2.1: Methodology

Our hybrid model is composed of two stages, feature extraction and classification. Feature extraction consists of removing of the data redundancy and filling the missing val-

ues. Data redundancy is the large amount of similar content in the dataset. Redundant records are removed to maintain the data redundancy. Missing values are filled by either by mean, mode and median in order to get the highest accuracy. The next part of feature extraction includes the encoding of the data which is performed by using label encoding and one-hot encoding techniques of scikit-learn library of python. All the categorical text data is converted into numerical data and all the categorical variables are converted into model understandable form.

2.1.1 Base Classifiers

Analyzing and implementing the base classifiers to find the best classifiers needed for the proposed model.

- **K-nearest neighbor**

It is a supervised machine learning algorithm that stores all the objects available in the dataset and classifies the new object or data according to the similarity measure. K represents the count of the nearest neighbor [4]. The classification of the new data or object is purely done by the vote measure (e.g. euclidean distance, overlap matrix (hamming distance)). Values of k for which the algorithm gives the maximum result can be selected by different heuristic techniques.

$$d(x, y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2.1)$$

- **Naive bayes**

Naive Bayes classifier is a conditional probabilistic machine learning model. The essence of the classifier is based on the Bayes theorem. With the help of Bayes theorem, we can find the probability of happening of event A , given that event B has occurred. The reason behind the word naive is that it is to be assumed that the features are independent [6], i.e., the presence of one predictor/features does not affect other. There are different types of naive Bayes classifier namely multinomial naive Bayes, Bernoulli naive Bayes and Gaussian naive Bayes. For deciding the best base learner for the proposed model, we use the multinomial naive Bayes.

- **Support vector classifier**

It is an efficient machine learning algorithm which can perform linear as well as non-linear classification using the perimeter called Kernel trick, which maps the input data into high dimensional feature spaces [18]. The parameter kernel helps in non-linear classification.

- **Random Forest classifier**

Random forest classifier is an ensemble learning method for classification which control overfitting and uses averaging to improve the predictive accuracy. The classifier operates by constructing a multitude of individual decision trees at training time and outputting the class which has the maximum vote count [17]. Various random decision forests present correct tree's habit of overfitting to their training set.

2.1.2 Ensemble Approaches:

Ensemble learning is an effective technique to increase classification and prediction accuracy. It combines several machine learning algorithms (stacking) or uses many algorithms of same type (boosting, bagging) into one model in order to reduce loss and variance. Jie Gu [3] proposed a framework for intrusion detection which uses SVM ensemble with feature enhancement that gives the robust performance with high precision and accuracy than any existing model only for binary classification.

2.1.2.1 Bagging:

This ensemble approach comprises of two steps, first it randomly selects bootstrapped samples from the given dataset and build a classifier for each bootstrap sample and then combines the results from all classifiers [12]. For combining the results it may use different methods like voting, mean and creates a powerful classifier. It used bootstrap samples so it is also known as bootstrap aggregation. Bootstrap aggregation method performs the following steps:

- Random selection of the sub-samples of dataset. (one value can be selected multiple times).
- Calculation of the mean of each sub-samples.
- Calculating the average value of means of all bootstrapped samples to predict the result.

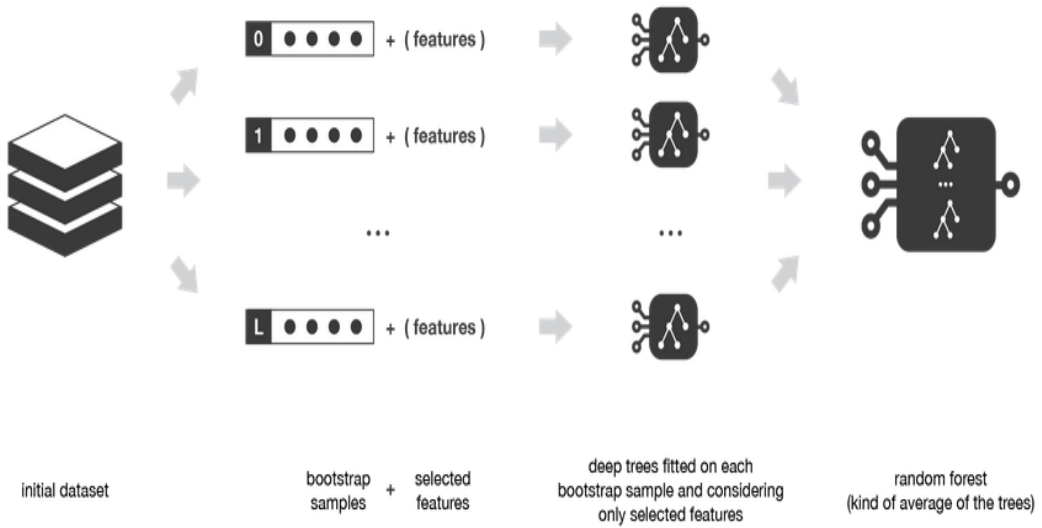


Figure 2.2: Bagging , adopted from [8]

Assume we have L bootstrapped samples and each of them having size B .

$$\{z_1^1, z_2^1, \dots, z_B^1\}, \{z_1^2, z_2^2, \dots, z_B^2\}, \dots, \{z_1^L, z_2^L, \dots, z_B^L\} \quad (2.2)$$

$z_b^l \equiv b - th$ observation of $l - th$ bootstrap sample

Fitting L classifiers on different sub-samples and collecting their mean.

$$w_1(.) , w_2(.) , \dots, w_L(.) \quad (2.3)$$

Aggregating the results of all these classifiers using an averaging method to get the final model. Resulted model can be defined as

$$s_l(.) = \frac{1}{L} \sum_{l=1}^L w_l(.) \quad (2.4)$$

(Simple average, for regression problem)

2.1.2.2 Boosting:

Boosting is a sequential learning algorithm which trains weak learners to convert them into the strong learner. In each iteration, it tries to increase the weights of poorly predicted Instances [2]. In each of its iteration we select some instances from the dataset

and train the model with these instances but it will poorly classify those instances which were not included in previous training so, In next iteration it may include those instances but it is sure for the strong learner all of instances of the database included at least once.

