International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT217239

Intrusion Detection System based on K-means, Classification and Regression Trees Algorithm

R. K. Borikar*, Dr. S. S. Sherekar, Dr. V. M. Thakare
PG Department of Computer Science and Engineering, Sant Gadge Baba Amravati University, Amravati, , Maharashtra, India

ABSTRACT

The security of computer networks has become a very important aspect in today's era. An intrusion detection system (IDS) is a type of security software designed to automatically inform administrators when someone is trying to compromise the information system through malicious activities or security policy violations. IDS works by monitoring the system functionalities by checking system vulnerabilities, the integrity of files, and performing analysis of patterns based on known attacks. Intrusion detection is the cycle called distinguishing intrusions. The activity which is entering a framework without consent is called interruption. Interruption Detection Systems are fundamental for security limits. This paper is focused on the analysis of five different techniques like k-means clustering and naive Bayes classification, Enhanced k-means algorithm, random forest, and weighted k-means, k-means clustering, regression trees algorithm, etc. But some problems are indicated by techniques. These methods analyzed the restrictions of intrusion detection systems by way of low accuracy, high incorrect alarm rate, and time consumption. So, to overcome these problems the proposed method 'K-means and Classification and Regression Trees Algorithm' is used to show good accuracy in performance analysis with time complexity by using a hybrid data mining method in the Weka tool. The proposed model of K-means and classification and regression trees algorithm depends on intrusion detection system that gives proper period popular huge information measure of these days' interruption data set.

Keywords : Intrusion detection, hybrid algorithm, CART, Kmeans, Clustering.

Article Info
Volume 7, Issue 2
Page Number : 163-171

Publication Issue :
March-April-2021

Article History
Accepted : 01 April 2021
Published : 03 April 2021

I. INTRODUCTION

Intrusion Detection System expects a huge part to accomplish higher security in distinguishing malicious activities for quite a long while. Existing irregularity recognition is oftentimes connected with high incorrect alerts by unobtrusive precision besides identification percentage once it cannot recognize a wide range of assaults effectively [1]. As of late different sorts of information mining strategies stayed functional to intrusion detection. Here stand dual significant ideal models aimed at preparing information mining-built interruption location frameworks that are abuse recognition and inconsistency discovery [2]. In oddity identification, the k-means grouping procedure remains utilized towards distinguishing original intrusion through grouping the organization associations information towards gathering the greatest intrusion together in at least single groups [3]. Abuse identification distinguishes interruption dependent on known examples while abnormality recognition centers around obscure examples. K-means is a common grouping procedure that takes continue demonstrated aimed at application to intrusion detection system [4]. Intrusion Detection can likewise be viewed as a characterization disadvantage. In this examination the utilization K-means procedure and classification and regression trees (CART) algorithm [5]. Principal Component Analysis is an ordinary quantifiable methodology aimed at information assessment and pre-dealing that takes stayed broadly functional in different turfs of examination. PCA is proposed to change the data in a reduced construction and keep most of the principal distinctions present in the hidden data. Complementing preventative technologies like firewalls, sturdy authentication, and user privilege. IDSs turned into a significant piece of big business IT security management. [6]. Network Intrusion Detection Systems take reliably stayed planned towards help and progress the suggestion safety matter via the workplace of assessment, recognizing, surveying and reportage any unapproved and ill-conceived network associations and exercises [7]. Intrusion Detection Systems (IDSs) have as of now pulled in the consideration of a significant segment of the world, determined their turn of events, and improving address a high need for association and investigators and science focuses [8]. Intrusion is one of the most dangerous to the net. Safety matters needed remaining a tremendous disadvantage. A lot of procedures and techniques are invented to manage the requirements of intrusion detection systems like low precision, high outburst rate, and time-consuming [9]. Partner in the nursing of Intrusion detection Systems (IDS) screens either network or alternative frameworks for malicious or strange practices. Supplementing protection advancements like firewalls, tough confirmation, and client advantage can be utilized for IDS. [10].

