Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 132 (2018) 323–332

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2018)

Rank Allocation to J48 Group of Decision Tree Classifiers using Binary and Multiclass Intrusion Detection Datasets

Ranjit Panigrahi[a]*, Samarjeet Borah[b]

[a]Department of Computer Applications, SMIT, Sikkim Manipal University, Majitar-737136, India
[b]Department of Computer Applications, SMIT, Sikkim Manipal University, Majitar-737136, India

Abstract

The wide acceptance of decision tree classifiers lies with their fast performance and simple nature. The J48 group of decision tree classifiers are widely used for classification and decision-making process. In this paper, three popular J48 group classifiers, namely J48, J48Consolidated and J48Graft are evaluated using both binary and multi-class datasets across thirteen performance matrices, which is unique in its area. In order to come across a versatile classifier, the evaluated results of these classifiers are nourished to a prominent multi-criteria decision-making module called Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) for suitable rank allocation.

Keywords: C4.5, J48, J48Consolidated, J48Graft, TOPSIS, Classifier Rank, ISCXIDS2012, NSL-KDD, Intrusion Detection

1. Introduction

Many appealing discrete and continuous features are inherent to decision trees [1]. The approach of decision tree classifications is considered to be non-parametric in nature. Therefore, it doesn't call for any prior assumptions about the type of probability distributions addressed by the class and other significant attributes. Calculating an ideal decision tree is regarded as an NP-complete problem [2]. The best part of decision trees is that the presence of noise doesn't hamper the performance of the classifier. Likewise, the accuracy of a decision tree isn't affected by the presence of redundant attributes. Various aspects of decision trees have been addressed by many researchers. A brief

* RanjitPanigrahi. Tel.: +91 95479 94659; fax: +91 3592 246112.
E-mail address: ranjit.panigrahi@gmail.com

10.1016/j.procs.2018.05.186

comparison between the C4.5 and ID3 decision tree classifiers was conducted using datasets of varying size [3]. In the study, the C4.5 classifier achieved the highest average accuracy, 87.94%, and the lowest average CPU execution time, 0.13s. Again, the C4.5 classifier was used to classify speech recognition errors [4]. Similarly, using both an unstandardized and a standardized pregnancy dataset, the performance of the C4.5 classifier was evaluated [5]. The C4.5 classifier had 66.08% classification accuracy on the unstandardized pregnancy dataset, whereas the same C4.5 classifier had an accuracy of 71.3% on the standardized pregnancy dataset. On the other hand, an extensive comparison of Decision Forest, Bayesian Network, C4.5, and NBTree classifiers for Course Registration Planning Model of Undergraduate Students was presented, wherein C4.5 languished with the lowest accuracy rate of 81.86% for computer science students [6].

In this research work three widely used decision tree classifiers of J48 groups,such as J48, J48Consolidated and J48Graft are evaluated across thirteen performance matrices using two binary and multi-class datasets. The evaluation was essential because of their wide use in literatures[3,4,5], as these classifiers produces high precision and accuracy rates. To come across a most versatile classifier the performance out comes of these classifiers are subjected to a popular multi-criteria decision-making algorithm called Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) for suitable weight allocation and subsequent ranking. The classifiers ranking is first ever in its kind and a unique approach followed in this research work.

The rest of the paper is arranged as follows: Section 2 highlights all the variations of C4.5 classifiers available in Weka 3.7.0. Section 3 provides a brief overview of the multi-criteria decision-making algorithm TOPSIS used for rank allocation. Section 4 deals with experimental setup followed by the Results and Conclusion in Sections5 and 6, respectively

2. Decision Tree Classifiers

The learning process of decision tree classification is based on the principle of splitting criteria. Decision trees are presented similarly to a flow chart, with a tree structure wherein instances are classified according to their feature values. A node in a decision tree represents an instance meant for classification. The outcomes of the tests are represented by branches, where each leaf node symbolizes the class label. There are various decision tree classifiers available in the literature. Out of those, three classifiers, namely, J48, J48Consolidated, and J48Grafted, are widely used due to their high precision and accuracy rates.