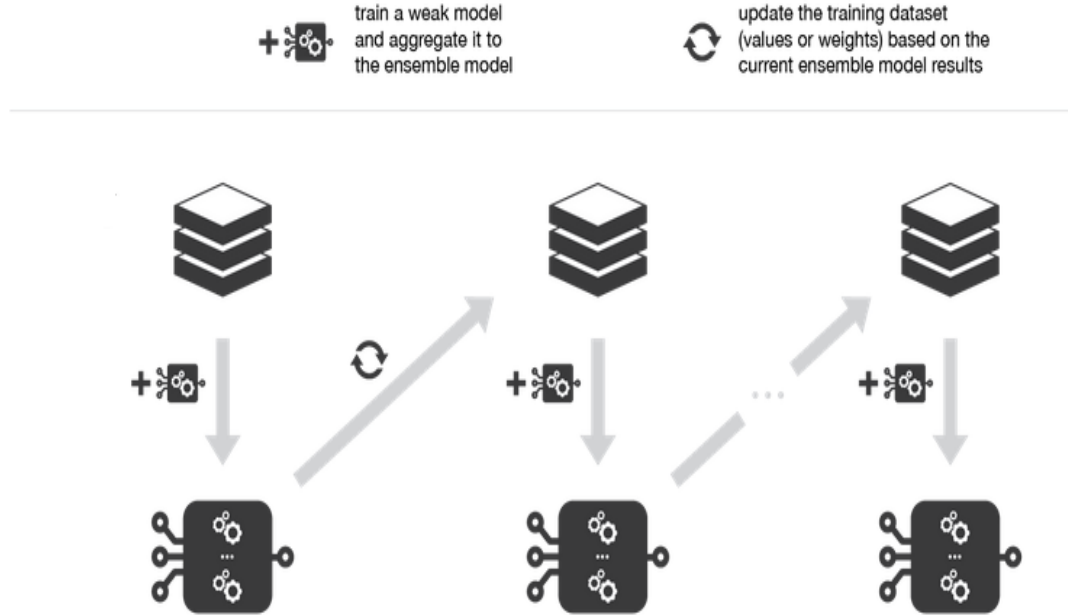


Figure 2.3: Boosting , adopted from [8]

As in Adaptive boosting algorithm, Final strong model is the weighted sum of all weak learners.

$$s_L(.) = \sum_{l=1}^L c_l \times w_l(.) \quad (2.5)$$

where c_l 's are coefficients and w_l 's are weak learners.

Instead of optimizing the final classifier we try to optimize the performance of local weak learners. It will automatically lead to build a strong classifier. In each iteration we transform the weak learner with strong learner. In other words, we can define a strong learner (s_l) over a weak learner ($s_{(l-1)}$) such that

$$s_l(.) = s_{(l-1)} + c_l \times w_l(.) \quad (2.6)$$

2.2 Mixture of Experts

Mixture of experts works on the principle that every expert is specialized in a particular domain of a input data space by imitating a gating network which is liable for learning the combined weight of the specialized experts for a particular input [11], by this method input space is dynamically divided and conquered by the gating network and the experts.

2.2.1 Architecture of MoE

Input to the architecture comprises of textual data with selected features using feature selection algorithm on the dataset, working of the model can be explained in the following steps:

- first all the selected experts predict the attack type for the corresponding tuple and after that to get the required result voting for the particular attack type is done [14].
- After voting if there exist a class which clearly outweigh all the class than that class is the required result if the probability of all the class to be of that particular attack type is equally probable than the output is taken from the experts which are expert in the particular attack type.
- All the predicted output from the experts are given the dynamic weight on the basis of that output whose probability is highest is accepted as the final output for the given tuple

The aim of the model is to correctly classify all the classes either by using the voting method or by using expertise of different experts to get the desired output using the above approach the probability of getting a wrong output minimizes and we get the high accuracy result.

