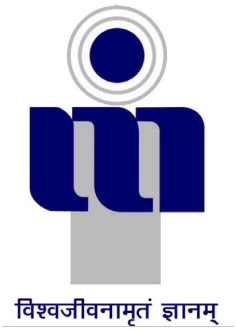
Analysis of Ensemble and Hybrid Approaches for Intrusion Detection

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)



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

2019

CANDIDATES DECLARATION

We thusly ensure that the work, which is being presented in the report, entitled as **Analysis of outfit and cross breed approaches for interruption detection**, in midway satisfaction of the essential for the respect of the Degree of **Bachelor of Technology** and our own one of a kind veritable record work finished during the period *May 2019* to *September 2019* under the supervision of **Dr. Saumya Bhadauria** and **Dr. Santosh Singh Rathore**.We tend to moreover refered to the reference with respect to the text(s)/figure(s)/table(s) from any place they need been taken.

Date: Signatures of the Candidates

This is to guarantee that the above proclamation made by the applicants is right supposedly.

Date: Signatures of the Research Supervisors

ABSTRACT

System security is a significant issue nowadays as the system utilization is expanding viably in multi-measurements because of increment in the quantity of clients.

Our motivation is to break down the different kinds of assaults and their parameters and assemble a proficient calculation to recognize these sorts of assaults. Our methodology depends on the blend of specialists system which takes the system of weighted classifiers and group the sort of assault dependent on its parameters.Four real assaults, either known

in the writing or new, are contemplated in detail: Denial of administration assault, User to attach attack,Remote to neighborhood assault and examining assault.

Taking parameters of the assault created on the framework ,and after that utilizing dynamic weighting and proper order calculations for grouping of assaults.

Picking the base classifiers as the contribution of the calculations for the proposed model . The proposed

framework considers every one of the parameters required for the assault characterization. It gives brief thought with respect to the conceivable information esteems for

which the characterization of assaults can be performed successfully. The outcomes show how the

proposed model utilizing blend of specialists method helps in arrangement of assaults productively in the system framework .

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 incredibly obligated to Dr. Saumya Bhadauria and Dr.Santosh Singh Rathore,and are compelled by a solemn obligation for giving us the self-rule of working and trying different things with thoughts. we may wish to take this risk to explicit our significant inclination to them not only for their instructive directing anyway also for their own enthusiasm for our undertaking and steady help including certainty boosting and rousing sessions which demonstrated frightfully productive and were instrumental in mixing confirmation and trust inside us. The sustaining and development of the present work is particularly a result of their significant direction, proposals, shrewd judgment, useful analysis and an eye fixed for flawlessness. Our guide constantly addressed heap of our questions with grinning charitableness and immense patience,their property feel that we keep an eye on our amateurs by consistently listening carefully to our perspectives, acknowledging and rising them and by giving us an opportunity in our undertaking. It's exclusively owing to their mind-boggling premium and helpful edge, the present work has earned the stage its.

Finally, we are grateful to our Institution and accomplices whose consistent relief served to revive our spirit, refocus our thought and imperativeness and helped us in doing this work.

(Abhinav Jaiswal) (Mohit Kumar) (Surendra Singh Gangwar)

TABLE OF CONTENTS

LIST OF FIGURES

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 cable 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 been 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

Blend of specialists deals with the rule that each master is had practical experience in a specific area of an information space by mimicking a gating system which is at risk for learning the joined load of the particular specialists for a specific information [11] ,by this technique info space is powerfully isolated and vanquished by the gating system and the specialists.

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, their is a need of consistent dataset so that their 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.

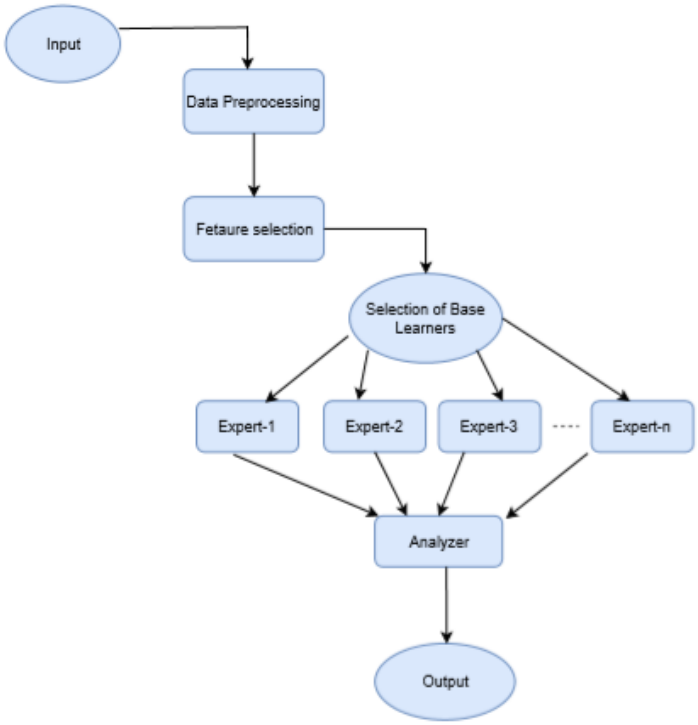


Figure 2.1: Methodology

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 is 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 labelencoding and onehotencoding techniques of scikit-learn library of python.All the categorical text data is converted into numerical data and all the categorical variables are converted 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 classify the new object or data according to the similarity measure.K represent the 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.