2.1. J48 Classifier

The J48 classifier is an implementation of the C4.5 decision tree algorithm[7]. J48 classifies a new instance by creating a decision tree from the attribute values of the given training set. The moment it come across the training set, it recognizes the attribute that is responsible for categorizing the various instances most accurately. The possible feature values with no ambiguity are assigned to the concern branch by terminating it.

2.2. J48Consolidated Classifier

The J48Consolidated classifier is based on the famous C4.5 decision tree algorithm, with a minor deviation in the knowledge extraction process [7,8]. The mechanism behind the knowledge extraction process is said to be built on the Consolidated Tree Construction (CTC) algorithm [9,10]. Unlike C4.5, J48Consolidated extracts knowledge from data using a set of samples.

2.3. J48Graft Classifier

The J48Graft classifier is based on the C4.5A algorithm [11]. It works on the principle that similar objects are very likely in the same class. Tree grafting achieves better classification models at the cost of producing highly complex trees [12].

3. TOPSIS-Multi-Criteria Decision Making

TOPSIS is one of the widely used classical Multi-Criteria Decision Making (MCDM) methods[13,14]. It is used to select the best from a set of alternatives, each of which is evaluated against multiple criteria. The algorithm allocates weight to each alternative and finally select the best one based on the highest weight. The weight allocation procedure by TOPSIS follows the following steps.

STEP I

Construct normalized decision matrix.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\left(\sum x_{ij}^2\right)}} \tag{1}$$

Where i = 1,…,m, j = 1,…,n

x_{ij}= Original score of decision matrix

r_{ij}= Normalized score of decision matrix

STEP II

Construct weighted normalized decision matrix.

$$V_{ij} = W_j r_{ij} \tag{2}$$

Where W_j =Weight for j criterion

STEP III

Calculate positive and negative ideal solutions.

$$A^* = \{V_1^*, \dots, V_n^*\}$$
$$A' = \{V_1', \dots, V_n'\}$$

Where

A^* = Positive ideal solutions

A' = Negative ideal solutions

$$V_j^* = \{max(V_{ij})\ if\ j\epsilon J;\ min(V_{ij})\ if\ j\epsilon J'\}$$

$$V_j' = \{min(V_{ij})\ if\ j\epsilon J;\ max(V_{ij})\ if\ j\epsilon J'\}$$

STEP IV

Determine the separation measure for each alternative.

A. The separation for positive ideal alternative is

$$S_i^* = \sqrt{\sum\left(V_j^* - V_{ij}\right)^2} \text{ for } i=1,\dots,m \tag{3}$$

B. The separation for negative ideal alternative is

$$S_i' = \sqrt{\sum\left(V_j' - V_{ij}\right)^2} \text{ for } i=1,\dots,m \tag{4}$$

STEP V

A. Calculate the closeness of each alternative to the ideal solution.

$$C_i^* = \frac{S_i'}{S_i^* + S_i'} \{0 < C_i^* < 1\} \tag{5}$$

B. Select the alternative where C_i^*closest to 1.

4. Experimental Setup

The J48 group of classifiers are evaluated using WEKA 3.8, and the dataset is exported to a Microsoft Excel file for better migration to the workspace[15].

4.1. Datasets under Evaluation

To evaluate the J48 group of classifiers, two widely used datasets were considered. Keeping the objective of this research work intact, the ISCXIDS2012 [16] and NSL-KDD [17,18,19] datasets were considered a binary and a multiclass dataset, respectively. These datasets were selected based on their class density, number of instances, and number of attributes.

ISCXIDS2012 is a binary class dataset provided by the University of New Brunswick for evaluating an intrusion detection system[16]. It consists of seven days of network activities and is labelled as normal and malicious. Again, a random collection of 5000 records are considered for evaluation, consisting of 181 instances of minority class with 3.62% density and 4819 instances of majority class with 96.38% density.

Similarly, NSL-KDD is a multi-class dataset[17,18]. This dataset is made available by the University of New Brunswick. The NSL-KDD dataset was recommended to address some of the natural problems of the KDD'99dataset[20]. Though it contains 4,898,431 records, for the purposes of the current experiment, a random total of 5000 records with 14 distinct classes were considered. The minority class has only one instance, with 0.04% density, whereas the majority class has 2678 instances, with 53.56% density.

The detailed characteristics of both ISCXIDS2012 and NSL-KDD are presented in Table 1.

