International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-1, November 2019

Classifier Rank Identification using Multi-Criteria Decision Making Method for Intrusion Detection Dataset

Priyanka Patsariya, Rajni Ranjan Singh

Abstract*: Network intrusion detection system (NIDS) tracks network traffic for suspicious activity and policy violations. It generates alerts whenever such activity found. The objective is to detect and report anomalies. Further intrusion prevention system can take action such as blocking traffic from suspected IP addresses. Classification of network traffic as is a tedious task. Existing classifiers are suffered by generating many/false alerts. It is paramount important to select best classification approach among set of available approaches. KDD 99 is the benchmark dataset utilized to test the classification capabilities of classifiers. However, many classifiers generate similar results by measuring performance on various criteria. Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is a traditional multi-criteria decision making (MCDM) approach which is widely used to rank classifiers from number of options that are assessed on various criteria. In this work, KDD 99 dataset is applied as input to bayes net, naive bayes, NB updateable, random forest, oneR, zeroR, adaboostM1, decision stump, J48 and decision table classifiers. The performance of each classifier is measured using 10 different criteria's such as accuracy, misclassification, RA error, RMS error, false positive rate, f- measure, precision, RRS error, mean absolute error and recall. In order to test the effectiveness of proposed approach weka utility is utilized for classification and classifier performance result are supplied to the TOPSIS. An application is designed to implement TOPSIS method using python. It is observed that J48 secured at the top position with performance score 0.5829.

Keywords : negative ideal solution (NIS), The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), KDD99, positive ideal solution (PIS), , multi-criteria decision making (MCDM).

I. INTRODUCTION

IDS is a form of monitoring system that alerts administrators about malicious operations and violations of privacy policies that are attempting to compromise the information system. There are various kinds of intrusion detection system: Signature based IDS, Stack based IDS, Network based IDS, Host based IDSand Anomaly based IDS. KDD99 is the benchmark dataset utilize to test the classification capabilities.

Revised Manuscript Received on November 30, 2019.
***** Correspondence Author
Ms. Priyanka Patsariya*, Research scholar, Department of Computer Science & Engineering, Madhav Institute of Technology & Science, Gwalior, India. patsariya.priya900@gmail.com
Mr. Rajni ranjan singh, Assistant Professor, Department of Computer Science & Engineering, Madhav Institute of Technology & Science, Gwalior, M.P., India. rrsingh@mitsgwalior.in

The KDD training dataset comprises of approximately 4,900,000connections vectors each with 41 features and ismarked as either normal or as an attack, with precisely one specific type of attack. Different combinations of feature selection and classification technique are utilized to reduce time and complexity. However, selection of best classifier from available classifiers is a challenging task. In order to resolve this problem, multi criteria decision making (MCMD) technique can be utilized. Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is a method suggested by Hwang and Yoon [27] with a resemblance to the ideal solution for order choice. It is focused on the concept that alternatives that are the smallest range from the positive ideal solution and the farthest range from the negative ideal solution. The ideal-positive alternative maximizes the requirement of advantage and minimizes the cost criteria. The negative-ideal alternative (NIS) maximizes the requirements for costs and minimizes the requirements for benefits. In summary, the positive ideal solution comprises of all the highest achievable criteria values and the NIS made up of all the worst possible requirements.

II. RELATED WORK

Ranjit Panigrahia,Samarjeet Borah [2018] Presented a wave in which J48 classifier group has been considered for ranking in this document. This research is crucial in order to apply a particular identification method straight to a domain for optimum results. Two frequently used security domain datasets have been selected as binary and multiclass information. Classification accuracy under varying test circumstances and matrices associated with tree creation, classification precision, error indicators, frequency of identification, etc., and the J48 classifier group was carefully assessed. To determine weight and rank, the test outcomes were sent to TOPSIS. J48 secured the top rank and both J48Graft and J48Consolidated outperformed the binary dataset, while J48Graft outperformed the multi-class dataset. Raman Singh, Harish Kumar, R.K. Singla [2014] Suggest selection algorithm depending on the technique for order of preference by similarity to ideal solution (TOPSIS). TOPSIS is used to propose a solution, with many characteristics, among some alternatives. To analyze the NSL-KDD network dataset, a total of ten feature selection methods were used. For this experiment, three classifiers were considered using the Weka data mining tool, namely Naïve Bayes, J48 and PART. Using MATLAB as an instrument, the ranking of methods using TOPSIS was calculated.