This paper focused on the analysis of techniques such as k-means clustering and naive Bayes classification, Enhanced k-means algorithm, random forest, and weighted k-means, k-means clustering, regression trees algorithm. This paper is applied to utilize k-means and classification and regression trees (CART) algorithms in the direction of organizing standard and attack datasets.

II. BACKGROUND

Several lessons on data mining representations take stayed complete towards advance the flexibility arrangement in current earlier years such arrangements are:

Group altogether information interested in the comparing set before put on a classifier intended for an order. An examination stands completed towards ascertaining the presence of the projected method utilizing the KDD Cup 99 dataset. The creator has projected a crossbreed learning approach dependent on the mix of K-Means grouping and Naïve Bayes arrangement towards advance present oddity-based location abilities in the stretch of exactness, recognition percentage just by way of incorrect alert rate [1].

Enhanced k-means grouping built intrusion detection technique which learns on unlabelled information to recognize new assaults. The aftereffect of investigations arranged the KDD Cup 1999 informational collection demonstrations the progression in identification percentage besides reduction in incorrect optimistic rate and to distinguish incomprehensible intrusion [2].

Fusion IDS dependent scheduled double renowned information removal procedures called arbitrary woods besides k-means. The proposed abnormal recognition strategy utilizes the k-means algorithm as a data mining clustering algorithm to recognize novel intrusion. It catches the organization associations information and converts it to an abnormality discovery dataset by pre-preparing, at that point, information is allocated into homogeneous groups utilizing the k-means algorithm. After that, the irregularity locator decides the strange and ordinary bunches. The framework raises inconsistency abnormal when irregularity bunches are recognized [3].

Another intrusion detection system has been presented utilizing the Minmax K-means grouping procedure which beats the lack of affectability towards beginning focuses in the K-means algorithm then expands the nature of grouping. The investigations happening in the NSL-KDD informational index show that the proposed procedure stands extra proficient than that dependent scheduled the K-means grouping procedure. The strategy takes an advanced identification percentage and brings down the incorrect optimistic location percentage [4].

Another ongoing intrusion detection system utilizing a choice tree approach with an effective information pre-preparing comprising of just 13 highlights. The Calculated execution including recognition rate and memory utilization under a continuous climate. The choice tree characterization algorithm was a reasonable methodology for constant intrusion detection [5].

This paper introduces some data mining techniques i.e., k-means clustering and naive Bayes classification, Enhanced k-means algorithm, random forest, and weighted k-means, k-means clustering, regression trees algorithm, etc. The paper is organized as follows. **Section I** Introduction. **Section II** talks about **the** Background**. Section III** explains previous work done before on data mining in the intrusion detection system. **Section IV** gives a brief about existing methodologies. **Section V** attributes and parameters and how these are affected on data mining techniques. **Section VI** gives the proposed method **Section VII** gives the result and expected result. **Section VIII** Conclude this paper. Finally, **Section IX** gives future Scope.

III. PREVIOUS WORK DONE

In the study literature, many flexibility representations take remained studied towards the offer various data mining techniques arrangements and advance the performance in terms of accuracy and also detection rate.

Z. Muda et al. (2011) [1] have proposed the methodology which will be a group of all information into the comparing cluster before putting on a classifier aimed at clustering reason. A study is completed to assess the presence of the projected method utilizing the KDD Cup '99 dataset. The outcome shows that the projected method achieved improvement regarding exactness, recognition rate with a sensible false alert rate.

Shenghui Wang et al. (2011) [2] present an enhanced k-means grouping constructed intrusion detection system that learns on unlabelled information to recognize new assaults. The reverberation of the trials track arranged the KDD Cup 1999 informational collection demonstrations the headway in identification rate and lessening in fake optimistic percentage beside the capacity to distinguish obscure intrusion.

Reda M. Elbasiony et al. (2013) [3] proposed a half breed system, the peculiarity fragment stands enhanced through supplanting the k-means procedure through an alternative termed the weighted k-means algorithm, besides, it utilizes a projected strategy in picking an irregular group through infusing recognized assaults interested in dubious associations information. The creator's methodologies are assessed ridiculous Discovery and Data Mining (KDD'99) datasets.