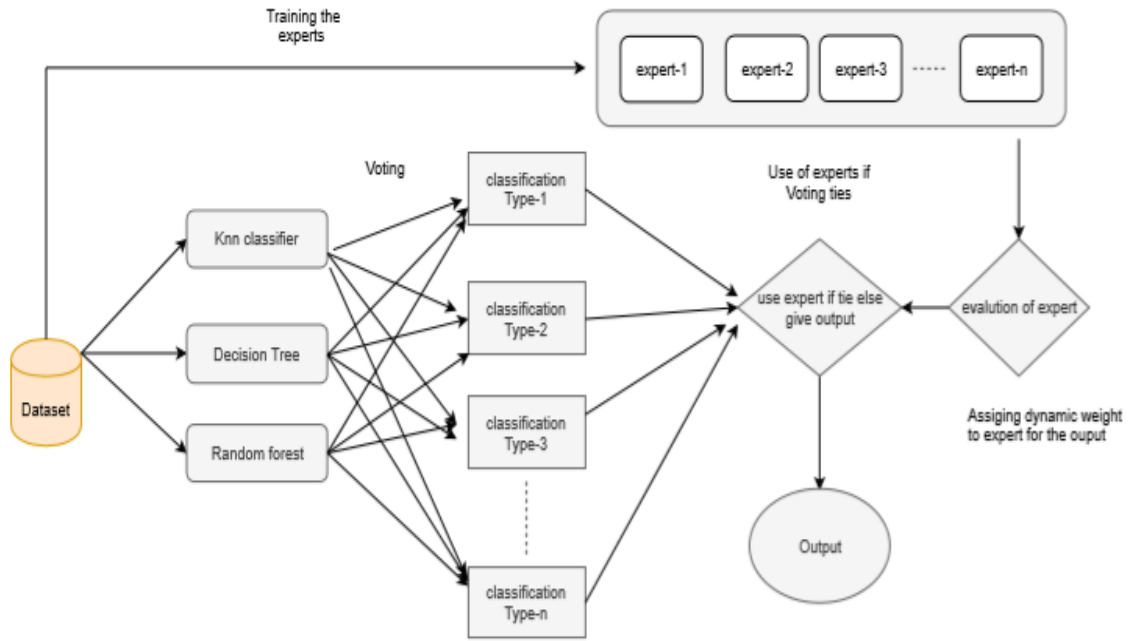


Figure 2.4: Mixture of Expert Architecture

2.3 Experimental Setup

All the models implementation is performed using the scikit-learn library. The model proposed in this paper is trained on a single Nvidia GeForce GTX 1080 GPU with 4GB of graphics memory.

CHAPTER 3

DATASET DESCRIPTION

3.1 Dataset

The proposed work utilizes KDD 1999 informational collection. The KDD 1999 dataset was created by the MIT Lincoln Labs and utilized for the Third International Knowledge Discovery and Data Mining Tools Competition [10]. The whole dataset is exceptionally enormous because of numerous autonomous highlights and requires a ton of calculation control. Our proposed work is utilizing 10 percent of the dataset for intrusion detection process. The KDD 99 comprises of numerous excess records which causes to anticipate the model to perform well. For averting these issues an extemporized variant of KDD 99 dataset was made, the NSL-KDD dataset [5]. In NSL-KDD target field is separated into four sorts of assault or an ordinary traffic.

3.1.1 Four categories of Attacks

1. **Denial of Service (DoS):** In dos attack, Intruder sends many requests to a server to exhausts all the resources of the server with the goal that real clients can't use resource and services. [9].
2. **Probe:**It is an attempt from intruders to gain access the information and vulnerabilities of the system, later on this information can be used to generate an attack.
3. **Remote to Local (R2L):**In this type of attack intruders try to gain unauthorized access to the system over the network by sending the packets to that system.
4. **User to Root (U2R):**By this type of attack intruders first access the system with normal privileges later on try to access the user with administrator privileges.

3.1.2 Features Categories

Training and testing sets both comprises of 41 highlights delegated typical traffic or explicit attack types[15]. These 41 parameters can be subdivided into four categories like Basic features, time sensitive traffic, content features and host-based traffic features [19].

1. **Basic features:** Protocol type, Service, duration and etc.
2. **Time-based traffic features:**count, srv count, Rerror rate and etc.
3. **Content features:** hot, num root, is guest login.
4. **Host-based traffic features:**dst host count, dst host srv count and etc.

3.1.3 Attack Distribution

1. Dataset contains a lot of irregularities like redundant records very low percentage of some attack types [7]. The percentage of attack type in the dataset has been pictorially shown by the pie chart.

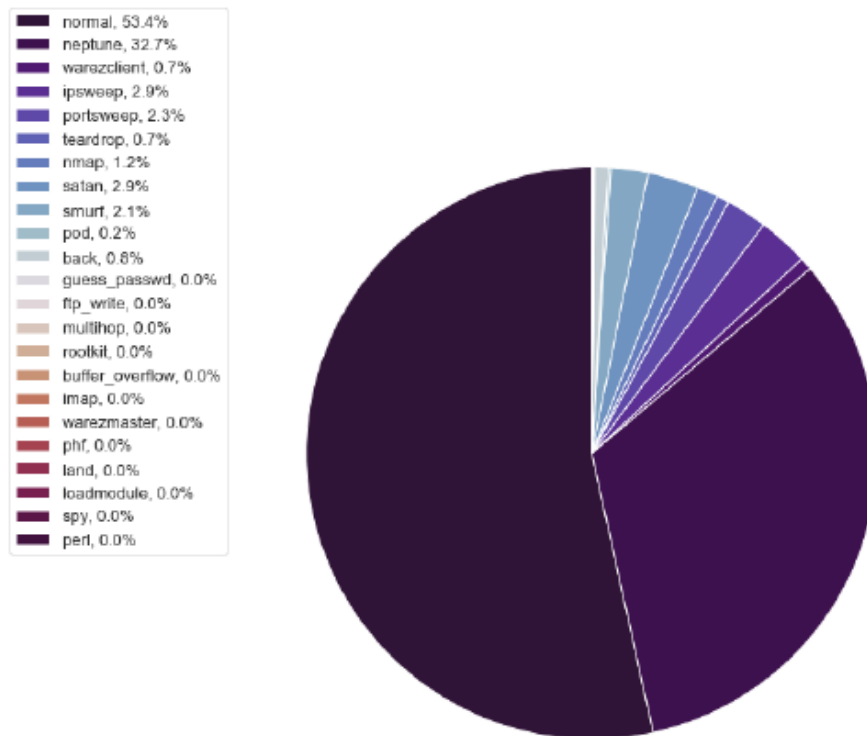


Figure 3.1: Attack Distribution

2. Attack Distribution in training dataset

Attack Classes	Instances of Attacks	Total Instances
Denial Of Service	back(956),land(18),neptune(41,214), pod(201),smurf(2,646),teardrop(892)	45,927
Probe	satan(3,633),ipsweep(3,599), nmap(1,493),portsweep(2,931)	11,656
R2L	guess-passwd(53),imap(658), phf(4),multihop(7),warezmaster(20), warezclient(890)	16,42
U2R	loadmodule(9),perl(3),bufferoverflow (30),rootkit(10)	52