Table 1. Dataset characteristics

Dataset	Type	No. of Instances	No. of Attributes	No. of Distinct Classes	Minority Class		Majority Class	
					Instances	%	Instances	%
ISCXIDS2012	Binary	5000	6	2	181	3.62	4819	96.38
NSL-KDD	Multi-class	5000	40	14	1	0.04	2678	53.56

4.2. Performance matrices

Most studies have used classification accuracy as a performance parameter for the classifiers. However, practically, a classifier's performance can be judged through additional parameters. In the case of a decision tree, the number of leaves and the size of the tree generated by the classifier are also considered good performance parameters. Apart from these basic performance parameters, the following matrices are also considered for optimum evaluation classifiers' performance:

- Build Time (Seconds): This is the time taken to build the model by a classifier using the training set.

- Correctly Classified Instances: This indicates the overall accuracy rate of the classifier. The rate of correctly classified instances can be calculated as

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ instances} \tag{6}$$

- Incorrectly Classified Instances (%): This indicates the overall misclassification rate of the classifier. Therefore, the rate of incorrectly classified instances can be calculated as

$$Misclassification\ Rate = \frac{False\ Positives + False\ Negatives}{Total\ instances} \tag{7}$$

- Mean Absolute Error(MAE): The mean absolute error is the sum of all instances' absolute errors divided by the number of instances in the test set with an actual class label. Mean absolute error is defined as

$$MAE = \frac{1}{n}\sum |x_i - x| \tag{8}$$

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

- Root Mean Squared (RMS) error: RMSE usually provides how far is the model is from giving the right answer. It represents average prediction error within the same scale (unit). It is calculated as

$$RMS = \sqrt{\frac{1}{n}\sum \epsilon_i^2} \tag{9}$$

Where ϵ_i is calculated as $(\widehat{x_i} - x_i)$

- Relative Absolute (RA) error (%): Relative absolute error is the ratio of total absolute error of the actual values with the absolute error of the simple predictor. Relative absolute error is represented as

$$RA = \frac{\sum_{i=1}^{N} |\widehat{x_i} - x_i|}{\sum_{i=1}^{N} |\bar{x} - x_i|} \tag{10}$$

- Root Relative Squared (RRS) error (%): RRS error is the ratio of RMSE with the RMSE obtained by just predicting the mean of target values and then multiplying by 100. Therefore, smaller values are better, and values larger than100% indicate a scheme is doing worse. It is calculated as

$$RRS = \sqrt{\frac{\sum_{i=1}^{N} (\widehat{x_i} - x_i)^2}{\sum_{i=1}^{N} (\bar{x} - x_i)^2}} \tag{11}$$

- True Positive (TP) Rate: True positive rate or sensitivity or recall is defined as

$$TP\ Rate = \frac{TP}{TP+FN} \tag{12}$$

- False Positive (FP) Rate: False positive rate is defined as

$$FP\ Rate = \frac{FP}{FP+TN} \tag{13}$$

- Precision: Precision or Positive Predictive Value (PPV) is calculated as

$$Precision\ or\ PPV = \frac{TP}{TP+FP} \tag{14}$$

- F-Measure: The F-measure or F-score is the harmonic mean between precision and recall.

$$F - Measure = \frac{2(Precision - Recall)}{(Precision + Recall)} \tag{15}$$

4.3. Rank allocation process

The allocation of rank to various J48 group classifiers began with dataset selection and was followed by performance evaluation and finally rank identification. Initially, both binary (ISCXIDS2012) and multi-class (NSL-KDD) datasets were imported to the WEKA workspace separately and tested against various performance matrices discussed earlier. The result of each classifier for each performance matrix was passed to a popular multi-criteria decision-making algorithm, TOPSIS, to allocate rank to the classifier. In TOPSIS, all the performance matrices of classifiers have equal weight (λ), i.e., 1. The TOPSIS criteria signs for all the performance matrices are presented in Table 2. It should be noted that +1 indicates the greater the measure, the better the classifier's result, and –1 indicates the greater the measure, the worse the classifier's result.