Filtered Subset Evaluation was discovered to be appropriate for intrusion detection with reasonable precision in terms of very less computational time. Saad Mohamed Ali Mohamed Gadal and Rania A. Mokhtar [2017] proposed a hybrid machine learning method which is the detection of anomaly. The hybrid algorithm of data Mining approach creates upon the composition of SMO classification & K-means clustering.TheNSL-KDD datasets utilized as input and a hybridtechnique (k-mean+ SMO) is performed. It diminishes the false alarm rate to 1.2% and reached the accuracy of 97.3695%. It also observed positive detection rate of 94.48%.

Renato A. Krohling, André G. C. Pacheco [2015] Proposed to fix the issue of ranking and comparing algorithms, suggest an alternative novel technique based on the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). Algorithms are performed several times in evolutionary computation and then a statistic is calculated in terms of mean values and standard deviations. It is very prevalent to manage such a problem through statistical trials to compare the efficiency of algorithms.

EwaRoszkowska[2007] Present multi-criteria models in TOPSISselection settlement is discussed. It systematizes the learning of choosing approaches with the use of TOPSIS methods within the extensions. Practical applications of differentviewpoints of this method display straightforward numerical models that refer to real conditions.

III. KDD CUP'99 DATASET DESCRIPTION

The KDD Cup '99 [24] dataset created in the year 1999 by MIT Lincoln Laboratory was known as the KDD CUP '99 Intrusion Detection Dataset. The KDD training dataset comprises of roughly 4,900,000 vectors each with 41 characteristics and is marked as normalor an attack. Attack categories are represented in four subcategories such as U2R, DOS, R2L and probe attack. KDD'99 is 50MB in size. Four attack sub categories are:

1) **Denial of service attack (DOS):** In denial of service attack, attacker makes some computing or memory resource too busy or too full to handle legitimate requests.

2) **User to root attack (U2R):**Is an exploit class in which an attacker start access a normal user account on the system (maybe gained through sniffing passwords, a dictionary attack, or social engineering) and can exploit some vulnerability to gain root access to the system.

3) **Remote to local attack (R2L):** It is an attack in which attacker has the capacity to send packets to a machine over a network without a having an account on that machine. It exploits the vulnerability to access local machine as a user of that machine.

4) **Probe Attack:** Probe is anattack that scans a machine or networking device to identify weakness for the apparent purpose of circumventing its security controls.

IV. PERFORMANCE PARAMETER FOR EVALUATION

Performance parameters to evaluate the capability of classifiers are as follow:

- Build Time (Seconds): It is a time taken by classifier to construct the model by the training dataset.
- Correctly classified instances: Shows classifier's accuracy rate. Properly classified instances could be determined as:

$$\textbf{Accuracy}=\frac{true\ positive+true\ negative}{total\ instance}$$

- Incorrectly classified instances (%): Shows classifier's general misclassification rate. It is, therefore, possible to calculate incorrectly classified instances as:

$$\textbf{Misclassification Rate}=\frac{False\ positive+false\ negative}{total\ instance}$$

- Mean absolute error (MAE): Summary of all instances the absolute errors divided by the number of cases through a real class label into the test set. MAE is characterized as:

$$\textbf{MAE}=\frac{1}{n}\sum|x_i-x|$$

Where n= with a class label number of error instances and $|x_i-x|$= per instance absolute error.