Table 2.1 : Attack Distribution in training dataset

3. Attack Distribution in testing dataset

Attack Classes	Attack Types(instances)	Total Instances
Denial Of Service	back(359),apache2(737),smurf(665), pod(41),neptune(4657),teardrop(12), processtable(685),mailbomb(293)	74,60
Probe	satan(3,633),ipsweep(3,599),nmap (1,493),portsweep(2,931)	11,565
R2L	guesspasswd(1,231),imap(307),named (17),multihop(18),xlock(9),snmpguess (331),warezmaster(944),sendmail(14),	3191
U2R	Bufferoverflow(20),loadmodule(2), xterm(13),rootkit(13),perl(2),sqlattack (2),ps(15)	67

Table 2.2 : Attack Distribution in testing dataset

CHAPTER 4

Results and Analysis

As mentioned before we are using three different approaches on NSL-KDD dataset. Result part is divided into three sections (Base Learners, Ensemble, Mixture Of Experts) each containing performance table, a line graph and AUC-ROC curve for comparing the results [2].

Auc-roc (Area Under the Curve for Receiver Operating curve) is used to measure the performance for the classification at various thresholds settings [20], ROC represents the probability curve and AUC speaks to degree or proportion of separability. It tells how much model is equipped for recognizing classes. Higher the AUC, better the model is at foreseeing 0s as 0s and 1s as 1s. By similarity, Higher the AUC, better the model is at recognizing different attack type.

4.1 Base Learner Performance

In first approach, We have performed different base learners on the dataset in order to find the best in its type to classify attacks. Table 4.1 below describes the findings of different learners on the basis of accuracy, precision, recall, F1-score. Figure 4.1 is the line graph representation of performances of different base learners with Accuracy, Precision, Recall and F1-score.

Classifier	Accuracy	Precision	recall	F1-score
LinearSVC	0.9221	0.9229	0.9147	0.9187
KNN	0.9590	0.9699	0.9747	0.9674
MultinomialNB	0.4172	0.4161	0.4159	0.3966
RandomForest	0.9688	0.9679	0.9680	0.9745
LogisticRegression	0.8959	0.8912	0.8899	0.8457
DecisionTree	0.9490	0.9474	0.9447	0.9447

Table 4.1 : Base Learners Results

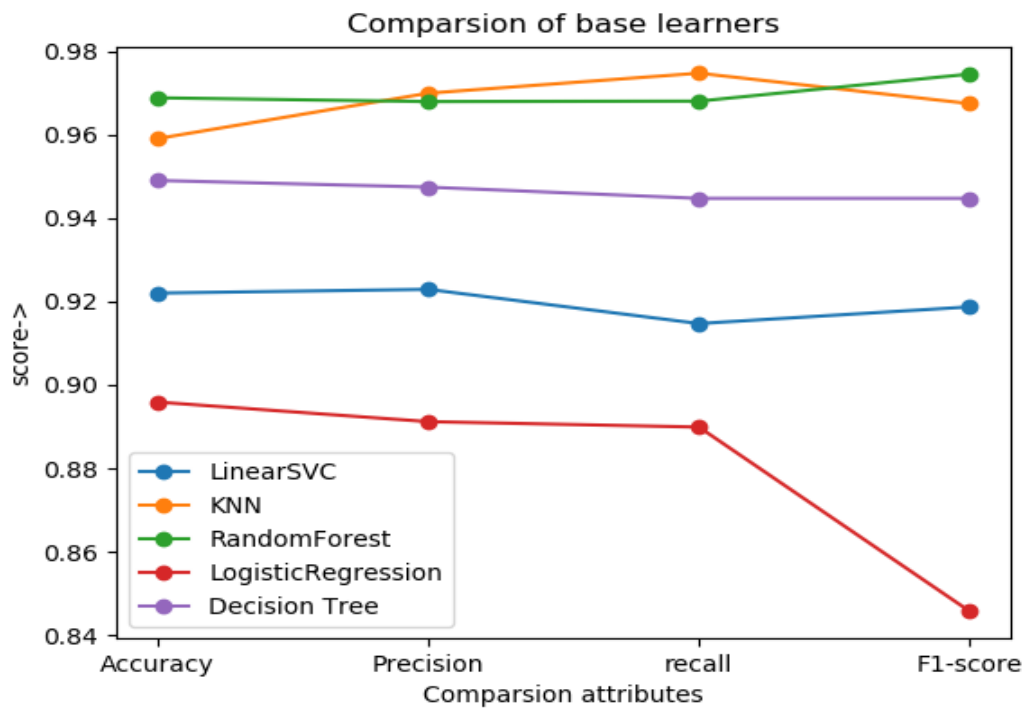


Figure 4.1: Comparison plot of Base Learners

4.2 Ensemble Approaches

In second part, Bagging and Boosting ensemble approaches are performed on dataset. Table 4.2 below shows the findings of ensemble approaches on the basis of accuracy, precision, recall, F1-score. Figure 4.2 represents comparison b/w bagging and boosting methods.