Mohsen Eslamnezhad et al. (2014) [4] another intrusion detection system is projected utilizing the Minmax K-means clustering procedure. The examination arranged the NSL-KDD informational index shows that the projected interruption location strategy can improve the Recognition Amount and lessens the Incorrect Optimistic Amount.

Yi Aung et al. (2018) [5] utilize the K-means algorithm and classification and regression trees (CART) algorithm. The motivation behind this paper remains towards demonstrating acceptable exactness in execution examination through while intricacy through utilizing a crossover information mining strategy. This model is confirmed by the KDD'99 informational collection.

IV. EXISTING METHODOLOGY

Many data mining structures take stayed implemented over the previous numerous years. Different procedures are applied aimed at different data mining models i.e. k-means clustering and naive Bayes classification, Improved k-means algorithm, random forest, and weighted k-means, k-means clustering, regression trees algorithm.

A) K-means clustering and naive Bayes classification

Here, the principal objective to use the K-Means clustering method remains towards the part then gathering information into ordinary and assault occurrences. K-Means clustering techniques segment an information dataset interested in k-cluster as indicated by an underlying worth recognized by way of the kernel focuses hooked on to each group's centroids or group focuses. The determined unkind estimation of mathematical information limited inside each group is termed centroids.

$$P(C_i|X) = P(x_1|C_i).P(x_2|C_i)\ldots.\ P(x_n|C_i).\ P(C_i)$$

For this situation, the creator picks k = 3 to group the information hooked on three groups (C1, C2, C3). The U2R and R2L assault designs stand normally very like ordinary cases and one additional group stays utilized to assemble U2R and R2L assaults [1].

B) Enhanced k-means algorithm

In this method, the deficiencies of k-means delicate towards starting focuses besides reliance scheduled a few clustering show the clustering impact under various introductory focuses. As per the principal deficiency of k-means, the creator proposes a way to deal with pick starting focuses towards progress the K-means algorithm. The fundamental thought stands to pick the underlying focuses as decentralized as conceivable to abstain from awful clustering.

$$\varphi_j = \max(r_{ij}) - \min(r_{ij}), 1 \le i \le n$$

$$r_{ij}' = (r_{ij} - \min(r_{ij})) / \varphi_j, i = 1, 2, \cdots, n$$

The enhanced k-means procedure brands around enhancement in the initial phase of the k-means procedure. Structures of intrusion detection dependent on arranged grouping. In a genuine organizational climate, ordinary associations are overwhelmingly bigger than unusual associations [2].

C) Random forest and weighted k-means

In this method, the proposed hybrid structure comprises two stages: the online stage, and the disconnected stage. The online stage is a piece of the abuse location strategy and it is principally liable for contrasting the organization associations information with the created interruption designs. If any intrusion is recognized, an abuse caution will be created, and the assault highlights will be shipped off the arbitrary assault selector segment of the anomaly detection part [3].

D) K-means clustering

This model incorporates essential modules of information assortment, information pre-handling, clustering, marking clusters, and intrusion identification. Distinctive information like kinds of bundles, hosts, and convention subtleties get to the information assortment module and are situated in the information stockroom. At that point, the information is partitioned into preparing and testing informational collections. Minmax K-means clustering algorithm allots to each preparation information towards a particular group. At that point, the groups are marked as one or the other types of an assault under the presumption that the quantity of ordinary associations is limitlessly more than the number of assaults or interruption [4].

E) Regression trees algorithm

In this method, k-means clustering is perhaps the most straightforward strategy to take care of clustering issues. The principal objective is to use the K-means clustering method is to part and gathering information hooked on typical assault cases. It ascertains the group community through utilizing the unkind estimation of the articles in each cluster. In Classification and Regression Trees, Algorithm Decision Trees are utilized in information mining to make a model that predicts the estimation of an objective reliant on the estimations of information the classification and Regressing Trees (CART) philosophy [5].

V. ANALYSIS AND DISCUSSION

The mixture knowledge method improves the precision aimed at a solitary classifier particularly aimed at the KM+NB mix, which demonstrations an increment while lessening the false alert rate [1].

Since a consequence of recognition amount and incorrect optimistic amount in regardless of whether the location rate isn't high beside a fake optimistic percentage is a tiny high, yet the creator actually could reason that interruption discovery-based clustering can recognize obscure assault in reality [2]. The consequences of the abuse identification stage are examined previously and cover the majority of the potential assumptions. The yield of this stage is changed physically to test the effectiveness of the general system under numerous conditions [3].

The Minmax K-means algorithm takes the superlative exhibition once the estimation of detection rate stands just about by way of extensive utilizing could be expected, and the estimation of false-positive rate is pretty much as little as conceivable [4].

In the examination of 10-overlay cross-approval, the effectively grouped occurrence records of choice k-means and classification and regression trees (CART) based methodology [5].

TABLE 1

COMPARISONS BETWEEN DIFFERENT DATA MINING TECHNIQUES

Methods and Techniques	**Characteristics**	
	Advantages	***Disadvantages***
K-means clustering and naive Bayes classification	Assurances merging.	Clustering outliers.
Enhanced k-means algorithm	Easy to implement	Difficult to predict K-Value.
Random forest and weighted k-means	k-Means may Developed clusters than hierarchical clustering	Scaling with number of dimensions
K-means clustering	Simple to implement.	Scaling with the number of dimensions.
Regression trees algorithm	Scales to large data sets.	They Do Not Procedure Encoded Packets

VI. PROPOSED METHODOLOGY

In this proposed strategy, the training stage utilizes a K-means clustering algorithm that appoints each preparation information to a particular cluster. The classification and regression trees strategy are unique of the perceptive presenting methods applied cutting-edge visions, data withdrawal, and Artificial Intelligent. Uncertainty trees the objective variable that can take persistent qualities are called regression trees. The classification and regression trees strategy are utilized to group the dataset as typical or inconsistent. What's more, in the testing stage, the interaction of interruption identification is done which depends on the choice of clustering information and information separating, information pre-preparing, clustering, CART algorithm, and data detection.

Basic steps of an algorithm

Step 1: Training and testing datasets are loaded

Step 2: Data Pre-processing

In this step, the training and testing data undergoes pre-processing.

Step 3: K-means Clustering

The K-means clustering plan is utilized to isolate an assortment of typical and inconsistent assault information that carry on comparatively into a few parts which are known as K-th group centroids.

Step 4: Classification and Regression Trees Algorithm

A classification & Regressing Trees (CART) methodology is used to classify the dataset as normal or anomaly.

Step 5: Intrusion detection

The intrusion detection procedure stays acted in the testing stage. In this interaction, the test information is perceived as an abnormality otherwise an assault category, else, it would be supposed as standard data.

Diagrammatic representation of the proposed method is shown as follows:

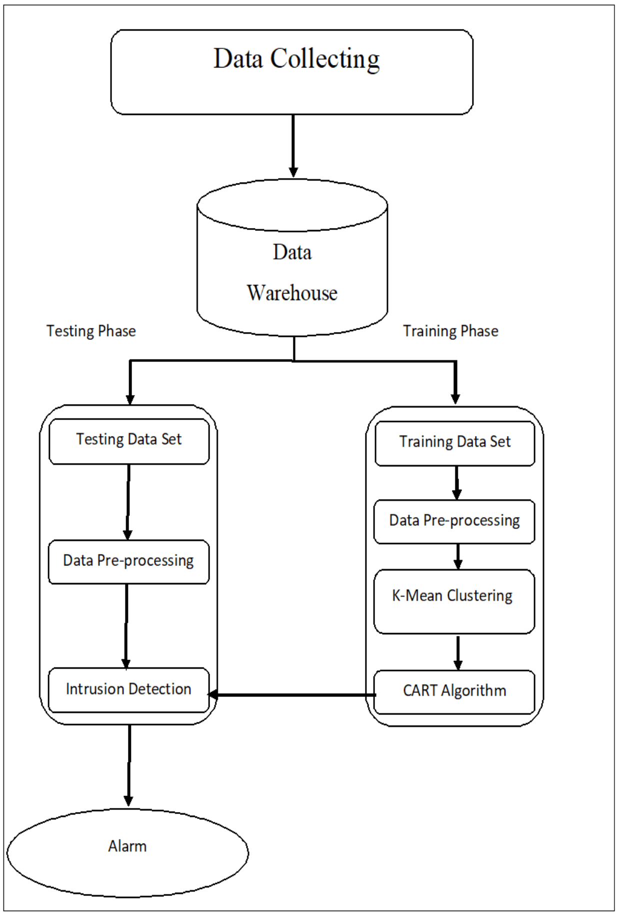
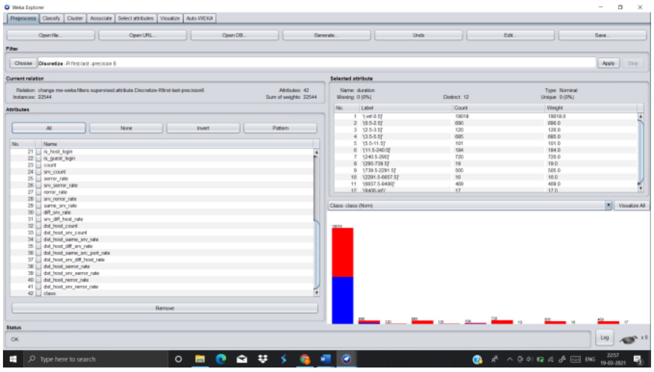


Figure 1. Proposed Model

VII. STIMULATION AND RESULT

The experimental results have shown that the direction of the prediction accuracy in the proposed method is satisfactory and its magnitude is proportional to the reliability and accuracy. The results got from this proposed strategy and approach beat extra through the encouraging detection percentage and decrease a false alarm percentage and give high accurateness.



Filter: Discretize -R first-last -precision 6

Current relation
Relation: change me-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6 Attributes: 42
Instances: 22544 Sum of weights: 22544

Attributes

No.	Name
21	is_host_login
22	is_guest_login
23	count
24	srv_count
25	serror_rate
26	srv_serror_rate
27	rerror_rate
28	srv_rerror_rate
29	same_srv_rate
30	diff_srv_rate
31	srv_diff_host_rate
32	dst_host_count
33	dst_host_srv_count
34	dst_host_same_srv_rate
35	dst_host_diff_srv_rate
36	dst_host_same_src_port_rate
37	dst_host_srv_diff_host_rate
38	dst_host_serror_rate
39	dst_host_srv_serror_rate
40	dst_host_rerror_rate
41	dst_host_srv_rerror_rate
42	class

Selected attribute
Name: duration Type: Nominal
Missing: 0 (0%) Distinct: 12 Unique: 0 (0%)

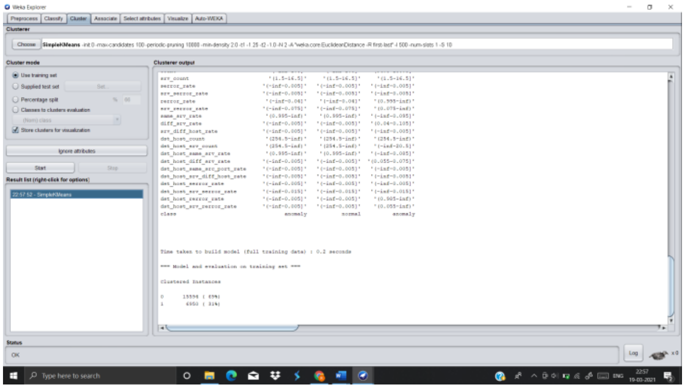
No.	Label	Count	Weight
1	'(-inf-0.5]'	19018	19018.0
2	'(0.5-2.5]'	690	690.0
3	'(2.5-3.5]'	120	120.0
4	'(3.5-5.5]'	685	685.0
5	'(5.5-11.5]'	101	101.0
6	'(11.5-240.5]'	184	184.0
7	'(240.5-295]'	720	720.0
8	'(295-739.5]'	19	19.0
9	'(739.5-2291.5]'	505	505.0
10	'(2291.5-6657.5]'	16	16.0
11	'(6657.5-8406]'	409	409.0
12	'(8406-inf)'	17	17.0

Class: class (Nom)

Status
OK

Figure 2. Datasets uploaded

In fig.2 KDD NSL standard datasets are uploaded in the weka tool. Pre-processing is done with discretize filter which converts real-valued attributes to ordinal attributes.



Clusterer: SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Cluster mode
● Use training set
○ Supplied test set
○ Percentage split
○ Classes to clusters evaluation
☑ Store clusters for visualization

Result list (right-click for options)
22:57:52 - SimpleKMeans

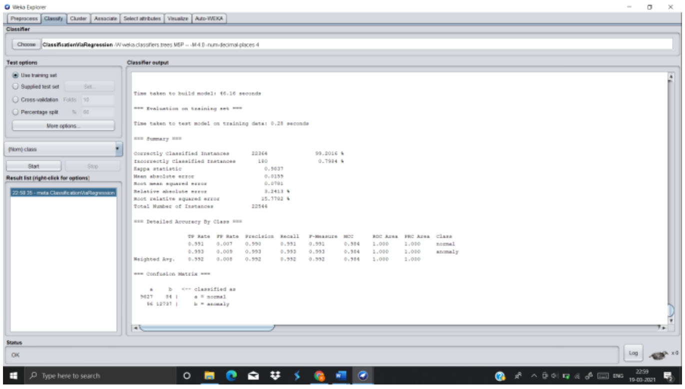
Clusterer output

```
srv_count                    '(1.5-16.5]'     '(1.5-16.5]'     '(1.5-16.5]'
serror_rate                '(-inf-0.005]'   '(-inf-0.005]'   '(-inf-0.005]'
srv_serror_rate            '(-inf-0.005]'   '(-inf-0.005]'   '(-inf-0.005]'
rerror_rate                 '(-inf-0.04]'    '(-inf-0.04]'   '(0.995-inf)'
srv_rerror_rate            '(-inf-0.075]'   '(-inf-0.075]'   '(0.075-inf)'
same_srv_rate               '(0.995-inf)'    '(0.995-inf)'  '(-inf-0.095]'
diff_srv_rate              '(-inf-0.005]'   '(-inf-0.005]'   '(0.04-0.105]'
srv_diff_host_rate         '(-inf-0.005]'   '(-inf-0.005]'  '(-inf-0.005]'
dst_host_count              '(254.5-inf)'    '(254.5-inf)'   '(254.5-inf)'
dst_host_srv_count          '(254.5-inf)'    '(254.5-inf)'   '(-inf-20.5]'
dst_host_same_srv_rate      '(0.995-inf)'    '(0.995-inf)'  '(-inf-0.085]'
dst_host_diff_srv_rate     '(-inf-0.005]'   '(-inf-0.005]' '(0.055-0.075]'
dst_host_same_src_port_rate '(-inf-0.005]'   '(-inf-0.005]'  '(-inf-0.005]'
dst_host_srv_diff_host_rate '(-inf-0.005]'   '(-inf-0.005]'  '(-inf-0.005]'
dst_host_serror_rate       '(-inf-0.005]'   '(-inf-0.005]'  '(-inf-0.005]'
dst_host_srv_serror_rate   '(-inf-0.015]'   '(-inf-0.015]'  '(-inf-0.015]'
dst_host_rerror_rate       '(-inf-0.005]'   '(-inf-0.005]'   '(0.905-inf)'
dst_host_srv_rerror_rate   '(-inf-0.005]'   '(-inf-0.005]'   '(0.055-inf)'
class                            anomaly           normal         anomaly



Time taken to build model (full training data) : 0.2 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      15594 ( 69%)
1       6950 ( 31%)
```

Status
OK

Figure 3. K-means Clustering

In Fig.3 the K-means clustering is applied to the pre-processed data. After K-means clustering the two clusters are formed i.e. cluster 0 and cluster 1.



Classifier: ClassificationViaRegression -W weka.classifiers.trees.M5P -- -M 4.0 -num-decimal-places 4

Test options
● Use training set
○ Supplied test set
○ Cross-validation Folds 10
○ Percentage split % 66

(Nom) class

Result list (right-click for options)
22:58:35 - meta.ClassificationViaRegression

Classifier output

```
Time taken to build model: 46.16 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.28 seconds

=== Summary ===

Correctly Classified Instances       22364               99.2016 %
Incorrectly Classified Instances       180                0.7984 %
Kappa statistic                          0.9837
Mean absolute error                      0.0159
Root mean squared error                  0.0781
Relative absolute error                  3.2413 %
Root relative squared error             15.7792 %
Total Number of Instances            22544

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.991    0.007    0.990      0.991    0.991      0.984    1.000     1.000     normal
                 0.993    0.009    0.993      0.993    0.993      0.984    1.000     1.000     anomaly
Weighted Avg.    0.992    0.008    0.992      0.992    0.992      0.984    1.000     1.000

=== Confusion Matrix ===

     a     b   <-- classified as
  9627    84 |     a = normal
    96 12737 |     b = anomaly
```

Status
OK

Figure 4. CART Classification

In fig.4 the Classification and Regression Trees Algorithm is applied and indicates the normal and anomaly datasets.

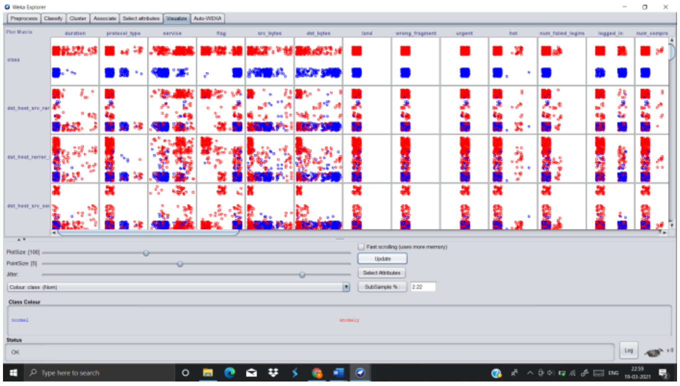


Figure 5. Clustering results

In Fig.5 the graph shows the results of K-means Clustering.

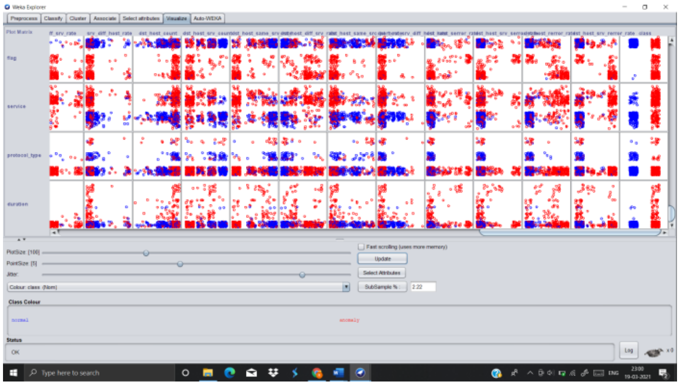


Figure 6. CART Classification results

In Fig.6 the graph shows the results of the Classification and Regression Trees Algorithm.

VIII. CONCLUSION

This proposed strategy utilizes K-means clustering to distinguish between attacks and assaults. Exploratory outcomes demonstrate that the exactness of classification and regression trees algorithm dependent on K-means stays best in the order of ordinary attacks and assaults. The model preparing of K-means and classification and regression trees algorithm depends on intrusion detection system that gives proper information measure of these days interruption data set. Experimental results show that the accuracy of classification and regression trees algorithm based on K-means is good in classification of normal and attacks. And also, the model training time of K-means and classification and regression trees algorithm-based intrusion detection system can provide appropriate time in large data amount of today intrusion database.