- Root mean squared error (RMS): RMS error generally offers the extent to which the model offers the correct response. It represents in the same scale (unit) an average prediction error. It is calculated where the calculation takes place.

$$\textbf{RMS}=\sqrt{\frac{1}{n}\sum\varepsilon_i^2}$$

- Relative absolute error (RA) (%): RA error is the proportion of the complete absolute error of real number through simple predictor absolute error. RAE is portrayed via:

$$\textbf{RA}=\frac{\sum_{i=1}^{N}|\hat{x}_i-x_i|}{\sum_{i=1}^{N}|\bar{x}-x_i|}$$

- Root Relative square error (RRS) (%): It is the RMSE-to-RMSE ratio acquired by anticipating target mean and multiplied through 100. Minimum values are better; values above 100% imply that a system is doing worse. It is considered as follow:

$$\textbf{RRS}=\sqrt{\frac{\sum_{i=1}^{N}(\hat{x}_i-x_i)^2}{\sum_{i=1}^{N}(\bar{x}-x_i)^2}}$$

- True Positive Rate: TPR determined via:

$$\textbf{TPR}=\frac{TP}{FN+TP}$$

- False Positive Rate: FPR calculated through:

$$\textbf{FPR}=\frac{FP}{TN+FP}$$

- Precision: Precision defined via:

$$\textbf{PPV /Precision}=\frac{TP}{TP+FP}$$

- F-Measure: Between recall and precision, the harmonic mean is known as the f-score/f-measure.

$$\text{F-Measure}=\frac{2(Precision-recall)}{(Precision-recall)}$$

V. PROPOSED WORK

In this section, various phases of proposed work are explained in detail.

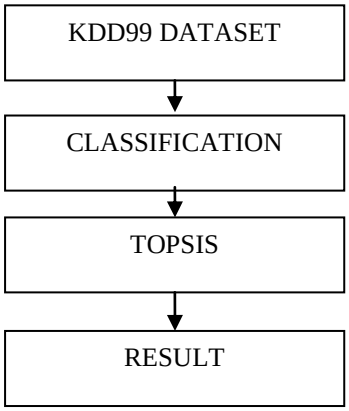

Fig1. Flow diagram of proposed work

Dataset KDD99: - KDD99 dataset is used as an input in this proposed work. KDD Cup 1999 dataset is separated into two parts- testing and training dataset. The KDD training dataset comprises of roughly 4,900,000 vectors and testing dataset contain 311028 vectors.

Classification: - Classification is performed by using weka 3.9 utility on Linux platform. In this phase 10 different classifiers are evaluated and result are stand in a .csv format.

TOPSIS: - TOPSIS based on the idea that the best choice should be the least distance from the ideal solution that is a Euclidean distance. A program is developed to implement TOPSIS algorithmusing python. The algorithm of TOPSIS is given below: -

TOPSIS Algorithm: -

Step1. Construct the decision matrix and determine the weight of criteria.

Let X = (x_j) be a decision matrix and [, , ... ,] W = [$w_1, w_2, ... w_n$] a vector of weight, where xij ∈ P , wj ∈ P and $w_1 + w_2 + + w_n = 1$.

Function criteria may be: benefit functions (more is better) or cost functions (less is better).

Step2. Compute normalized decision matrix.

Its covert dimensions of different characteristics into non-dimensional attributes that enable criteria to be compared. Because different criteria are generally evaluated in different units, it is necessary to transform the result in the assessment matrix X into a normalized scale. One of the several recognized standardized formulas can be used to normalize vales. Some of the techniques used most frequently to calculate the normalized value n_{ij} are as follows:

$$\mathbf{n_{ij}} = \sqrt{\sum_{j=1}^{n} x_{ij}^2}$$

Step3. Compute weighted normalized decision matrix.

The weighted normalized value ij v is calculated as follows:

$$\bar{x}_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{n} x_{ij}^2}} \times \mathbf{w_j}$$

j = 1, ..., n, i = 1, ..., m. Where w_j is the weight of the j-th criterion, 1.

$$\sum_{j=1}^{n} w_j$$

Step4. Establish the positive ideal solution (PIS) and the negative ideal solutions (NIS) that are negative.

Identify the positive ideal option (each criterion should have peak performance) and the negative ideal option (each criterion requires reverse intense performance). The positive ideal solution can be defined as the solution in which the profit requirements is maximum and requirement of cost is minimum on the other hand the ideal negative solution have high cost and the benefit requirements is less.