Classifier	Accuracy	Precision	recall	F1-score
Bagging	0.98947	0.98947	0.99847	0.98945
Adaptive Boosting	0.97954	0.95218	0.97654	0.97836
Gradient Boosting	0.98891	0.98894	0.98891	0.98892

Table 4.2 : Ensemble Results

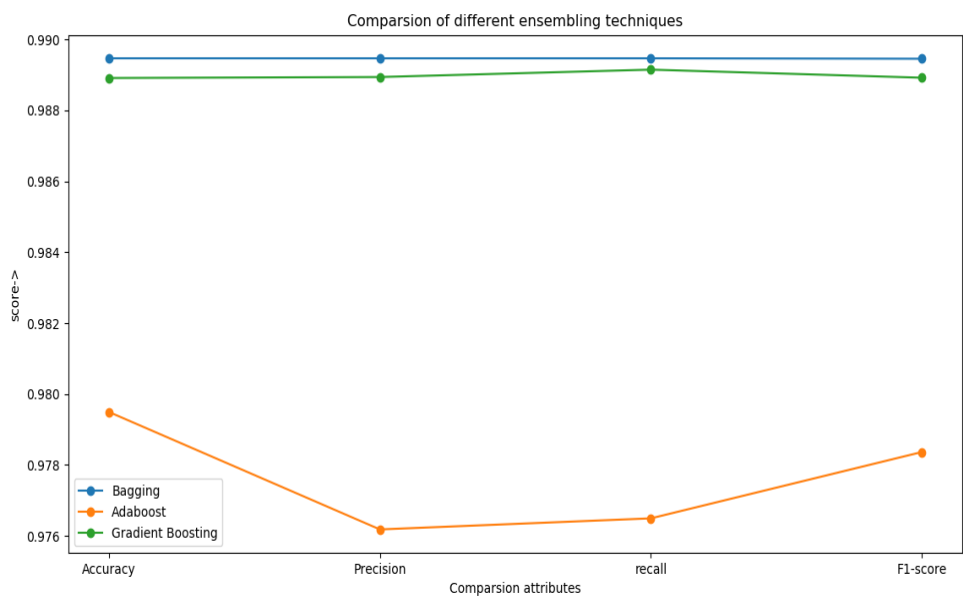


Figure 4.2: Comparison plot of Ensembles

4.2.1 AUC-ROC curve of Bagging

Below roc curve represents the performance of Bagging Ensemble for differentiating different classes.

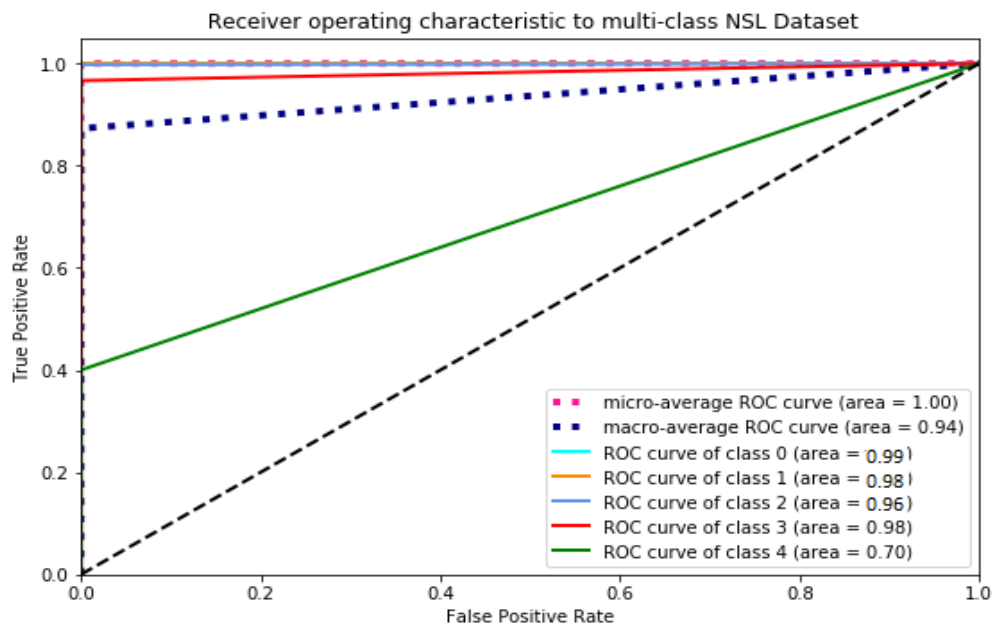


Figure 4.3: AUC-ROC of Bagging

4.2.2 AUC-ROC curve of Adaptive Boosting

Below roc curve represents the performance of Adaptive Boosting Ensemble for differentiating different classes.