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

- **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 occoured.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 predictors/features does not effect 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 proposing 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.

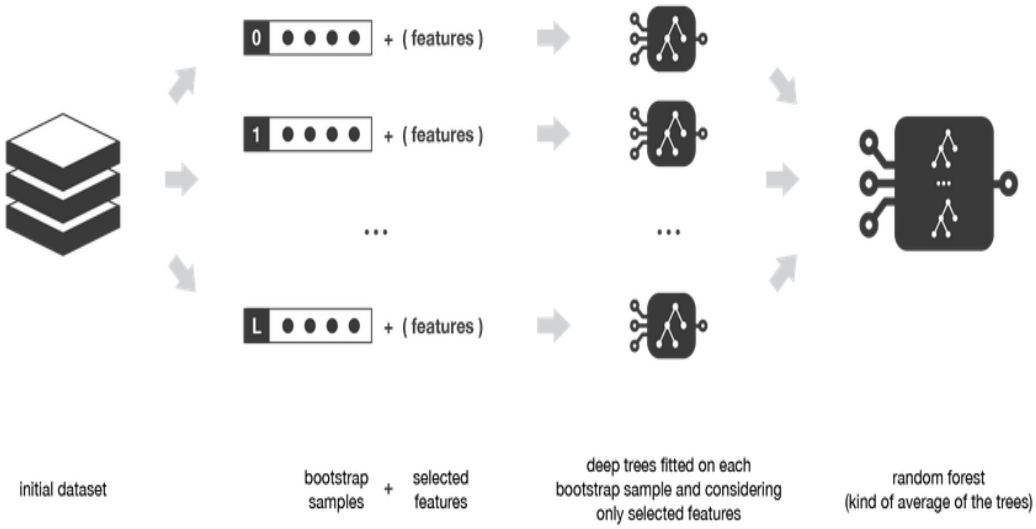


Figure 2.2: Bagging , adopted from [8]

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

$$\{z_1^1, z_2^1 .....z_B^1\}, \{z_2^2, z_2^2 .....z_B^2\}, \{z_1^L, z_2^L .....z_B^L\} \tag{2.2}$$

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

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

$$w_1(.), w_2(.), ..., w_L(.) \tag{2.3}$$

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

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

(Simple average, for regression problem)

2.1.2.2 Boosting:

Boosting is a sequential learning algorithm which trains weak learners to converts 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.

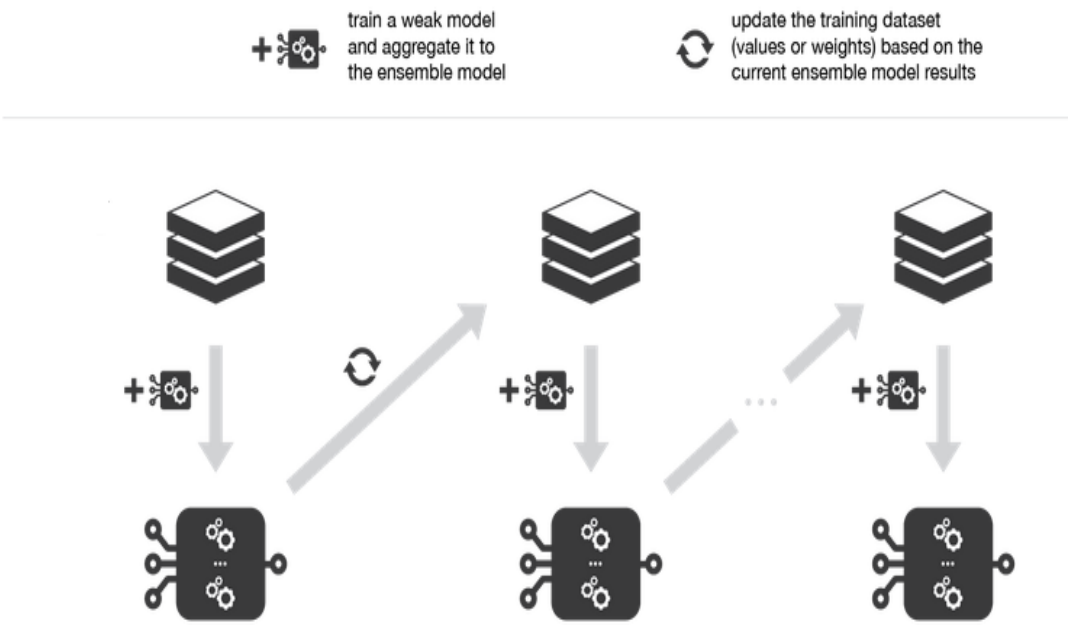


Figure 2.3: Boosting , adopted from [8]