PIS (S^+) has the following equation:

$$S^+ = (v_1^+, v_2^+ ..., vn^+) = \left((maxv_{ij} | j \in I), (minv_{ij} | j \in J) \right)$$

NIS (S^-) has the following equation:

$$S^- = (v_1^-, v_2^- ....., vn^-) = \left((minv_{ij} | j \in I), (maxv_{ij} | j \in J) \right) \backslash$$

Here I associate to profit requirement and the cost requirement with J,

i=1…, m; j=1……, n.

Step5. Determine the separation steps from the PIS and the NIS.

Separating every option via PIS is defined as-

$$\mathbf{s_i^+} = \left[\sum_{j=1}^{m} \left(\mathbf{V_{ij}} - \mathbf{V_j^+} \right)^2 \right]^{0.5}$$

Separating every option via NIS defined as-

$$\mathbf{s_i^-} = \left[\sum_{j=1}^{m} \left(\mathbf{V_{ij}} - \mathbf{V_j^+} \right)^2 \right]^{0.5}$$

Step6. Compute the relative proximity for the PIS.

The i-th alternative S_j relative proximity in respect of S^+ is given in the following equation -

$$\mathbf{P_i} = \frac{\mathbf{S_i^-}}{\mathbf{S_i^+ + S_i^-}}$$

Step7. Rank the order of preference or select the option closest to 1.

By sorting the value order of P_i in descending order, a set of alternatives can be sorted.

VI. EXPERIMENTAL RESULT

Firstly, take different classification technique with their performance criteria.

Table-I: Different classifiers with criteria

Performance Matrices	Accuracy (%)	Misclassification (%)	Recall	Mean Absolute error	RMS-error	RA-error (%)	RRS-error (%)	false-positive rate	Precision	f-Measure
Naïve Bayes	92.215	7.784	0.922	0.031	0.174	23.247	67.425	0.01	0.982	0.948
Bayes Net	98.795	1.205	0.988	0.004	0.064	3.613	24.842	0	0.944	0.991
NBUpdateable	92.215	7.784	0.922	0.031	0.174	23.247	67.425	0.01	0.982	0.948
J48	99.983	0.017	1	0	0.008	0.096	3.107	0	1	1
Random Forest	99.815	0.185	0.998	0.996	0.003	2.784	10.486	0.001	0.998	0.998
Decision Table	99.841	0.158	0.998	0.002	0.026	1.906	10.057	0.003	0.998	0.998
ZeroR	79.239	20.76	0.792	0.133	0.258	100	100	0.792	0.5	0.613
OneR	98.178	1.821	0.982	0.007	0.085	5.465	33.063	0.063	0.939	0.96
Decision Stump	97.623	2.376	0.976	0.018	0.095	13.526	36.778	0.023	0.976	0.976
AdaboostM1	97.623	2.376	0.976	0.032	0.096	24.096	37.511	0.023	0.976	0.976

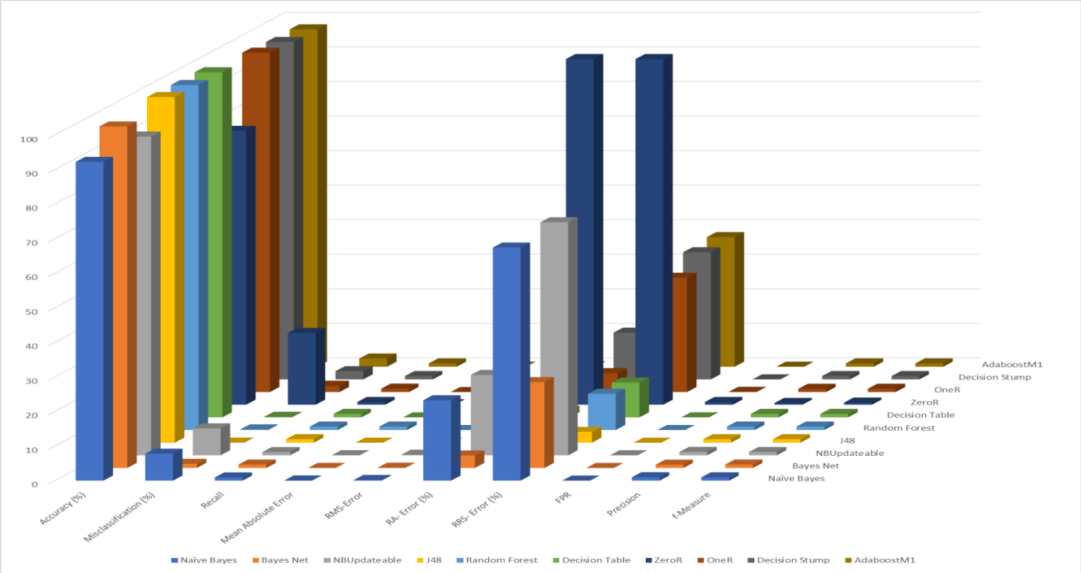

Fig.2 Different classifiers with performance criteria

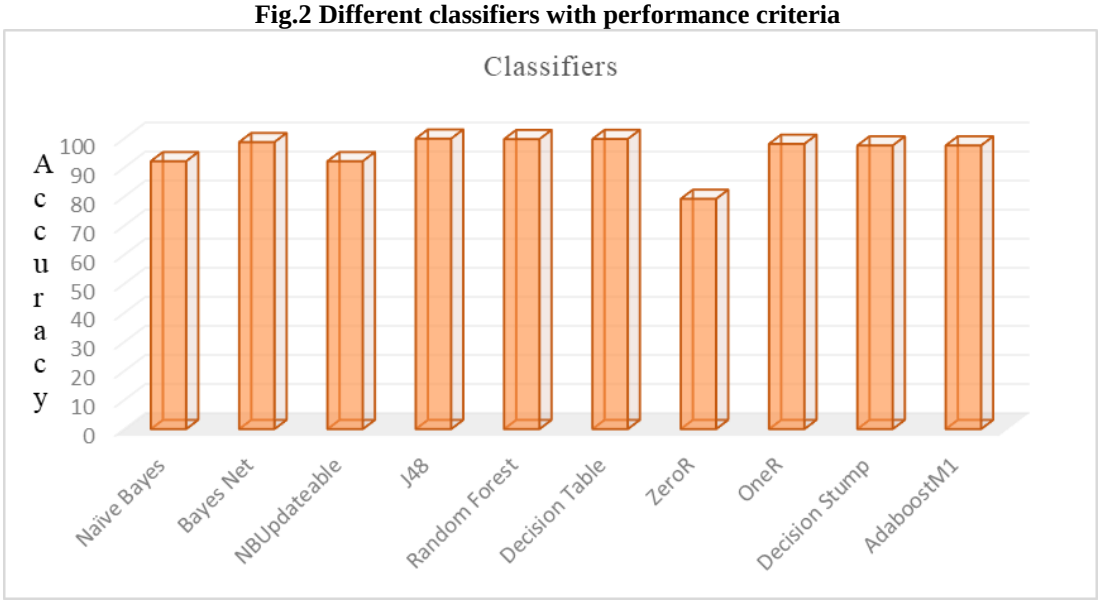

Fig.3 Different classifiers with accuracy

After normalizing the data,TOPSIS is utilized to select the best Classifier out of all available 10 classifier technique based on 10 criteria that are accuracy, misclassification, RMS-error, recall, RA-error, False-negative rate, Precision, F-measure, mean absolute error, RRS-error.
Now the values were substituted in TOPSIS vector normalization. Formula for the vector normalization: -

$$\sqrt{\sum_{j=1}^{n} x_{ij}^2}$$

Do vector normalization using this denominator.
The square root of the squared summation gives the denominator. All values of accuracy are squared and added and the root of the value is taken.

For Accuracy: -
$(92.2151)^2+(98.795)^2+(92.2151)^2+(99.983)^2+(99.815)^2+$
$(99.8411)^2+(79.2393)^2+(98.1784)^2+(97.6238)^2+(97.6238)^2$
= 91674.26175

Now take square root of value is calculated: -

$\sqrt{91674.26175}$ = **302.77757292798293**

Similarly, calculate values for other criteria also which are given below: -

Table-II: Vector Normalization

S. N.	Criteria	Values
1	Accuracy	302.77757292798293
2	Misclassification	428.1921500171623
3	Recall	524.4035824976028
4	Mean Absolute Error	608.222437218753
5	RMS Error	671.8215074975793
6	RA Error	723.6338062642734
7	RRS Error	754.5127341503987
8	False Positive Rate	810.6078228065101
9	Precision	863.4929041370926
10	F-Measure	913.9948617695344

The performance value in each cell is divided by the rooted summation of square value and multiplies the weights of each criterion with the normalized performance value of each cell.

$$\bar{x}_{ij}=\frac{x_{ij}}{\sqrt{\sum_{j=1}^{n} x_{ij}^2}}$$

Calculate weight using this formula: -

1/10= 0.1

Because there is 10 criteria and the total summation of the entire alternative becomes 1.
For the normalized weighted matrix, calculate accuracy is: -

92.2151/ 302.77757292798293×0.1= 0.03045638

The performance value 92.2151 is divided by 302.77757292798293 then it will be multiplied by the weight 0.1, which gives the weighted normalized matrix. Similarly, do it for all of the cells.

Table-III: Normalized Weighted Matrix of Classifiers

Performance Matrices	Accuracy (%)	Misclassification (%)	Recall	Mean Absolute Error	RMS-Error	RA-Error	RRS-Error	false-positive Rate	Precision	F-Measure
Naïve Bayes	0.0304564	0.0215359	0.0175819	0.0159317	0.0122934	0.0106065	0.0043173	0.0122131	0.0113724	0.0103721
Bayes Net	0.0326296	0.0230726	0.0188405	0.0163624	0.0139308	0.0133198	0.0099611	0.0123364	0.0109323	0.0108425
NBUpdateable	0.0304564	0.0215359	0.0175819	0.0159317	0.0122934	0.0106065	0.0043173	0.0122131	0.0113724	0.0103721
J48	0.0330219	0.02335	0.0190693	0.0164397	0.0147658	0.0138058	0.0128417	0.0123364	0.0115809	0.010941
Random Forest	0.0329664	0.0233108	0.0190311	0.0163805	0.0144815	0.0134343	0.0118637	0.0123241	0.0115577	0.0109191
Decision Table	0.0329751	0.0233169	0.0190311	0.0164003	0.0144979	0.0135557	0.0119206	0.0122994	0.0115577	0.0109191
ZeroR	0.0261708	0.0185056	0.0151029	0.0142497	0.0110416	0	0	0.0024772	0.0057904	0.0067068
OneR	0.0324259	0.0229286	0.018726	0.0163213	0.0136137	0.0130638	0.0088715	0.0115592	0.0108825	0.0105075
Decision Stump	0.0322427	0.0227991	0.0186116	0.0161454	0.0134708	0.0119499	0.0083791	0.0120527	0.0113134	0.0106838
AdaboostM1	0.0322427	0.0227991	0.0186116	0.0159136	0.0134441	0.0104892	0.008282	0.0120527	0.0113134	0.0106838

Next calculate the best and worst ideal value in case of beneficial criteria like accuracy maximum value is desired so ideal best will be the maximum value and the ideal worst will be the minimum value

V_j^+ = indicate the Ideal best value (Positive Ideal Solution)

V_j^- = Indicate the Ideal worst value (Negative Ideal Solution)

Table-IV:. Best and worst ideal solution

Criteria	Ideal BestV_j^+	Ideal WorstV_j^-
Accuracy	0.0330219	0.0261708
Misclassification	0.0185055	0.02335
Recall	0.0190693	0.0151029
MAE	0.0142497	0.0164397
RMS-Error	0.0110416	0.0147658
RA-Error	0.0104892	0.0138058
RRS-Error	0.0043172	0.0128417
FPR	0.0024772	0.0123364
Precision	0.0115809	0.0057904
F-Measure	0.010941	0.0067068

The formula for calculating the best and worst idle Euclidean distance: -

$$s_i^+ = \left[\sum_{j=1}^{m}\left(V_{ij} - V_j^+\right)^2\right]^{0.5}$$

$$s_i^- = \left[\sum_{j=1}^{m}\left(V_{ij} - V_j^+\right)^2\right]^{0.5}$$

First, there's a formula for calculating the Euclidean distance from idle best so here calculate the Euclidean distance first.

$$s_i^- = \left[\sum_{j=1}^{m}\left(V_{ij} - V_j^+\right)^2\right]^{0.5}$$

Calculate the Euclidean distance for the first alternative that is naive Bayes. It takes the difference and then square it. Finally, sum all the squared values and then take the root or the square root.

$$(0.03045638\text{-}0.03302193059846656)^2+(0.02153592\text{-}0.01850554710001667)^2+(0.017581880.019069282388141796)^2+(0.015931670.014249720940305974)^2+(0.012293440.0110416232841827 1)^2+(0.010606480.01048920038601 4006)^2+(0.00431725\text{-}0.004317249865461769)^2+(0.01221306\text{-}0.002477153493347552)^2+(0.011372420.011580871078486996)^2+(0.01037205\text{-}0.010940980544069535)^2$$

= 0.012715475206395887

This is the value for the naive bayes. Similarly, calculate all the distance of Euclidean as of the idle best. Similarly, for the ideal worst, calculate the first alternative that negative ideal solution (NIS) is naive bayes. It is being squared and summed up and then the square root is being taken to get a value

$$(0.030456380.02617079568797767)^2+(0.021535920.023350031054047253)^2+(0.017581880.015102871651408304)^2+(0.015931670.016439709205275113)^2+(0.012293440.014765826769896531)^2+(0.010606480.013805795021621055)^2+(0.004317250.012841718318923194)^2+(0.012213060.012336421779619286)^2+(0.011372420.005790435539243498)^2+(0.010372050.006706821073514626)^2$$

= 0.010841691079090906

Similarly, calculate all Euclidean distance and the performance score is alsocalculated.

This is the formula for performance Score it is denoted by P_i

$$P_i = \frac{S_i^-}{S_i^+ + S_i^-}$$

Add S_i^+ and S_i^- values and then the PIS from Euclidean distance are divided with a sum:-

$$P_i = \frac{0.010841691079090906}{0.012715475206395887+0.010841961079090906} = 0.460228999859389$$

Table-V: Ideal solution

Classifiers	Positive Ideal Solution S_i^+ from Euclidean distance	Negative Ideal Solution S_i^- from Euclidean distance	Performance Score P_i $\frac{S_i^-}{S_i^+ + S_i^-}$
Naïve Bayes	0.012715475	0.010841691	0.460229
Bayes Net	0.010418373	0.013091079	0.556843226
NB Updateable	0.012715475	0.010841691	0.460229
J48	0.010683057	0.014933395	0.582961098
Random Forest	0.010667587	0.014214568	0.571275591
Decision Table	0.010664004	0.014262366	0.572179823
ZeroR	0.000000000	0.000000000	0.000000000
OneR	0.010538714	0.011894258	0.530213199
Decision Stump	0.011000242	0.011798582	0.517508375
AdaboostM1	0.011388476	0.011633515	0.505321841

So, based on the performance score, rank the alternative. J48 is the best alternative as it is having the maximum value.

Table-VI:Rank Allocation of Classifiers

Classifier Name	Score	Rank
J48	0.5829610978532870	1
Decision Table	0.5721798234242056	2
Random Forest	0.5712755909093559	3
Bayes Net	0.5568432262574255	4
OneR	0.5302131987876587	5
Decision Stump	0.5175083745460517	6
AdaboostM1	0.5053218413675555	7
Naïve Bayes	0.4602289998593891	8
NB Updateable	0.4602289998593891	9
ZeroR	0.0000000000000000	10

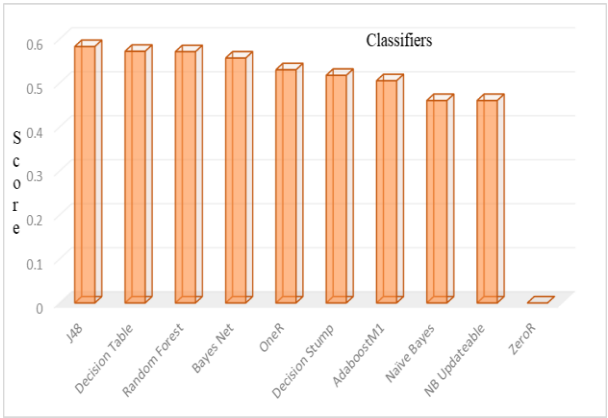

Fig.4 Classifiers with their performance score

VII. OBSERVATION

It is observed that J48 achieve highest performance score0.582 among 10 different classifiers. It is also observed that decision stump and adaboostM1 have same accuracy but adaboostM1 have high error rate (mean absolute error 0.032**,** RA- error 24.096, RRS- error 37.511) as compare to decision stump. Therefore, decision stump ranked higher than adaboostM1. It is noticed that ZeroR falls at last position because of high percentage of RA error and RRS error rate. Decision table secured 2nd highest performance score.

VIII. CONCLUSION& FUTURE WORK

Proposed approach is beneficial to select best classifier for intrusion detection based on multi-criteria decision-making approach. In futures it is proposed to extend the work by including feature selection and hybrid model of classification**.**

REFERENCES

1. Ravipati Rama Devi and MuntherAbualkibash "Intrusion Detection System Classification using different machine learning algorithms on KDD-99 and NSL-KDD datasets - A Review Paper" International Journal of Computer Science & Information Technology (IJCSIT) Vol 11, No 3, June 2019.
2. RanjitPanigrahia,Samarjeet Borah, "Rank Allocation to J48 Group of Decision Tree Classifiers using Binary and Multiclass Intrusion Detection Datasets"International Conference on Computational Intelligence and Data Science (ICCIDS 2018).
3. EwaRoszkowska, "Multi-Criteria Decision Making Models By Applying TheTopsis Method To Crisp And Interval Data".
4. Saad Mohamed Ali, Mohamed Gadal, Rania A. Mokhtar, "Anomaly Detection Approach using Hybrid Algorithm of Data Mining Technique"2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, Sudan.
5. Raman Singh, Harish Kumar, R.K. Singla, "TOPSIS Based Multi-Criteria Decision Making of Feature Selection Techniques for Network Traffic Dataset "International Journal of Engineering and Technology (IJET).
6. Renato A.Krohlinga,AndréG.C.Pacheco, "A-TOPSIS-An approach Based on TOPSIS for Ranking Evolutionary Algorithms"Information Technology and Quantitative Management (ITQM 2015).
7. RajniRanjan Singh, Deepak Singh Tomar,"Network forensics: detection and analysis of stealth port scanning attack", International Journal of Computer Networks and Communications Security, 3 (2), February 2015.
8. Shrikantupadhyay, rajniranjansingh, "Comparative Analysis based Classification of KDD'99 Intrusion Dataset" (IJCSIS) International Journal of Computer Science and Information Security, Vol. 13, No. 3, March 2015.
9. BhavnaDharamkar, RajniRanjan Singh "A Review of Cyber Attack Classification Technique Based on Data Mining and Neural Network Approach" International Journal of Computer Trends and Technology (IJCTT) – volume 7 number 2– Jan 2014.
10. Megha Jain Gowadiya, Anurag Jain, "Intrusion Detection in KDD99 Dataset: A Review"International Journal of Scientific & Engineering Research, Volume 6, Issue 11, November-2015 ISSN 2229-5518.
11. S. SobinSoniya,Maria Celestin Vigila, "Intrusion Detection System: Classification and Techniques",2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT].

AUTHORS PROFILE

Ms. PriyankaPatsariya, Research scholar, Department of computer science & Engineering, Madhav Institute of Technology & Science, Gwalior.

Mr. Rajni ranjan singh, Assistant Professor, Department of Computer Science & Engineering, Madhav Institute of Technology & Science, Gwalior, M.P., India. His research interest includes AlgorithDesign, Network Security, Network Forensics and Computer Networks.