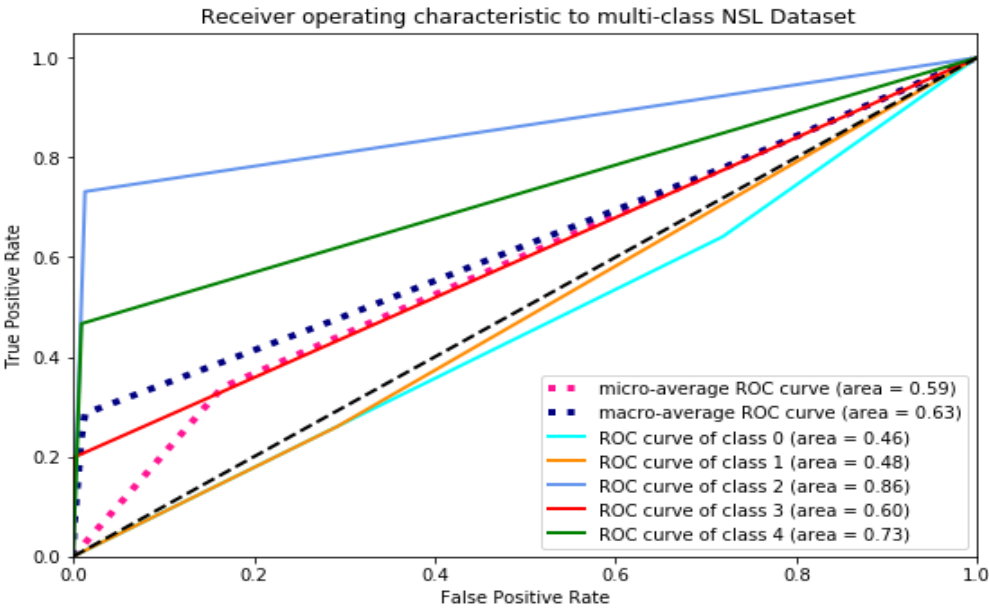


Figure 4.4: AUC-ROC of Adaptive Boosting

4.2.3 AUC-ROC curve of Gradient Boosting

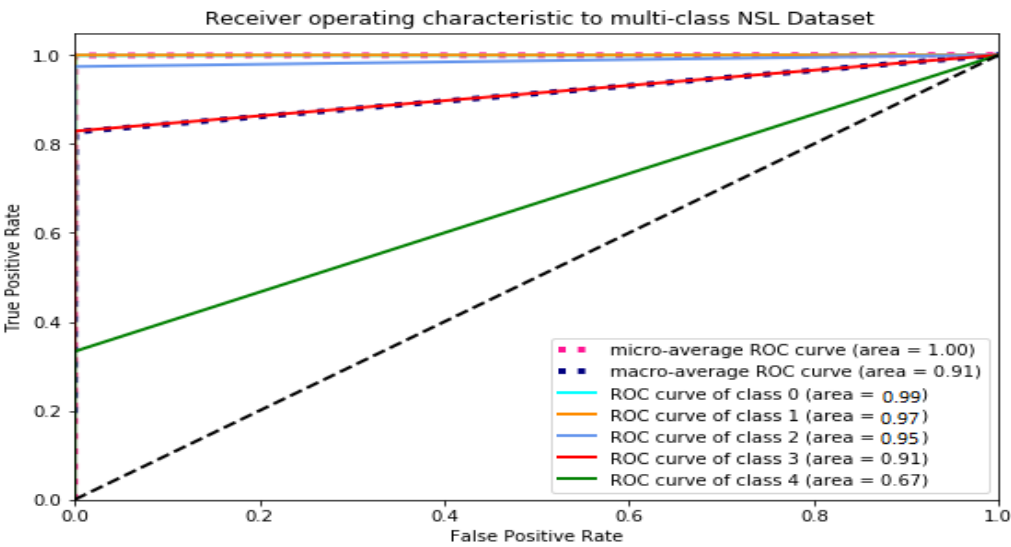


Figure 4.5: AUC-ROC of Gradient Boosting

Above roc curve represents the performance of Gradient Boosting Ensemble for differentiating different classes.

4.3 Mixture of Experts

Below are the findings after using MoE on the NSL-KDD dataset. Performance of proposed model is evaluated on the basics of Accuracy, precision, recall and F1-score. separability of MoE is shown with the help of AUC-ROC that signifies how well model differentiate the classes.

Classifier	Accuracy	Precision	recall	F1-score
MoE	0.99901	0.99745	0.99847	0.99987

Table 4.3 : MOE Results

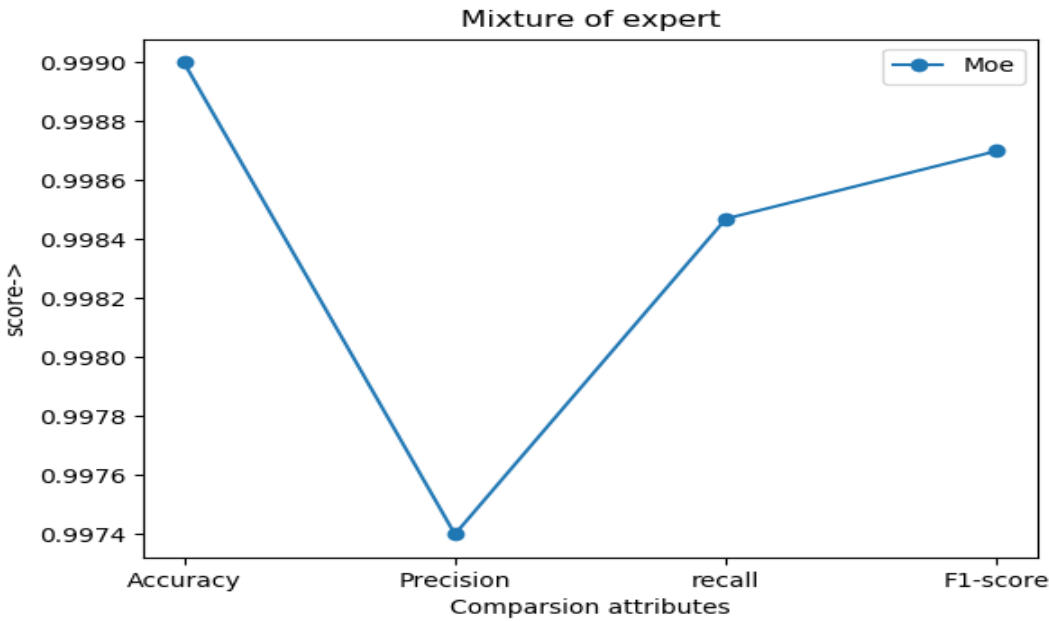


Figure 4.6: Comparison plot of MOE

4.3.1 AUC-ROC curve of MOE

Below roc curve represents how well MoE is capable of differentiating different classes in comparison of other approaches previously used.