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

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

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

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

2.2 Mixture of Experts

Blend of specialists deals with the rule that each master is had practical experience in a specific area of an information space by mimicking a gating system which is at risk for learning the joined load of the particular specialists for a specific information [11] ,by this technique info space is powerfully isolated and vanquished by the gating system and the specialists.

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 their exist a class which clearly outweight 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 epxertise 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.

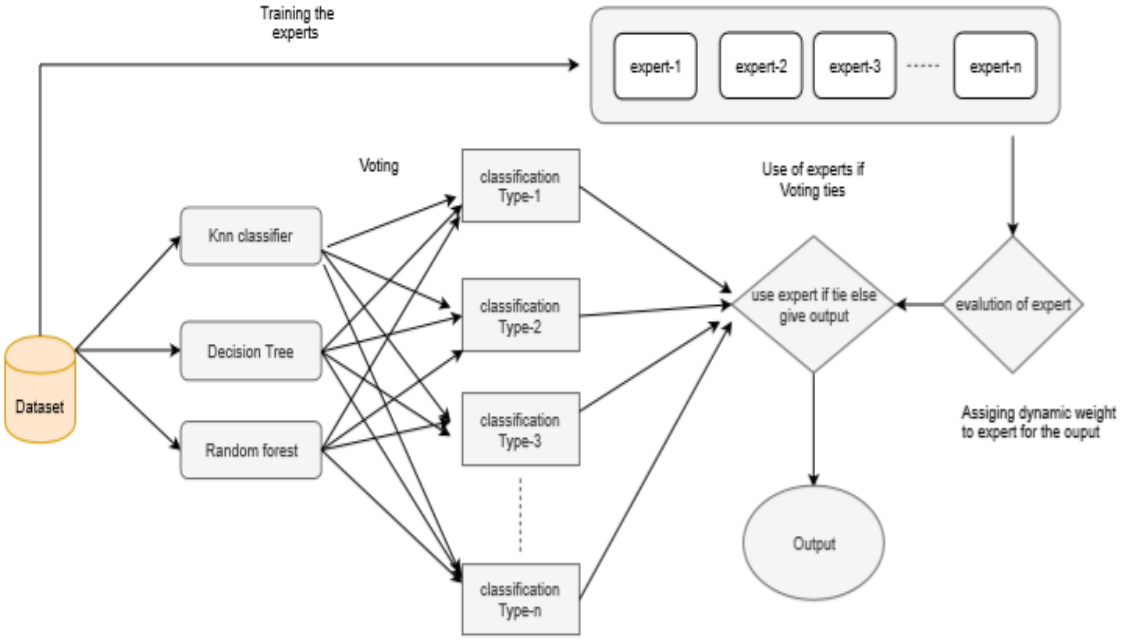


Figure 2.4: Mixture of Expert Architecture

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 kind of assault interlopers attempt to increase unapproved access to the framework over the system by sending the bundles to that framework.

4. **User to Root (U2R):**By this sort of assault interlopers first get to the framework with typical benefits later on attempt to get to the client with director benefits.

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 classes like Basic highlights, time delicate traffic, content highlights and host-based traffic highlights [19].

1. **Basic features:** Protocol type, Service, span and so forth.
2. **Time-based traffic features:**count, srv tally, Rerror rate and so forth.
3. **Content features:** hot, num root, is visitor login.
4. **Host-based traffic features:**dst host check, dst have srv tally and so on.

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.