Table 2. TOPSIS criteria signs for various classifiers' performance measures

Sl. No	Classifiers Performance Matrices	Criteria sign range	Sl. No.	Classifiers Performance Matrices	Criteria sign range
1	Number of Leaves	+1	8	RA Error (%)	–1
2	Size of the Tree	+1	9	RRS Error (%)	–1
3	Build Time (s)	–1	10	True Positive Rate	+1
4	Accuracy (%)	+1	11	False Positive Rate	–1
5	Misclassification (%)	–1	12	Precision	+1
6	Mean absolute error	–1	13	F-Measure	+1
7	RMS Error	–1			

5. Results and discussion

Table 3 and Table 4 present all the J48 group classifiers' outcomes for 13 different performance matrices employing the binary (ISCXIDS2012) and multi-class (NSL-KDD) datasets.

5.1. Classifiers performance for binary datasets

The binary (ISCXIDS2012) dataset was used to calculate the performance matrices outlined in Table 3 for the J48 group of classifiers.

Table 3. Performance outcome by J48, J48Consolidated, and J48Graft classifier for binary class dataset

Category	Performance Matrices	J48	J48Consolidated	J48Graft
Tree formation related matrices	Number of Leaves	4	2	14
	Size of the tree	7	3	27
	Build Time (Secs)	0.01	0.21	0.04
Classification matrices	Accuracy (%)	99.96	99.8	99.96
	Misclassification (%)	0.04	0.2	0.04
Error rate	Mean absolute error	0.0005	0.0021	0.0005
	RMS Error	0.0191	0.0446	0.0191
	RA Error (%)	0.6658	2.9576	0.6658
	RRS Error (%)	10.2198	23.8985	10.2198
Measure of relevance	True Positive Rate	1	0.998	1
	False Positive Rate	0.005	0.0001	0.011
	Precision	1	0.998	1
	F-Measure	1	0.998	1

It is evident that J48 claimed the very least amount of time, i.e., 0.01 seconds, to build the model, whereas J48Consolidated required 0.21 seconds. Meanwhile, J48Graft took only 0.04 seconds to build the model with the optimum size of the tree, 27, and number of leaves, 14. This is an important feature of both the algorithms, since tree formation time is a major issue in decision tree classification techniques. Thus, it can be concluded and suggested that for binary datasets, J48 and J48Graft are to be used.

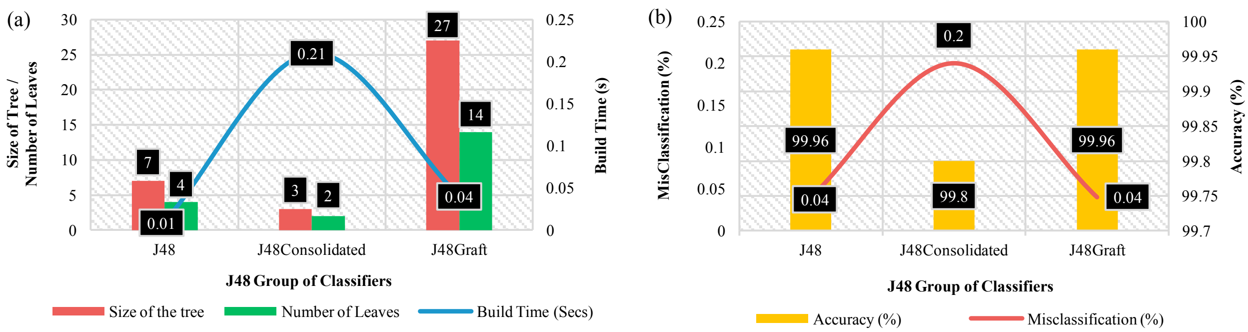

Fig. 1. (a) Tree formation related matrices of J48 group of classifiers for ISCXIDS2012 dataset; (b) Classification matrices of J48 group of classifiers for ISCXIDS2012 dataset