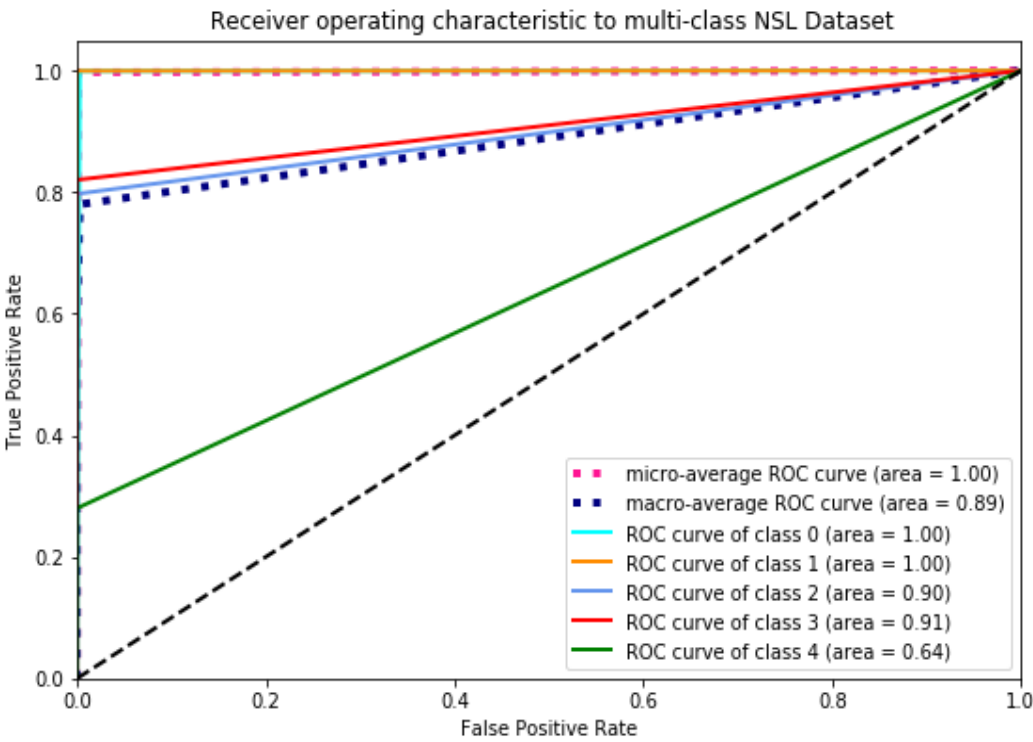


Figure 4.7: AUC-ROC curve of MOE

4.4 Web interface for Intrusion Detection

- A web interface is developed for the classification of attacks. The portal consists of multiple pages includes the current scenario of machine learning algorithms for intrusion detection and IDS.

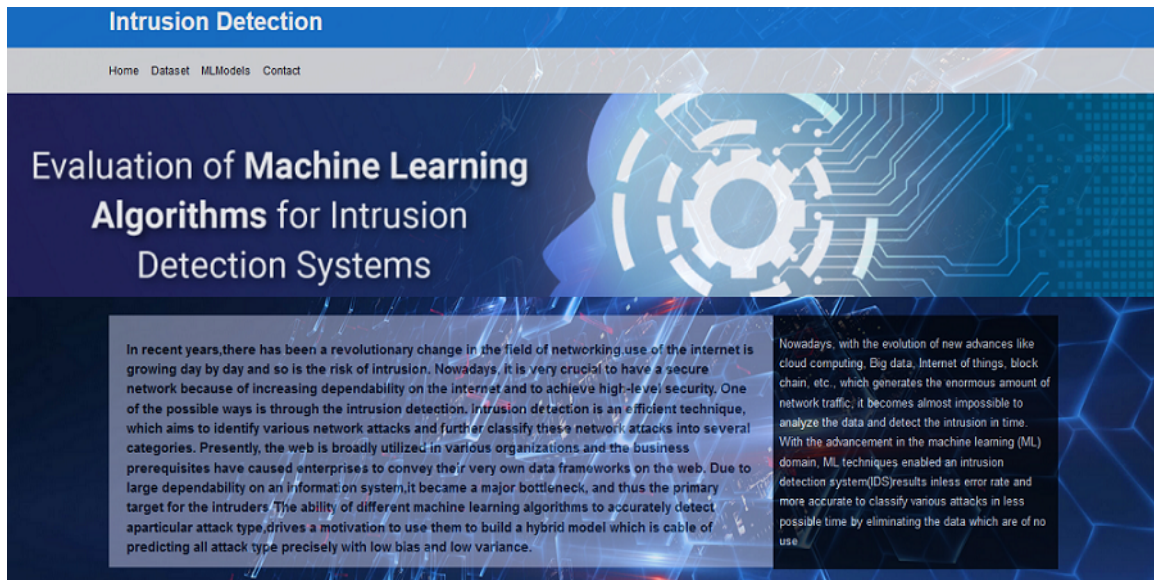


Figure 4.8: Home Page

- Next page describes the dataset description and at the separate page attacks will be classified into categories of Denial of service (DoS), Probe, Remote to Local (R2L) and User to root attack.