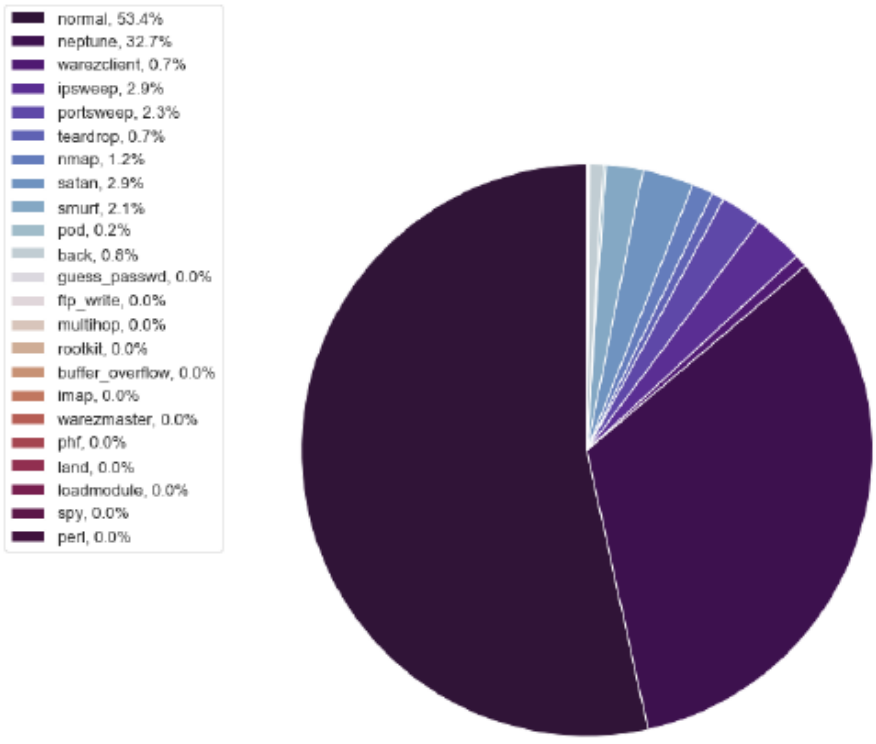


Figure 3.1: Attack Distribution

2. **Attack Distribution in training dataset** 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 their exist a class which clearly outweight 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 epxertise 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.

3. **Attack Distribution in testing dataset**

3.1.4 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 their exist a class which clearly outweight 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 epxertise 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.

3.1.5 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 their exist a class which clearly outweight 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 epxertise 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.

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.