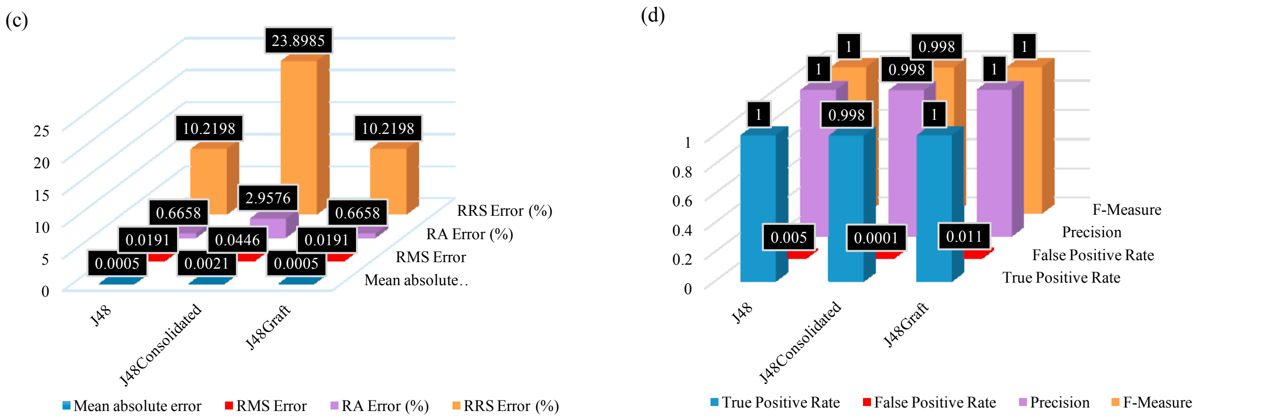

Fig. 1. (c) Error rate of J48 group of classifiers for ISCXIDS2012 dataset; (d) Measure of relevance of J48 group of classifiers for ISCXIDS2012 dataset

Similarly, so far as accuracy is concerned, both J48 and J48Graft showed the highest and same amount of correctly classified instances (99.96%), whereas the J48Consolidated lagged behind in this aspect. Moreover, both J48 and J48Graft showed much lower misclassification rates. This is due to the favorable classification rules for binary data. Let us conduct two more analyses before concluding.

There are several matrices for measuring classification error other than misclassification, which is basically overall incorrect classification rate against total instances. In this analysis, various classification error matrices, such as RRS, RA, MAE, and RMS, are considered.Fig. 1(b) shows that in all forms of classification error matrices, J48 and J48Graft had lower amounts of error compared to J48Consolidated.

In Fig. 1(d), some other important measures of relevance, such as TPR, FPR, Precision, and F-Measure, are compared for all three classifiers with the binary dataset. In this case, surprisingly, J48Consolidated showed improved FPR compared to J48 and J48Graft, whereas J48Graft had the worst FPR. In the other cases, both J48 and J48Graft showed similar effects.

It is clearly realized that one classifier is a better performer in some areas of comparison but fails significantly in other areas too. Therefore, in order to come across a conclusion about a versatile tree classifier for binary dataset the results from the above analysis can be fed into TOPSIS to calculate weight and, subsequently, rank. The calculated weight and rank are represented in Fig. 2.

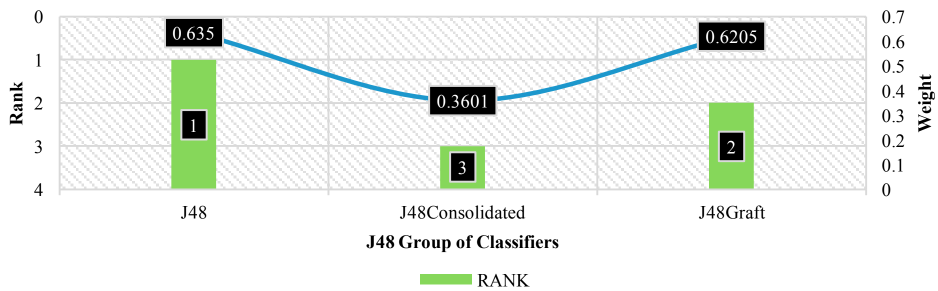

Fig.2. Weight and corresponding rank of J48 group of classifiers for ISCXIDS2012 dataset

5.2. Classifiers Performance for Multi-Class Dataset

In the previous section, a binary dataset was considered for evaluating the J48 group of classifiers. In this section, the same classifiers are tested with a multiclass dataset. The results of the same are consolidated in Table 4. Again, the same is compared graphically based on matrix/index category.