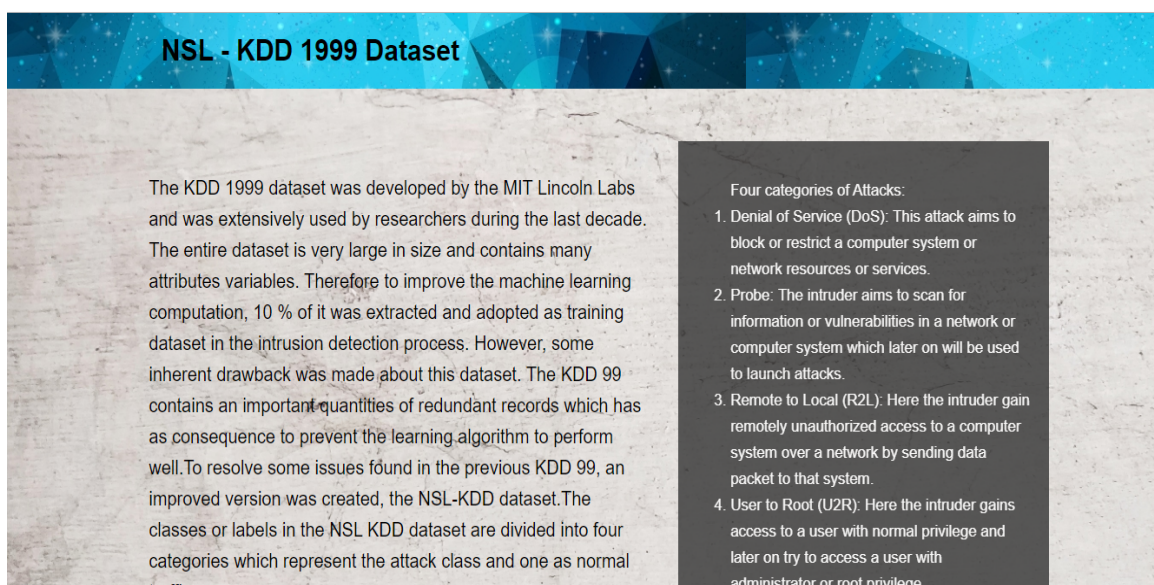


Figure 4.9: Dataset Page

- On the next page,the user has to give the parameters as the input to the four categories as basic features like protocol type, service and flag.The next category, time features includes count,server count and error rate.The thired category includes source bytes,num roots and guest login, and the last category includes count, host server count and destnation bytes.

Attack Type Classification

1 Basic Features

udp

shell

OTH

2 Time Features

2

3

5

3 Content Features

587

5

1

4 Host Features

2

35

958

submit

Prediction

Root_to_Local

Figure 4.10: Prediction Page

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this paper we have proposed the Mixture of expert approach for the intrusion detection and its comparison with the previous work. Our proposed model outclass the previous proposed methods by accurately classifying the attacks types by 99 % which can be clearly seen from the roc curve shown in result section.

5.2 Future Activities

1. Implementation of the proposed model with the large dataset, rich in wide variety of attacks.
2. Applications of mixture of experts technique has certainly exhibit tremendous results. Therefore, try to implementing this technique on other technologies.
3. Building a web based application to detect the type of attack generated dynamically on the system.

REFERENCES

- [1] K.Muneeswaran B.Selvakumar. "firefly algorithm based feature selection for network intrusion detection". *Science Direct*,, 81:148–155, 2018.
- [2] Peter Bühlmann. "bagging, boosting and ensemble methods". *Handbook of Computational Statistics*, 01 2012.
- [3] Chia-Ying Lin Wei-YangLin Chih-Fong Tsai, Yu-Feng Hsu. "intrusion detection by machine learning: A review". *Science Direct*,, 10:11994–12000, 2009.
- [4] Padraig Cunningham and Sarah Delany. "k-nearest neighbour classifiers". *Mult Classif Syst*, 04 2007.
- [5] Theodoros Evgeniou and Massimiliano Pontil. "support vector machines: Theory and applications". 2049:249–257, 01 2001.
- [6] Jie Gu, Lihong Wang, Huiwen Wang, and Shanshan Wang. "a novel approach to intrusion detection using svm ensemble with feature augmentation". *Computers Security*, 2019.
- [7] Y. P. N. L. S. S. L. C. Guo. "a two-level hybrid approach for intrusion detection". *Research Gate*,, 214:9, 2016.
- [8] joseph rocca. "ensemble-methods-bagging-boosting-and-stacking". *medium*, 2018.
- [9] Pouria Kaviani and Sunita Dhotre. "short survey on naive bayes algorithm". *International Journal of Advance Research in Computer Science and Management*, 04, 11 2017.
- [10] G. C. Kessler. "defenses against distributed denial of service attacks". *Science Direct*,, 321:12, 2002.
- [11] Euntai Kim, Heejin Lee, Minkee Park, and Mignon Park. "a subset feature elimination mechanism for intrusion detection system". *International Journal of Advanced Computer Science and Applications*,, 7:148–157, 2016.

- [12] Rohit Kumar Singh Gautam and Amit Doegar. "an ensemble approach for intrusion detection system using machine learning algorithms". *International Conference on Cloud Computing, Data Science Engineering*, 14:554 – 574, 2003.
- [13] W. Lu M. Tavallaei, E. Bagheri and A. Ghorbani. "a detailed analysis of the kdd cup 99 data set". *IEEE international conference on Computational intelligence for security and defense applications*, 1:53–58, 2009.
- [14] Kristina Machova, Miroslav Puszta, Frantisek Barcák, and Peter Bednár. "a comparison of the bagging and the boosting methods using the decision trees classifiers". *Comput. Sci. Inf. Syst.*, 3:57–72, 01 2006.
- [15] Arnu Pretorius, Surette Bierman, and Sarel Steel. "a meta-analysis of research in random forests for classification". pages 1–6, 11 2016.
- [16] Danijela Protic. Review of kdd cup '99, nsl-kdd and kyoto 2006+ datasets. *Vojnotehnicki glasnik*, 66:580–596, 07 2018.
- [17] Riyad.A.M and M.S Irfan Ahmed. "an ensemble classification approach for intrusion detection". *International Journal of Computer Applications*, 80:37–42, 2013.
- [18] A. Nowzari-Dalini M. Ganjtabesh-R. Ebrahimpour. S. R. Kheradpisheh, F. Sharifzadeh. "Mixture of feature specified experts", volume 20. 2014.
- [19] Mohit Tiwari, Raj Kumar, Akash Bharti, and Jai Kishan. "intrusion detection system". *International Journal of Technical Research and Applications*, 5:2320–8163, 2017.
- [20] Shengping Yang and Gilbert Berdine. "the receiver operating characteristic (roc) curve". *The Southwest Respiratory and Critical Care Chronicles*, 5:34, 05 2017.