Classifer	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

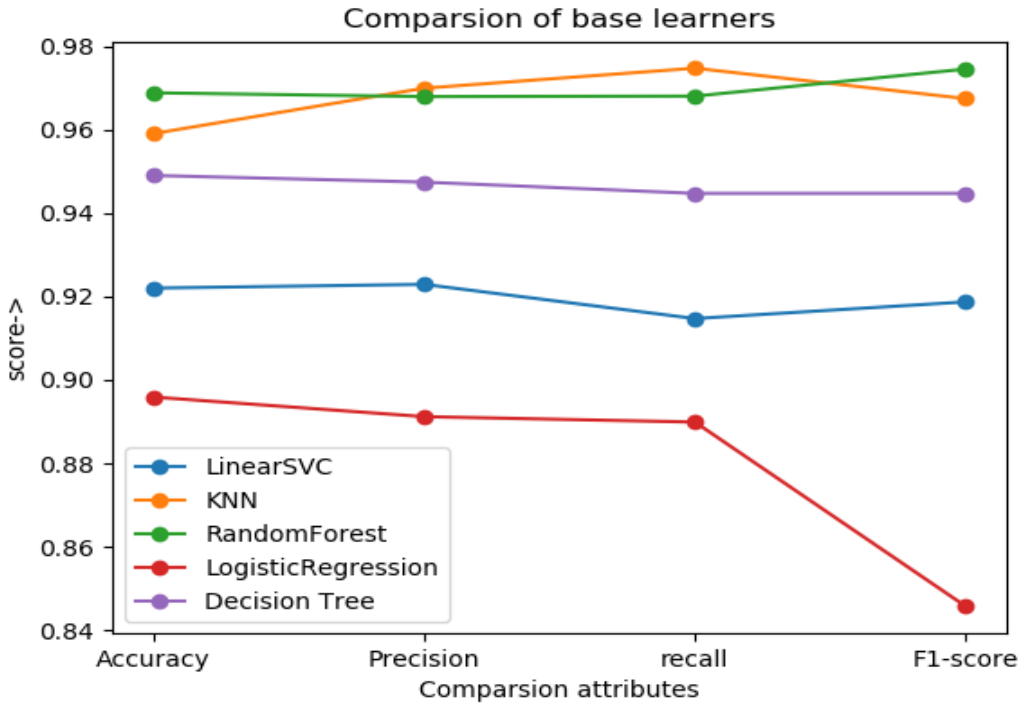


Figure 4.1: Comparison plot of Base Learners

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

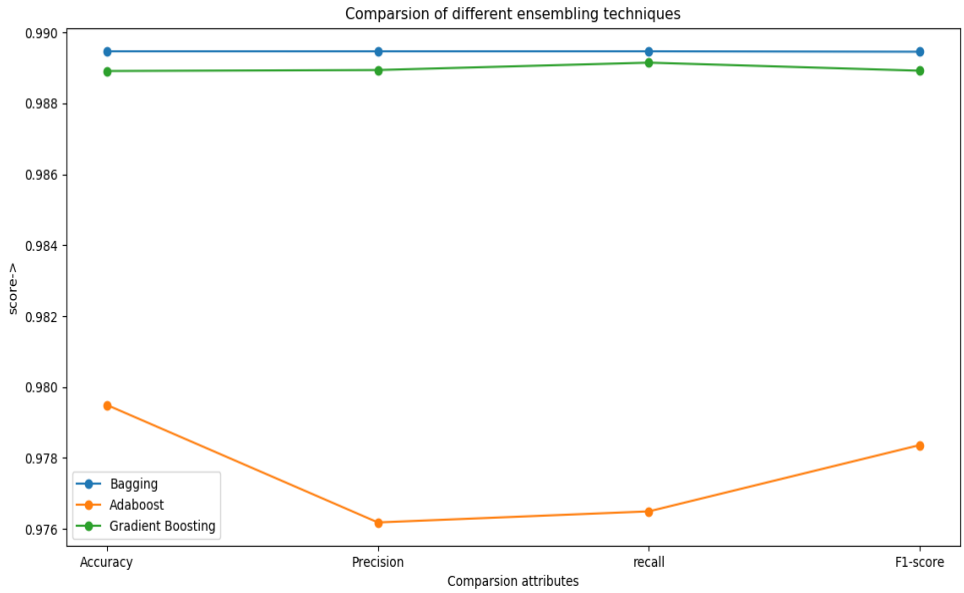


Figure 4.2: Comparison plot of Ensembles

4.2.1 AUC-ROC curve of Bagging

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

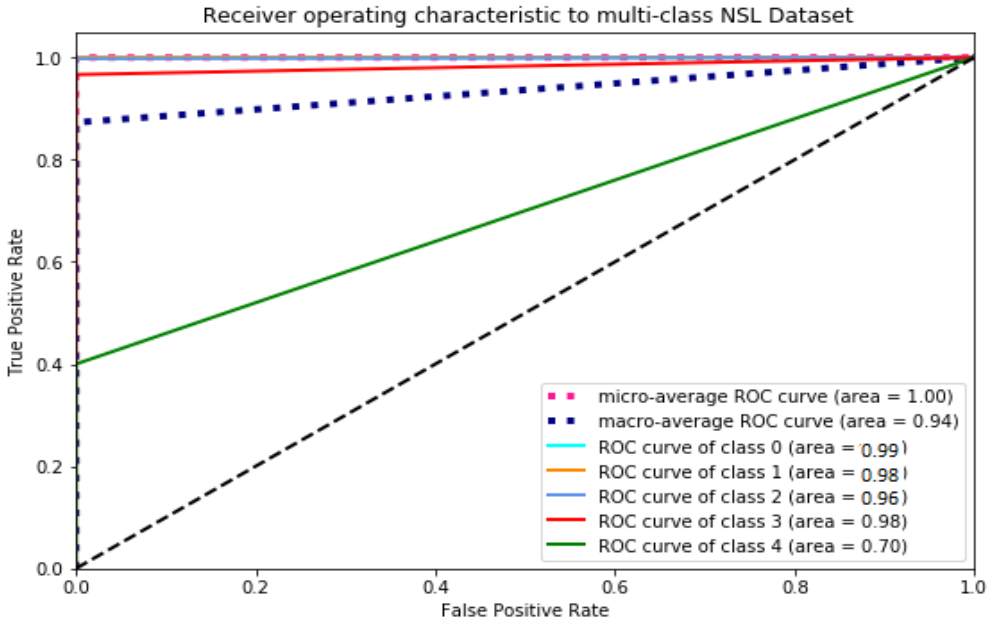


Figure 4.3: AUC-ROC of Bagging

4.2.2 AUC-ROC curve of Adaptive Boosting