IX. FUTURE SCOPE

From observations of the proposed method, the future work will the study proceeding in what way to raise the detection rate for detecting unidentified attacks or new attacks efficiently.

IX. ACKNOWLEDGMENT

It gives immense pleasure to express my gratitude to ***Dr. V. M. Thakare***, ***Dr. S. S. Sherekar*** my guide who provided me constructive criticism and positive feedback during the preparation of this paper. I am indebted to ***Dr. V. M. Thakare***, Head of P. G. Department of computer science and engineering at SGBAU, Amravati, and other teaching and non-teaching staff who were always there whenever I needed any help. Without them and their co-operation, completion of this research work would have been inevitable ad their presence behind me is indispensable.

X. REFERENCES

[1]. Z. Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir, "Intrusion Detection based on K-Means Clustering and Naïve Bayes Classification" Conference: Information Technology in Asia (CITA 11), 2011 7th International Conference,10.1109/CITA.2011.5999520

[2]. Shenghui Wang, "Intrusion detection with unlabeled data using clustering" 2011 Second International Conference on Innovations in

Bio-inspired Computing and Applications, 10.1109/IBICA.2011.72

[3]. Reda M. Elbasiony, Elsayed A. Sallam, Tarek E. Eltobely, Mahmoud M. Fahmy, “A hybrid network intrusion detection framework based on random forests and weighted k-means.” Ain Shams Engineering Journal (2013),10.1016/j.asej.2013.01.003

[4]. Mohsen Eslamnezhad and Ali Yazdian Varjani“ Intrusion Detection Based on MinMax K-means Clustering” 7'th International Symposium on Telecommunications (IST'2014), 10.1109/ISTEL.2014.7000814

[5]. Yi Aung and Myat Myat Min,” Hybrid Intrusion Detection System using K-means and Classification and Regression Trees Algorithms” 2018 IEEE 16th International Conference on Software Engineering Research, Management and Applications (SERA), 10.1109/SERA.2018.8477203

[6]. A. Hadri, K. Chougdali, and R. Touahni, "Intrusion detection system using PCA and Fuzzy PCA techniques," 2016 International Conference on Advanced Communication Systems and Information Security (ACOSIS), Marrakesh, Morocco, 2016, pp. 1-7, doi: 10.1109/ACOSIS.2016.7843930.

[7]. H. M. Tahir, A. M. Said, N. H. Osman, N. H. Zakaria, P. N. '. M. Sabri, and N. Katuk, "Oving K-Means Clustering using discretization technique in Network Intrusion Detection System," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 2016, pp. 248-252, doi: 10.1109/ICCOINS.2016.7783222.

[8]. E. Ariafar and R. Kiani, "Intrusion detection system using an optimized framework based on data mining techniques," 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, Iran, 2017, pp. 0785-0791, doi: 10.1109/KBEI.2017.8324903.

[9]. S. M. A. M. Gadal and R. A. Mokhtar, "Hybrid Method utilizes the Anomaly Detection Technique of data mining technique," 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, 2017, pp. 1-6, doi: 10.1109/ICCCCEE.2017.7867661.

[10]. F. Salo, M. Injadat, A. B. Nassif, A. Shami, and A. Essex, "Data Mining Techniques in Intrusion Detection Systems: A Systematic Literature Review," in IEEE Access, vol. 6, pp. 56046-56058, 2018, doi: 10.1109/ACCESS.2018.2872784

Cite this article as :

R. K. Borikar, Dr. S. S. Sherekar, Dr. V. M. Thakare, "Intrusion Detection System based on K-means, Classification and Regression Trees Algorithm", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 2, pp. 163-171, March-April 2021. Available at doi : https://doi.org/10.32628/CSEIT217239
Journal URL : https://ijsrcseit.com/CSEIT217239