Table 4. Performance Outcome by J48, J48Consolidated, and J48Graft Classifiers for Multi-Class dataset

Category	Performance Matrices	J48	J48Consolidated	J48Graft
Tree formation related matrices	Number of Leaves	187	175	580
	Size of the tree	229	209	1015
	Build Time (Secs)	0.32	5.63	0.44
Classification matrices	Accuracy (%)	98.56	93.88	98.56
	Misclassification (%)	1.44	6.12	1.44
Error rate	Mean absolute error	0.0028	0.0102	0.0028
	RMS Error	0.0436	0.0901	0.0444
	RA Error (%)	3.2438	11.8455	3.2747
	RRS Error (%)	21.097	43.5707	21.4449
Measure of relevance	True Positive Rate	0.986	0.939	0.986
	False Positive Rate	0.009	0.006	0.011
	Precision	0.985	0.965	0.985
	F-Measure	0.985	0.948	0.985

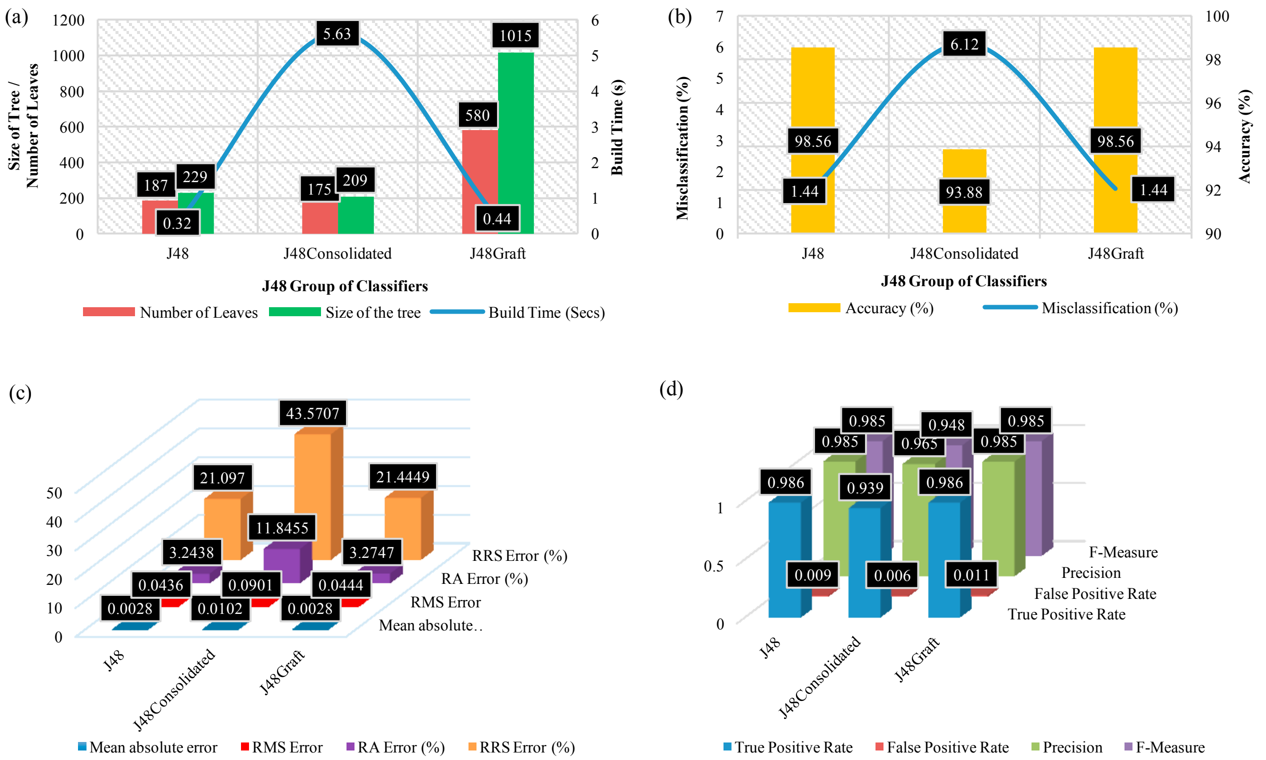

Fig. 3. (a) Tree formation related matrices of J48 group of classifiers for NSL-KDD dataset; (b) Classification matrices of J48 group of classifiers for NSL-KDD dataset; (c) Error rate of J48 group of classifiers for NSL-KDD dataset; (d) Measure of relevance of J48 group of classifiers for NSL-KDD dataset

Fig. 3(a) shows that J48 had the minimum time complexity in the given multi-class dataset. The smallest tree with the minimum number of leaves was generated by J48Graft. J48Graft generated the maximum number of conditions with the minimum time. Therefore, in this case, it was the best performer.