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

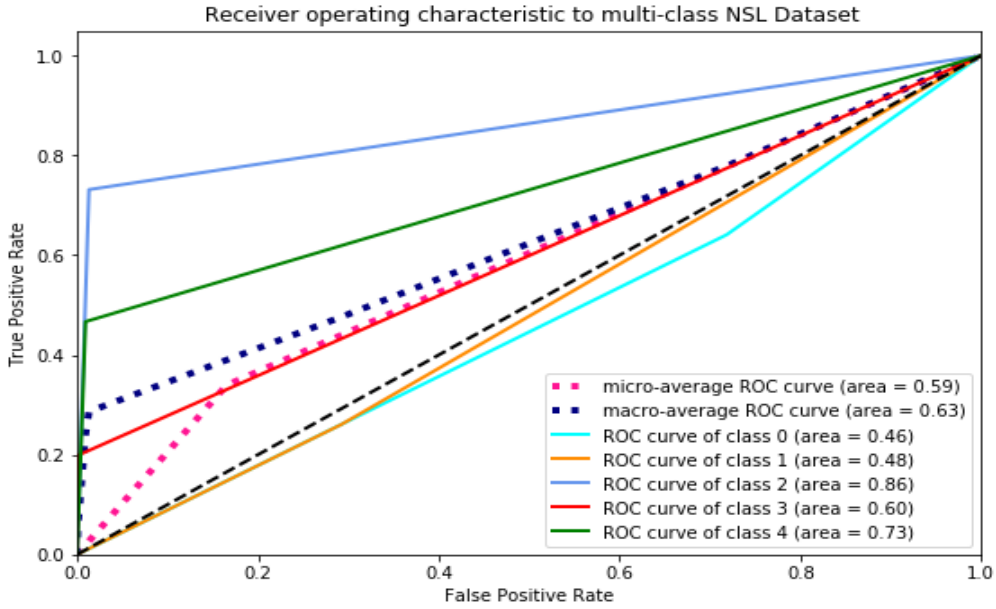


Figure 4.4: AUC-ROC of Adaptive Boosting

4.2.3 AUC-ROC curve of Gradient Boosting

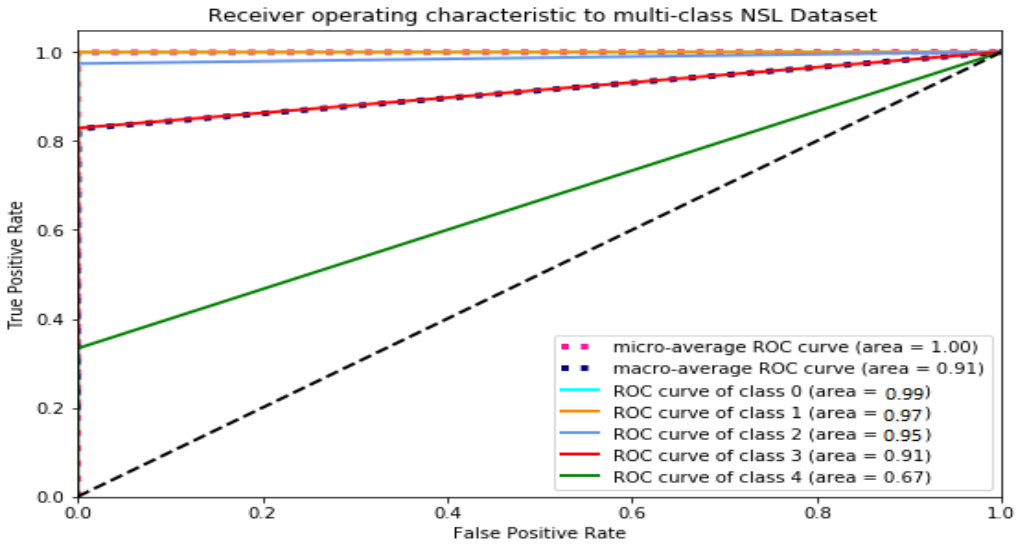


Figure 4.5: AUC-ROC of Gradient Boosting

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

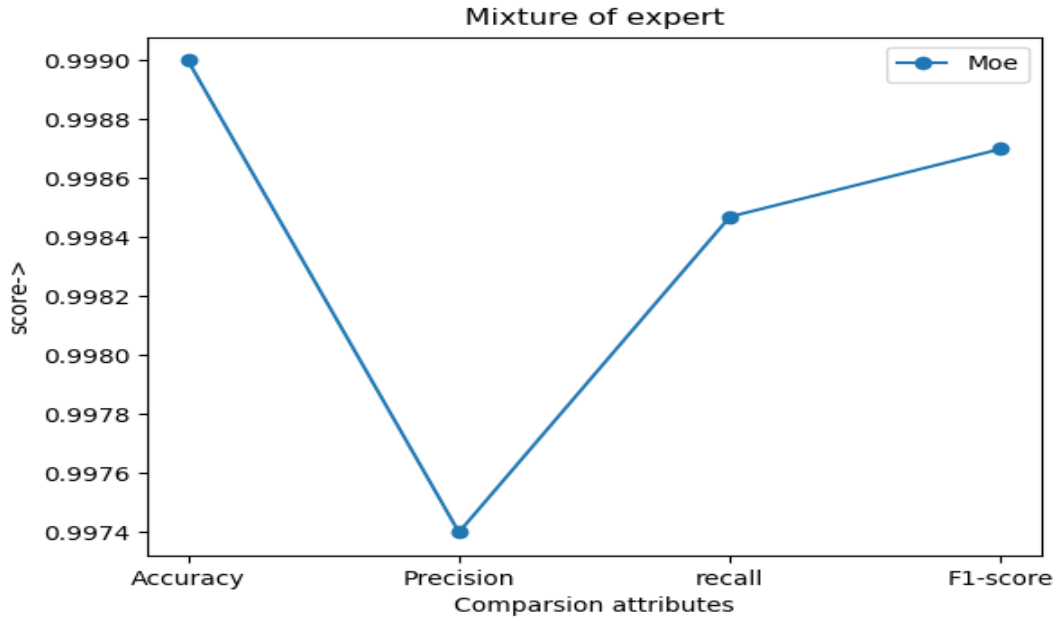


Figure 4.6: Comparison plot of MOE

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.

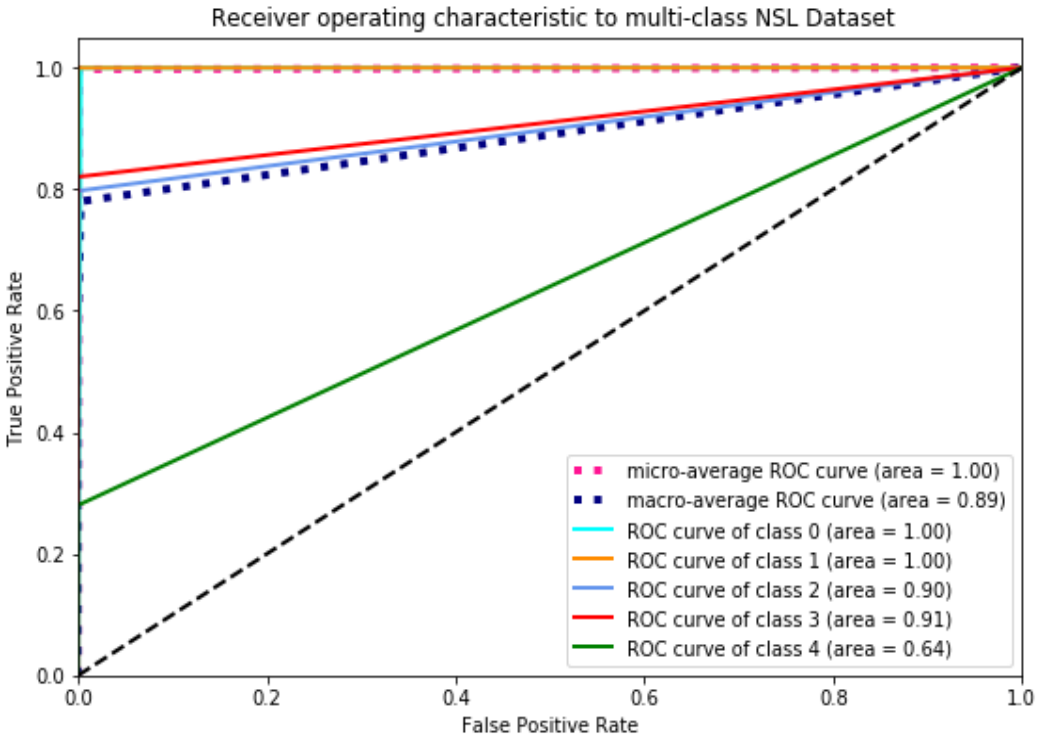


Figure 4.7: AUC-ROC curve of MOE

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.

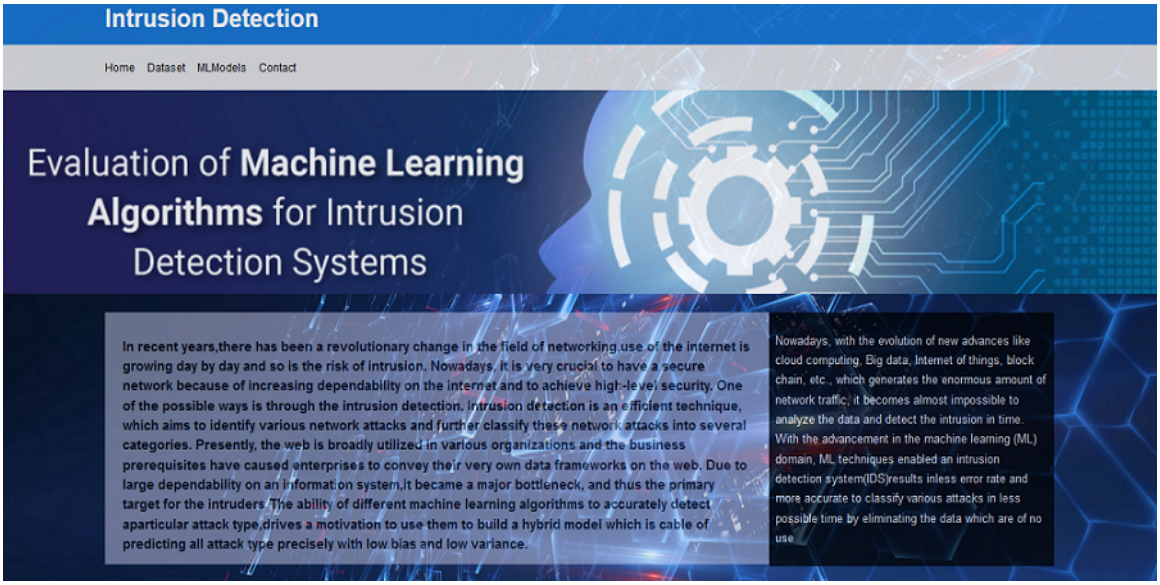


Figure 4.8: Home Page

- 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.