Again, in terms of accuracy and misclassification, both J48 and J48Graft outperformed J48Consolidated.Further, four different types of classification error were tested using the J48 group of classifiers on the multiclass dataset. Here, the lowest values were found with J48 and J48Graft in MAE, similar to binary dataset. The lowest RMS error was shown by J48, and the highest by J48Consolidated. Similarly, J48 outperformed the other two in terms of RA and RRS error. Finally, it can be said that J48 classifies a multiclass dataset with minimum errors.

All the classifiers were again tested for FPR, TPR, precision, and f-measure on the multiclass dataset, the result of which is shown in Fig. 3(d). In this case, both J48 and J48Graft yielded similar results, whereas J48Consolidated had the finest rates in all cases.

Finally, all the results from the experiment on the multiclass dataset were consolidated and fed to TOPSIS to compute the weight and assign rank. Fig. 4 represents the weight and corresponding rank of tree classifiers, where J48Graft found to be the best performer.

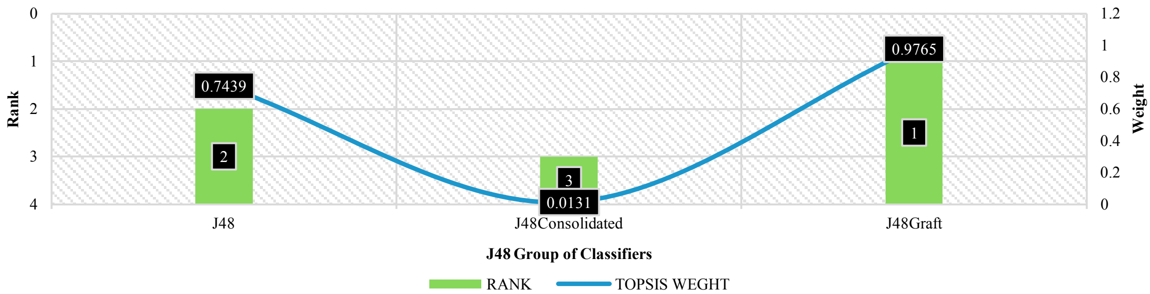

Fig. 4. Weight and corresponding rank of J48 group of classifiers for NSL-KDD dataset

6. Conclusion

In this paper, the J48 group of classifiers were considered for ranking. This kind of study is essential to apply a specific classification technique directly to a domain for an optimal result. Two widely used datasets from the security domain were selected for the said purpose as a binary and multiclass data. J48 group of classifiers were rigorously tested in various test conditions and matrices related to tree formation, classification accuracy, error indices, detection rate, etc. The test results were sent to TOPSIS to ascertain weight and, subsequently, rank. J48 achieved the highest rank and outperformed both J48Consolidated and J48Graft for the binary dataset, while J48Graft performed the best in the multi-class dataset and thus acquired the highest rank.

As a future study, it would be a better idea to test these classifiers with datasets of varying dimensions and densities. Datasets from diversified domains could also be considered for this kind of evaluation. Other ranking methods, such as FuzzyTOPSIS and VIKOR, could be employed to justify the ranks allocated to these classifiers. Moreover, the similar ranking mechanism can be extended for other group, such as Bayes, function based, lazy, meta and rule-based classifiers.

References

[1] Piltaver,Rok, Luštrek,Mitja, Gams,Matjaž and Martinčić-Ipšić,Sanda (2016) "What makes classification trees comprehensible?", Expert Systems with Applications62(16): 333-346.

[2] Tan, Pang-Ning, Steinbach, Michael and Kumar Vipin (2006) "Introduction to Data Mining", Pearson Addison-Wesley, United States

[3] Hssina, Badr, Merbouha, Abdelkarim,Ezzikouri, Hanane and Erritali, Mohammed(2014) "A comparative study of decision tree ID3 and C4.5",International Journal of Advanced Computer Science and Applications, Special Issue on Advances in Vehicular Ad Hoc Networking and Applications:13-17.