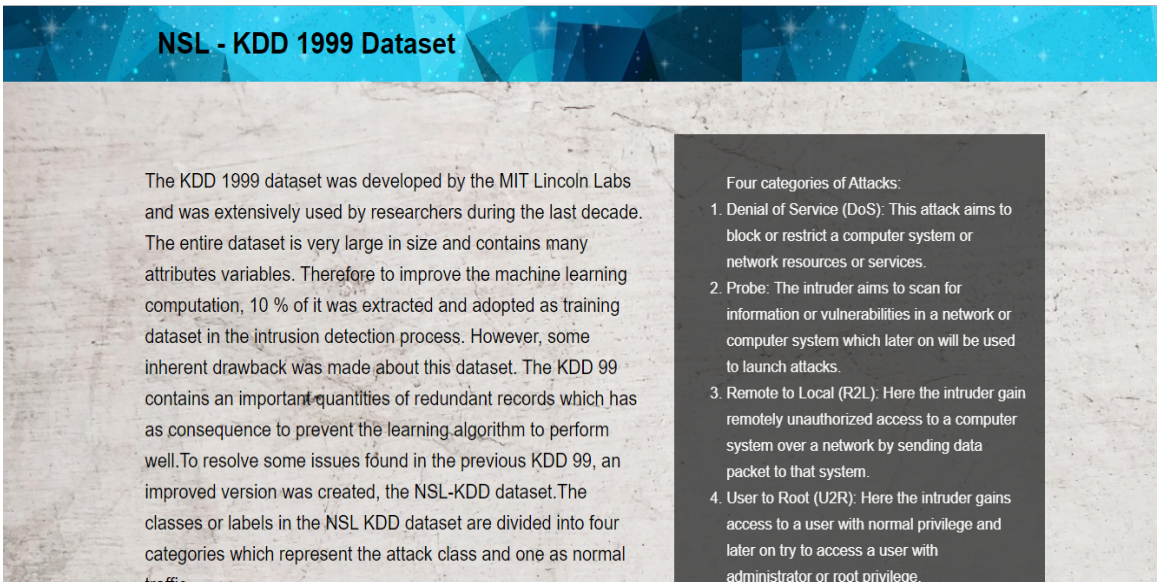


Figure 4.9: Dataset Page

- 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.

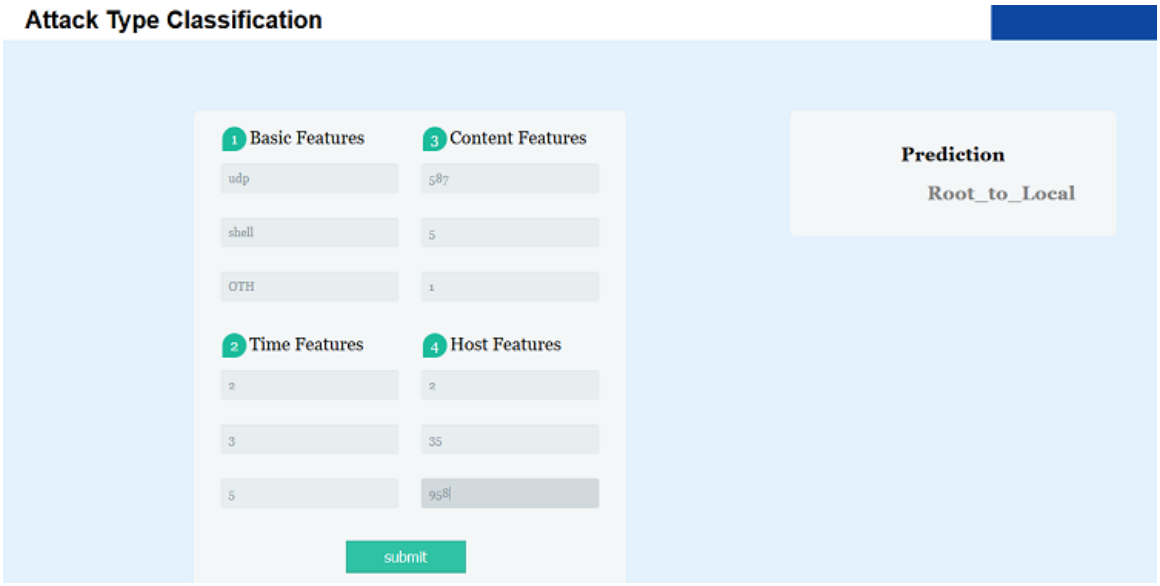


Figure 4.10: Prediction Page

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. Tavallaee, 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. Sharifizadeh. *"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.