[4] Cerňak, Milos (2010) "A Comparison of decision tree classifiers for automatic diagnosis of speech recognition errors",Computing and Informatics29(3): 489–501.

[5] Lakshmi, BN, Indumathi, TS, Nandini, R (2016), "A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy", Procedia Technology24: 1542-1549.

[6] Meedech, Phanupong ,Iam-On, Natthakan and BoongoenTossapon(2008) "Comparisons of classifier algorithms: Bayesian network, C4.5, decision forest and NBTree for Course Registration Planning model of undergraduate students", IEEE International Conference on Systems, Man and Cybernetics5: 3647-3651.

[7] Quinlan J Rose (1993) "C4.5: Programs for Machine Learning", Morgan Kaufmann Publishers Inc.(eds), San Mateo, California

[8] Arbelaitz, Olatz, Gurrutxaga, Ibai, Lozano, Fernando, Muguerza, Javier, Pérez, Jesus Maria M (2013) "J48Consolidated: An implementation of CTC algorithm for WEKA", Technical Report EHU-KAT-IK-05-13, University of the Basque Country (UPV/EHU). http://www.sc.ehu.es/aldapa/weka-ctc/Weka-CTC.pdf, Accessed on21Dec, 2017

[9] Ibarguren,Igor, Pérez,Jesús M., Muguerza,Javier, Gurrutxaga,Ibai and Arbelaitz,Olatz2015) "Coverage-based resampling: Building robust consolidated decision trees",Knowledge Based Systems79(C):51-67.

[10] M.Pérez,Jesús, Muguerza,Javier, Arbelaitz,Olatz, Gurrutxaga,Ibai and I.Martín, José (2007) "Combining multiple class distribution modified subsamples in a single tree", Pattern Recognition Letters28(4): 414-422.

[11] Webb, Geoffrey I (1999) "Decision Tree Grafting from the All-Tests-But-One Partition",Proc. 16th International Joint Conference on Artificial Intelligence 2:702–707.

[12] Hayashi,Yoichi, Tanaka,Yuki, Takagi,Tomohiro,Saito,Takamichi, Iiduka,Hideaki, Kikuchi,Hiroaki, Bologna,Guido andMitra,Sushmita(2016) "Recursive-rule extraction algorithm with J48graft and applications to generating credit scores", Journal of Artificial Intelligence and Soft Computing Research6(1): 35-44.

[13] Hwang,Ching-Lai and Yoon Kwangsun(1981) "Multiple Attributes Decision Making",Methods and Applications, Springer, Berlin Heidelberg

[14] AmeriSianaki, Omid (2015) "Intelligent Decision Support System for Energy Management in Demand Response Programs and Residential and Industrial Sectors of the Smart Grid". Dissertation, Curtin University

[15] Mark,Hall, Eibe,Frank, Geoffrey,Holmes, Bernhard,Pfahringer, Reutemann,Peter and Witten,Ian H. (2009),"The WEKA Data Mining Software: An Update",SIGKDD Explorations11(1):10-18.

[16] Shiravi,Ali, Shiravi,Hadi, Tavallaee,Mahbod and A.Ghorbani,Ali(2012) "Toward developing a systematic approach to generate benchmark datasets for intrusion detection", Computers & Security31(3): 357-374.

[17] Ibrahim,Laheeb M, Basheer,Dujan T andMahmod,Mahmod S (2013) "A Comparison Study for Intrusion Database (KDD99 NSL-KDD) Based on Self Organization Map (SOM) Artificial Neural Network",Journal of Engineering Science and Technology8(1): 107-119

[18] Ingre, Bhupendra& Yadav, Anamika. (2015) "Performance analysis of NSL-KDD dataset using ANN", Signal Processing and Communication Engineering Systems (SPACES) 92-96

[19] Aggarwala, Preeti and Sharma ,Sudhir Kumar(2015) "Analysis of KDD Dataset Attributes - Class wise for Intrusion Detection", Procedia Computer Science57:842-851

[20] Tavallaee,Mahbod, Bagheri,Ebrahim, Lu,Wei, Ghorbani, Ali A. (2009) "A Detailed Analysis of the KDD CUP 99 Data Set". Second IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA).