Journal of King Saud University – Computer and Information Sciences xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Music rhythm tree based partitioning approach to decision tree classifier

Shankru Guggari [a,*], Vijayakumar Kadappa [c], Umadevi V. [a], Ajith Abraham [b]

[a] Department of Computer Science and Engineering, B.M.S. College of Engineering, Bull Temple Road, Bengaluru, India
[b] Machine Intelligence Research Labs (MIR Labs), Scientific Network for Innovation and Research Excellence, P.O. Box 2259, Auburn, Washington 98071, USA
[c] Department of Computer Applications, B.M.S. College of Engineering, Bull Temple Road, Bengaluru, India

ARTICLE INFO

Article history:
Received 29 November 2019
Revised 12 March 2020
Accepted 31 March 2020
Available online xxxx

Keywords:
Decision tree
Music rhythm tree
Class imbalance
Vertical partitioning

ABSTRACT

Decision tree is a widely used non-parametric technique in machine learning, data mining and pattern recognition. It is simple to understand and interpret, however it faces challenges such as handling higher dimensional and class imbalanced datasets, over-fitting and instability. To overcome some of these issues, vertical partitioning approaches like serial partitioning, theme based partitioning are used in the literature. A vertical partitioning approach divides the feature set into subsets of features (blocks) and makes use of these subsets for subsequent tasks. In this work, we use the ideas of music rhythm tree to propose a novel vertical partitioning technique. It orders the features based on the average correlation strength of the features before partitioning the feature set. The proposed method is proved to be superior by showing an average of $13.8\%, 6\%, 9.8\%, 19.7\%, 9.4\%$, and 29.4% higher classification accuracy over C4.5, Random Forest, Bagging, Adaboost, an ensemble technique and a vertical partitioning technique respectively. Our empirical results on 15 datasets demonstrate that the proposed vertical partitioning method is more stable and better in handling class-imbalanced data. Finally, some popular statistical tests are conducted to validate the statistical significance of the results of the proposed method.

1. Introduction

Decision trees are well-known classification techniques that are used in various real world applications. Prediction of a given data instance is performed based on its learnt experience. They are applied in various domains such as predicting success rate of the students (Natek and Zwilling, 2014), security assessment (Oliveira et al., 2017), price prediction (Liu et al., 2017) and electrochemical systems (Erdem Günay et al., 2018). Decision trees are simple to understand and can be easily interpreted by the users. However, they suffer from various challenges like handling high-dimensionality data, overfitting, instability, and class-imbalance problem.

Decision trees are popularly known as unstable because a small change in training dataset leads into extreme variation in the tree structure and produce different prediction. Metrics such as dissimilarity measure, misclassification rate, number of terminal nodes, standard deviation, dissimilarity measure and depth of the tree are used to quantify the structural stability of the decision tree. Decision trees have two kinds of stabilities - Semantic and Structural stabilities. Two classifiers are said to be semantically stable if they show the same prediction rate for similar data, whereas structural stability measures whether two classifiers have similar topology (Mirzamomen and Kangavari, 2016). An ensemble-trees technique (Zimmermann, 2008) is proposed to avoid both overfitting and structural instability of the decision tree. In other study, semantic stability is proposed based on feature selection (Paul et al., 2012).

Classification of imbalanced dataset is an important issue in data mining, machine leaning and pattern analysis. Classification accuracy of the decision tree technique is immensely affected from class imbalanced datasets. Several ensemble techniques are proposed to handle such datasets. Recently, re-sampling and AdaBoost ensemble classifiers are used to classify five groups of heart beats based on AAMI standard EC57:1998. Shigang Liu et al. discuss twitter spam detection (Liu et al., 2016) for class imbalanced data using fuzzy based oversampling, random oversampling, and undersampling techniques. Support vector machine, k-Nearest Neighbour (kNN), Nave Bayes (NB) and Random Undersampling Boost (RUSBoost) are used to build the models and combine the predictions of each model based on majority voting technique. The metrics like, F-measure, class imbalance rate, true positive rate and false positive rate are used to measure the performance of a classifier

* Corresponding author.
E-mail addresses: shankru.16pej@bmsce.ac.in (S. Guggari), vijaykumark.mca@bmsce.ac.in (V. Kadappa), umadevi.cse@bmsce.ac.in (V. Umadevi), ajith.abraham@ieee.org (A. Abraham).

https://doi.org/10.1016/j.jksuci.2020.03.015
1319-1578/

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

for class-imbalanced datasets. Similarly, Real value Negative Selection Oversampling (RNSO) (Tao et al., 2019) technology is used to generate artificial minority data to improve classification accuracy. G-Mean and F-Measure are the metrics used to evaluate the performance of the eleven datasets from UCI repository.

To overcome these issues, partitioning approaches are proposed in the literature (Kumar and Minz, 2015; Guggari et al., 2018; Rokach, 2008). Feature partitioning is a useful method for developing a novel ensemble with diversified classifiers. There are two types of partitioning approaches: (1) Dimensions based partitioning technique (a.k.a. Vertical partitioning), where subsets are obtained based on the feature set, (ii) Data instances based partitioning technique (a.k.a. Horizontal partitioning), where subsets of data instances are created (Kumar and Minz, 2015). It is observed that partitioning methods extensively use local information to recognize a pattern (Kumar and Minz, 2015; Seetha and Narasimha Murty, 2016; Rokach and Maimon, 2006; Kumar and Minz, 2017; Rokach, 2008) and able to produce higher recognition rate as compared to classical approaches. Multi-view approach (Rokach, 2008), serial partitioning (Seetha and Narasimha Murty, 2016) and theme based partitioning (Vijayakumar et al., 2015) methods are some of the vertical partitioning techniques in the literature.

Present study focuses on bringing out a novel vertical partitioning technique using the ideas from *Music rhythm* tree to improve the efficiency for class imbalanced datasets, structural stability and classification accuracy of the decision tree. Music rhythm tree is a simple structure to describe the relationship between notes (Whole, Half, Quarter notes etc.). Music rhythm trees are successfully applied in varied contexts like, processing of term rewriting system (Jacquemard et al., 2015), Machine based music composition (Dostál, 2013), integer ratios (Boenn, 2018), expression of emotions (Alexander et al., 2015) and Electroencephalography (EEG)(Balasubramanian et al., 2018) etc. Recently, a study has been carried out to understand behavioural and neural features of humans using music rhythms and reveals the similarities in music, speech and animal communication (Kotz et al., 2018). The proposed partitioning method shows improved classification, and higher structural stability as compared to other well-established approaches. We also investigate the applicability of the proposed method on class-imbalanced datasets.

The rest of the paper is structured into 7 sections as follows: A detailed related work is discussed in Section 2 and we introduce our proposed method in Section 3. Experimental results and analysis are described in 4–7 Finally, we conclude in Section 8.

2. Related work

In this section, we discuss some key related works which are proposed in the literature for vertical partitioning to address the instability and class imbalance problem of the decision tree techniques. Initially Kusik introduces partitioning techniques in mechanical industry based on the data types of the features (Kusiak and Larson, 1995). In other work, feature partitioning is carried out based on Meta learning, where classifier selects the features depending on the characteristics of the data set (Rokach, 2006). Lior Rokach et al. show quality improvement in manufacturing using feature set decomposition method. It makes use of Breadth-Oblivious-Wrapper search technique to select the features (Rokach and Maimon, 2006). Similarly, in other work, classification of web pages is performed based on feature partitioning and uses Co-training method which makes use of both labelled and unlabelled data instances (Blum and Mitchell, 1998). Evolutionary algorithm such as genetic algorithm is used to select the features and Vapnik-Chervonenkis dimension bound to assign the features into a particular partition (Rokach, 2008).

An ensemble technique known as multi-view ensemble learning (MEL) improves the classification accuracy for both low (Kumar and Minz, 2015) and high dimensional (Kumar and Minz, 2015) datasets. The MEL method uses Naive Bayes (NB), k-Nearest Neighbour(kNN) and Support Vector Machine (SVM) methods to classify the data instances. In another work, Vijayakumar et al. divide the features of the data set based on the themes. Identification of the themes depends on the knowledge of the domain, for e.g., if we identify research and work experience are the themes in teacher recruitment dataset. The features related to research theme, number of publications (National/International), R & D Projects and number of consultancy works etc., are assigned to the same block (Vijayakumar et al., 2015). More recently, an evolutionary algorithm known as constrained particle swarm optimization method (OMEL-C-PSO) is used to select the features and partitions the features based on Bell number for high dimensional datasets. The performance of the method is evaluated using ten high dimensional datasets from UCI repository and Support vector machine is used to classify the data instances with the help of 10-fold cross-validation procedure (Kumar and Minz, 2017). Non-sequential partitioning methods are proposed to design decision tree classifiers using Ferrer diagram and Bell triangle concepts. These methods are applied on low and medium dimensional datasets from UCI repository and indicate performance improvement in terms of classification accuracy and structural stability (Guggari et al., 2018). In another study, an ensemble technique is designed for both low and medium dimensional datasets. It describes classification error rate with the uniform distribution of features space and uses over 60 datasets to conduct experiments. Spearman's correlation coefficient is used to find the relation between error and variance and the study indicates stronger relation for medium datasets (Cervantes et al., 2018). More recently, partition selection is performed to improve the image classification based on the content using sparse autoencoders. It identifies salient features that help in recognition of image and uses min-max normalization fusion technique. Superiority of the method is established by comparing the results with the SVM and extreme learning machine. It demonstrates the results by using both tenfold cross-validation and leave one out methods (Das and Walia, 2019).

Instability is an important issue with decision trees. Braiman is the first researcher who identified this problem in decision trees and used bagging technique to resolve it. Bagging is an ensemble of classifiers and combine them using majority voting technique to identify the class label (Leo, 1996; Leo, 1996). Similarly, Info-fuzzy network is proposed to produce stable decision tree and is compared with meta learning scheme (Last et al., 2002). In other study, Fuzzy min-max decision tree-HB (FMMDT-HB) is proposed to improve the structural stability and nodes are split based on the Hoeffding bound (Mirzamomen and Kangavari, 2016). A boosting technique, called as cross-split technique, is used to stabilize the decision tree (Mirzamomen et al., 2015). Baranauskas observes that the dataset with less than 5 classes provides stable decision trees (Baranauskas, 2015). Structural stability is improved based on region compatibility for the decision tree. It uses probability assignments of evidence theory to measure the stability (Wang et al., 2018).

The ensemble learning is gaining popularity to solve class imbalance problem (Zhang et al., 2018). Synthetic neighbourhood sample generation technique is used to rebalance the dataset. It adds a synthetic sample for a majority class and multiple samples to a minority class. Superiority of the method is evaluated with 5-fold cross validation technique using KEEL datasets (Chen et al., 2018). Paired ensemble technique is used to address both class

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

imbalance and concept drift problems. It uses two classifiers - a long term stable classifier and a dynamic classifier. Further, 10 real world along with 31 synthetic datasets are considered to evaluate the performance using area under the ROC curve (AUC) measure (Zhang et al., 2018). An evolutionary inversion method is presented to address multi-class imbalance learning. It selects few instances from majority classes in the overlapping areas and uses N1byClass measure (overlap with a class with other class in percentage) to address overlapping areas (Zhang et al., 2019). Similarly, multiple classifier system is introduced to solve the classification problem. It uses homogeneous classifiers with bagging sampling technique where 50% of the data instances and 50% of the features are used to build the model. Weighted majority voting technique is implemented to combine the output of the classifies (Mohammed et al., 2020). In another work, a multi-matrices ensemble technique is performed to address imbalanced problem. It uses entropy as an objective function to filter the instances to obtain better decision boundary for the base classifier. It uses 55 KEEL binary classification datasets to demonstrate the novelty of the method (Wang et al., 2019).

2.1. Brief review of benchmark methods

In this section, we introduce short summary of the popular decision tree techniques such as Classification and Regression Trees (CART), C4.5 and C5.0. The CART method generates regression trees and predict a real number but not a class. It splits based on the minimum squared error and use weighted mean for a node to predict a class (Rokach, 2001). The C4.5 method uses information gain ratio to build the tree and reduce over-fitting problem using pessimistic pruning (Xindong et al., 2008). Similarly, C5.0 is the improved version of C4.5 with enhancement of boosting and helps to prune the nodes which are not useful to improve the classification accuracy (winnowing) (C5.0, 1993).

The well-known ensemble of models (Adaboost, Bagging, etc.) are used to compare with our MRTPDT method in terms of classification accuracy, structural stability and handling class imbalanced data. Adaboost is a popular ensemble technique, where it focuses on each training set based on the weights. At the beginning, similar weight is assigned to each data instance. Further, it decreases the weight to correctly classified instances and increases the weight to all misclassified data instances at each iteration. Bagging is a composite classifier, where each classifier is trained from sample instances with replacement. Random Forest is an extension of bagging technique. It randomly selects a subset of features and construct an unpruned decision tree for each subset (Rokach, 2010). For comparison, we use an ensemble technique (Catal et al., 2015) which is combination of C4.5 (J48), Logistic regression, and Multi-layer perceptron classifiers. The vertical partitioning technique (Seetha and Narasimha Murty, 2016), a sequential partitioning approach, wherein the support vector machine (SVM) is used as a base classifier. It divides the features of the dataset sequentially into mutual exclusive subsets.

3. Music rhythm tree based partitioning technique to decision tree classifier (MRTPDT)

In this section, we propose Music Rhythm Tree based Partitioning Technique (MRTPDT), a vertical partitioning approach, where the feature set is partitioned into non-empty and mutual exclusive subsets (blocks) based on the ideas of music rhythm tree technique (Sebö and Waksman, 1999). It logically divides the feature set into subsets with different characteristics.

Let F = $\{F_1, F_2, \ldots, F_m\}$ denotes the set of m features, and $C = \{C_1, C_2, \ldots, C_s\}$ denotes the set of s class labels. The algorithm of the proposed method is given as follows:

3.1. Algorithm

1. Compute correlation coefficient, r_x^y, for the features, F_x and F_y, where $\forall x, y = 1, 2, \ldots, m; x < y$ as given by:

$$r_x^y = \frac{cov(F_x, F_y)}{\sigma(F_x)\sigma(F_y)} \tag{1}$$

and find A^y, the average correlation value of the feature $F_y, \forall y = 1, 2, \ldots, m$ (Step-1 of Fig. 1) as given by:

$$A^y = \frac{1}{m}\sum_{x=1}^{m} r_x^y \tag{2}$$

2. Arrange the average correlation values of the features in ascending order (Step-2 of Fig. 1). Let $f_1, f_2, \ldots, f_m$ be the features after re-arrangement. Here f_i stores an original feature, $F_j, i, j = 1, 2, \ldots, m$.

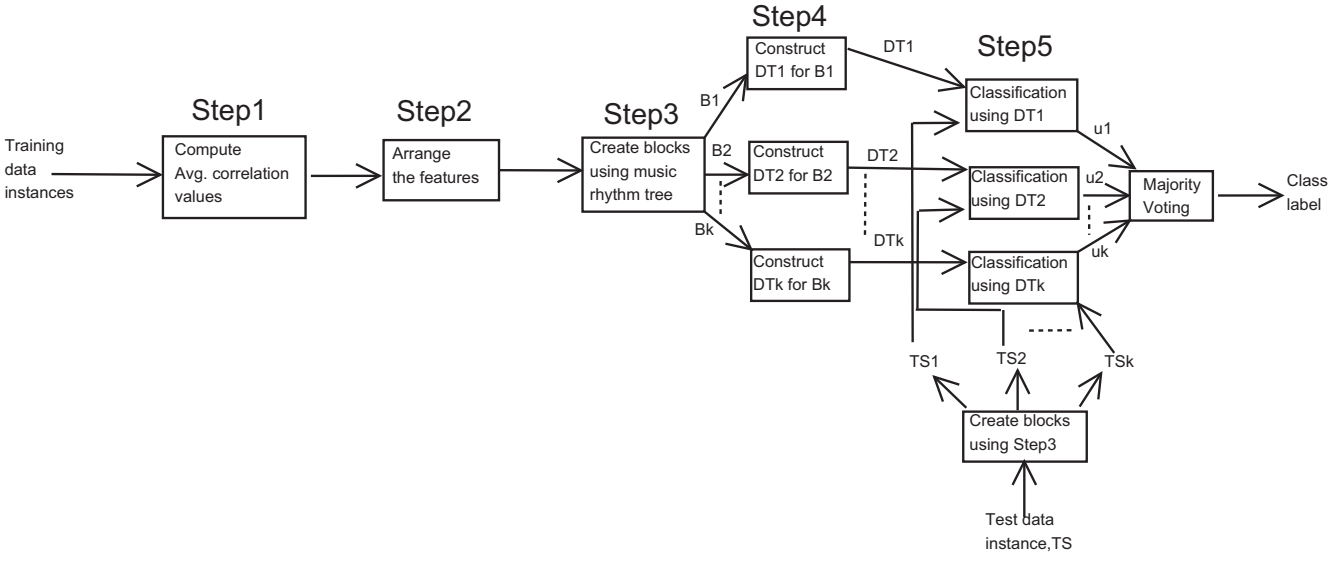


Fig. 1. Visualization of the proposed MRTPDT method.

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

Table 1
Characteristics of the Datasets.

Sl. No.	Data set	No. of Features	No. of Instances	No. of Classes
1	MicroMass (Aha and Murphy, 1994)	1025	871	21
2	Colon Tumor (Andres Cano and Andres, 2005)	2000	62	2
3	Thyroid (Zhu et al., 2010)	2000	168	4
4	SRBCT (Zhu et al., 2010)	2309	83	4
5	Lymphoma (Zhu et al., 2010)	4027	66	3
6	Gisette (Guyon, 2003)	5000	6000	2
7	NIC60 (Zhu et al., 2010)	6115	61	8
8	Central Nervous System (Andres Cano and Andres, 2005)	7130	60	2
9	DLBCL Harvard (Andres Cano and Andres, 2005)	7130	58	2
10	Duke Breast Cancer (Chih-Chung and Chih-Jen, 2011)	7130	44	2
11	Leukemia (Andres Cano and Andres, 2005)	7130	38	2
12	Lung Cancer (Andres Cano and Andres, 2005)	7130	96	2
13	Arcene (Guyon, 2003)	10000	100	2
14	Brain (Zhu et al., 2010)	10368	50	4
15	Prostate (Andres Cano and Andres, 2005)	12600	136	2

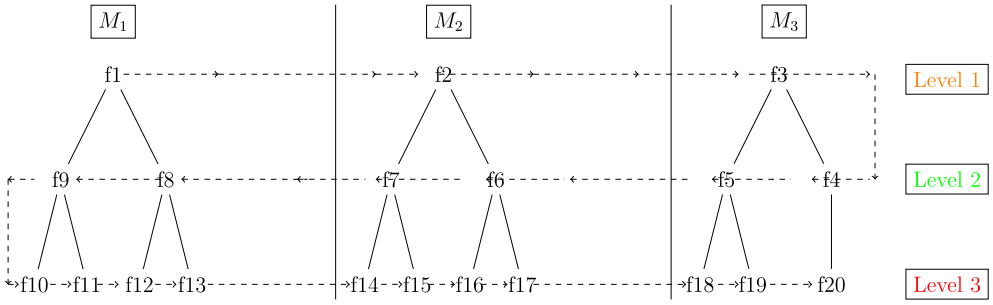


Fig. 2. Music rhythm tree based vertical partitioning for 20 features. Here, M_1, M_2 and M_3 represent music rhythm trees. Each tree represents a block.

3. Use music rhythm tree technique to create k blocks (partitions) (Step-3 of Fig. 1), B = $\{B_1, B_2, B_3, \ldots, B_k\}$ as shown in Fig. 2. Each block is a subset of original feature set.
4. Construct Decision Trees, DT_i, for each of k blocks, $B_i, i = 1, 2, \ldots, k$, using training data instances and develop a collection of decision trees (Step 4 of Fig. 1).

$$DT_i \longleftarrow Build\_DT(B_i) \tag{3}$$

5. Use the ensemble of decision trees of step 4 to classify a test instance, *TS*. (Fig. 1):
   5.1. Build k blocks, $TS_1, TS_2, \ldots, TS_k$, for a test data instance, *TS* as described in step 3.
   5.2. Use DT_i to classify TS_i, a test data block and stores the class label C_l in u_i

$$u_i \leftarrow DT_i(TS_i) \tag{4}$$

   5.3. $u \leftarrow$ majority-vote $(u_1, u_2, \ldots, u_k)$, where u represents the class label assigned to *S*.

3.2. The vertical partitioning concept using music rhythm tree (Step-3 of the proposed method)

The feature set partitioning using Music rhythm tree pattern is elaborated in this section. Music rhythm tree is a hierarchical structure used to represent the information in a tree pattern. Its notations are based on a logarithm scale: a whole note is the double of the half note and it is two times more than a quarter note. A non-linear tree structure is used to represent the pattern of a melody, where tree contains a root node which is denoted by a whole note and two child nodes are represented by two half notes. Similarly, each note is split into two notes recursively. This type of tree coding of melodies perform better than the string coding (Rizo and Quereda, 2002; Rizo et al., 2652). In another study, tree structure is used to represent the polyphonic music (Rizo et al., 2009).

Music rhythm binary trees (say $M_i, i = 1, 2, \ldots, k$) are created as follows: Here, we consider the feature order as obtained in step 2 of the algorithm (Section 3.1). The features are assigned to music rhythm tree in snake-movement fashion as shown in Fig. 2. At level 1, f_i is assigned to a music rhythm tree, M_i (Fig. 2). At level 2, two features are assigned to a music rhythm tree, M_i and so on. At level l, a maximum of 2^l features are assigned to a music rhythm tree. For example, consider a dataset which has 20 features, $f = \{f_1, f_2, f_3, \ldots, f_{20}\}$ and suppose we create 3 blocks using music rhythm tree partitioning approach as shown in Fig. 2. The features in these 3 blocks are shown as follows: $B_1 = \{f_1, f_8, f_9, f_{10}, f_{11}, f_{12}, f_{13}\}, B_2 = \{f_2, f_6, f_7, f_{14}, f_{15}, f_{16}, f_{17}\}, B_3 = \{f_3, f_4, f_5, f_{18}, f_{19}, f_{20}\}$. Please note that the music rhythm tree, M_i, represents the block, B_i.

4. Results and analysis

In this section, experiments are carried out on standard benchmark datasets to demonstrate the performance of the proposed method in comparison to the established classification techniques

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

(CART, C4.5, C5.0, Bagging, Adaboost, an ensemble technique (Catal et al., 2015) and a vertical partitioning technique (Seetha and Narasimha Murty, 2016)). Detailed explanation on experimentation set-up and analysis of the results is given in the following subsections:

4.1. Benchmark datasets used

Effectiveness of the MRTPDT method is analysed on 15 benchmark datasets from UCI repository (Aha and Murphy, 1994), Kent Ridge Bio-medical datasets (Andres Cano and Andres, 2005) and NIPS 2003 workshop challenge datasets (Guyon, 2003). The characteristics of the datasets used are listed in the Table 1 with number of samples, number of dimensions (features) and number of target classes. These datasets are high-dimensional in nature, with the range of 1025 to 12600 features.

4.2. Experimental set-up

Features of each dataset are vertically partitioned from 5 to 100 blocks (with a step of 5) based on the concept of music rhythm tree. We estimate the average classification accuracy, standard

Table 2
Classification accuracies of MRTPDT + CART method, CART, Bagging, AdaBoost, an Ensemble technique and a Vertical partitioning technique.

Sl. No.	Datasets	Classification rate (%)						
		CART	Random Forest	Bagging	Adaboost	Ensemble Technique (Catal et al., 2015)	Vertical partitioning technique (Seetha and Narasimha Murty, 2016)	MRTPDT + CART
1	MicroMass (Aha and Murphy, 1994)	73.38	**84.52**	81.46	17.23	84.85	78.21	81.27
2	Colon Tumor (Andres Cano and Andres, 2005)	65	79.14	76.12	79.31	73.55	64.76	**80.71**
3	Thyroid (Zhu et al., 2010)	66.1	71.01	69.78	55.1	73.06	53.52	**76.1**
4	SRBCT (Zhu et al., 2010)	76.53	98.35	95.38	56.36	99.76	45.14	**100**
5	Lymphoma (Zhu et al., 2010)	75.48	89.5	79.83	82.19	**96.57**	53.57	83.33
6	Gisette (Guyon, 2003)	92.07	**96.66**	95.53	88.53	64.34	68.71	94.72
7	NIC60 (Zhu et al., 2010)	26.67	49.9	47.26	16.71	47.57	47.33	**50**
8	Central Nervous System (Andres Cano and Andres, 2005)	63.33	62.67	59.5	57.83	61.83	65	**66.67**
9	DLBCL Harvard (Andres Cano and Andres, 2005)	50	54.77	53.3	58.63	50.5	55.33	**69.67**
10	Duke Breast Cancer (Chih-Chung and Chih-Jen, 2011)	68.5	79.25	81.95	79.95	84.55	52	**91**
11	Leukemia (Andres Cano and Andres, 2005)	85	86.83	80.33	87	86.92	70.83	**97.50**
12	Lung Cancer (Andres Cano and Andres, 2005)	96.89	96.99	95.54	97.03	96.59	10.14	**100**
13	Arcene (Guyon, 2003)	73	**78.5**	71.5	69.2	72.2	54	78
14	Brain (Zhu et al., 2010)	62	73.8	58.8	51	63.2	58	**76**
15	Prostate (Andres Cano and Andres, 2005)	87.64	87.54	86.65	88.25	84.65	63	**93.52**
	Average	**70.77**	**79.30**	**75.53**	**65.62**	**76.01**	**55.97**	**82.57**

Table 3
Classification accuracies of MRTPDT + C4.5 method, C4.5, Bagging, AdaBoost, an Ensemble technique and a Vertical partitioning technique.

Sl. No.	Datasets	Classification rate (%)						
		C4.5	Random Forest	Bagging	Adaboost	Ensemble Technique (Catal et al., 2015)	Vertical partitioning technique (Seetha and Narasimha Murty, 2016)	MRTPDT + C4.5
1	MicroMass (Aha and Murphy, 1994)	77.23	84.52	81.46	17.23	84.85	78.21	**86**
2	Colon Tumor (Andres Cano and Andres, 2005)	81.19	79.14	76.12	79.31	73.55	64.76	**84.05**
3	Thyroid (Zhu et al., 2010)	54.23	71.01	69.78	55.1	73.06	53.52	**74.38**
4	SRBCT (Zhu et al., 2010)	86.81	98.35	95.38	56.36	99.76	45.14	**100**
5	Lymphoma (Zhu et al., 2010)	66.43	89.5	79.83	82.19	**96.57**	53.57	93.81
6	Gisette (Guyon, 2003)	93.67	96.66	95.53	88.53	64.34	68.71	**97.5**
7	NIC60 (Zhu et al., 2010)	30	49.9	47.26	16.71	47.57	47.33	**58.33**
8	Central Nervous System (Andres Cano and Andres, 2005)	60	62.67	59.5	57.83	61.83	65	**71.67**
9	DLBCL Harvard (Andres Cano and Andres, 2005)	55	54.77	53.3	58.63	50.5	55.33	**62.67**
10	Duke Breast Cancer (Chih-Chung and Chih-Jen, 2011)	72.5	79.25	81.95	79.95	84.55	52	**93**
11	Leukemia (Andres Cano and Andres, 2005)	87.5	86.83	80.33	87	86.92	70.83	**97.5**
12	Lung Cancer (Andres Cano and Andres, 2005)	99	96.99	95.54	97.03	96.59	10.14	**100**
13	Arcene (Guyon, 2003)	71	78.5	71.5	69.2	72.2	54	**83**
14	Brain (Zhu et al., 2010)	60	73.8	58.8	51	63.2	58	**82**
15	Prostate (Andres Cano and Andres, 2005)	78.85	87.54	86.65	88.25	84.65	63	**96.46**
	Average	**71.56**	**79.30**	**75.53**	**65.62**	**76.01**	**55.97**	**85.36**

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

Table 4
Classification accuracies of MRTPDT + C5.0 method, C5.0, Bagging, AdaBoost, an Ensemble technique and a Vertical partitioning technique.

Sl. No.	Datasets	Classification rate (%)						
		C5.0	Random Forest	Bagging	Adaboost	Ensemble Technique (Catal et al., 2015)	Vertical partitioning technique (Seetha and Narasimha Murty, 2016)	MRTPDT + C5.0
1	MicroMass (Aha and Murphy, 1994)	76.35	84.52	81.46	17.23	84.85	78.21	**85.64**
2	Colon Tumor (Andres Cano and Andres, 2005)	**82.38**	79.14	76.12	79.31	73.55	64.76	80.95
3	Thyroid (Zhu et al., 2010)	56.58	71.01	69.78	55.1	73.06	53.52	**75.7**
4	SRBCT (Zhu et al., 2010)	86.81	98.35	95.38	56.36	99.76	45.14	**100**
5	Lymphoma (Zhu et al., 2010)	64.76	89.5	79.83	82.19	**96.57**	53.57	90.48
6	Gisette (Guyon, 2003)	94.08	96.66	95.53	88.53	64.34	68.71	**97.45**
7	NIC60 (Zhu et al., 2010)	28.33	49.9	47.26	16.71	47.57	47.33	**54.52**
8	Central Nervous System (Andres Cano and Andres, 2005)	60	62.67	59.5	57.83	61.83	65	**66.67**
9	DLBCL Harvard (Andres Cano and Andres, 2005)	48	54.77	53.3	58.63	50.5	55.33	**66**
10	Duke Breast Cancer (Chih-Chung and Chih-Jen, 2011)	70	79.25	81.95	79.95	84.55	52	**91**
11	Leukemia (Andres Cano and Andres, 2005)	**95**	86.83	80.33	87	86.92	70.83	**95**
12	Lung Cancer (Andres Cano and Andres, 2005)	98	96.99	95.54	97.03	96.59	10.14	**100**
13	Arcene (Guyon, 2003)	66	78.5	71.5	69.2	72.2	54	**85**
14	Brain (Zhu et al., 2010)	64	73.8	58.8	51	63.2	58	**84**
15	Prostate (Andres Cano and Andres, 2005)	87.64	87.54	86.65	88.25	84.65	63	**96.37**
	Average	**71.86**	**79.30**	**75.53**	**65.62**	**76.01**	**55.97**	**84.58**

deviation, misclassification rate and F-score using 10-fold cross validation technique. All experimental results are performed on Intel i5 core processor with 12 GB RAM and used R tool (version 3.6.0) (The R project for statistical computing, 1993) which is running on window 7 operating system. In addition, the experimental results of Bagging and AdaBoost methods are obtained using WEKA Tool (Hall et al., 2009).

4.3. Classification accuracies of MRTPDT method against traditional Decision tree methods, Bagging, AdaBoost, an Ensemble technique (Catal et al., 2015) and a Vertical partitioning technique (Seetha and Narasimha Murty, 2016)

Tables 2–4 describe average classification accuracies of MRTPDT method, CART, C4.5, C5.0 and other classification techniques respectively. For each dataset, experiments are conducted by varying partitions from 5 to 100 with a step of 5 and chosen highest classification accuracy obtained by MRTPDT method among these partitions.

MRTPDT + CART method uses CART method to build an ensemble of tree models. MRTPDT + CART shows approximately an average of 11.8% improvement in the classification accuracy as compared to classical CART decision tree across all standard benchmark datasets as indicated in Table 2. It is observed that MRTPDT + CART shows better accuracy for maximum number of datasets and exhibits 23% improvement in classification accuracy for SRBCT and NIC60 datasets over CART method. It also shows improvement in classification accuracy over other techniques. It records 3.3%, 7%, 16.9%, 6.6% and 26.6% improvement in average classification accuracy over Random forest, Bagging, Adaboost, an ensemble technique (Catal et al., 2015) and a vertical partitioning technique (Seetha and Narasimha Murty, 2016) respectively.

Table 3 compares the performance of MRTPDT + C4.5 with other known standard techniques. It shows an average of 13.8% improvement in classification accuracy as compared to C4.5 decision tree and uses C4.5 decision tree to build tree models. MRTPDT + C4.5 shows approximately more than 9% improvement in classification accuracy as compared the Ensemble technique (Catal et al., 2015). Similarly the proposed method outperforms Random Forest, Bagging, Adaboost, and vertical partitioning (Catal et al., 2015) techniques by showing approximately 6%, 9.9%, 19.7%, 29.4% improvement over them respectively in classification accuracy.

Average classification accuracies of MRTPDT + C5.0 are described in Table 4. It exhibits an average of 12.7% improvement in classification accuracy as compared to C5.0 decision tree. It uses C5.0 to build tree models. MRTPDT + C5.0 shows highest improvement of 25.7% and 26.2% in average classification accuracy for Lymphoma and NIC60 datasets and does not show any improvement in classification accuracy for both colon tumor and Leukemia datasets. MRTPDT + C5.0 records 5.3%, 9.1%, 19%, 8.6% and 28.6% improvement in classification accuracy for Random Forest, Bagging, Adaboost, an ensemble Technique (Catal et al., 2015) and a vertical partitioning technique (Seetha and Narasimha Murty, 2016) respectively.

It observed that MRTPDT + C4.5 and MRTPDT + C5.0 methods achieve 100% classification accuracy for both SRBCT (Zhu et al., 2010) and Lung Cancer (Andres Cano and Andres, 2005) datasets.

Figs. 3–5, indicate the average classification accuracies of the proposed method with varied number of partitions from 5 to 100. Here, average is computed by considering the classification accuracies obtained for all datasets for each block number. The proposed method outperforms other methods in terms of classification accuracy. It shows that 15, 25 and 20 partitions have highest average classification accuracies for MRTPDT + CART, MRTPDT + C4.5 and MRTPDT + C5.0 decision trees respectively. Figs. 3–5, also give direction to decide the right number of partitions based on classification accuracy.

4.3.1. Statistical significance of average classification accuracies

In this section, we study the statistical significance of average classification accuracies obtained by MRTPDT method in comparison to other methods. Friedman, Wilcoxon sign, ANOVA, Spearman's and Pearson statistical tests are explored for our analysis. We use $p - value$ as a indicator to gauge the significance of the

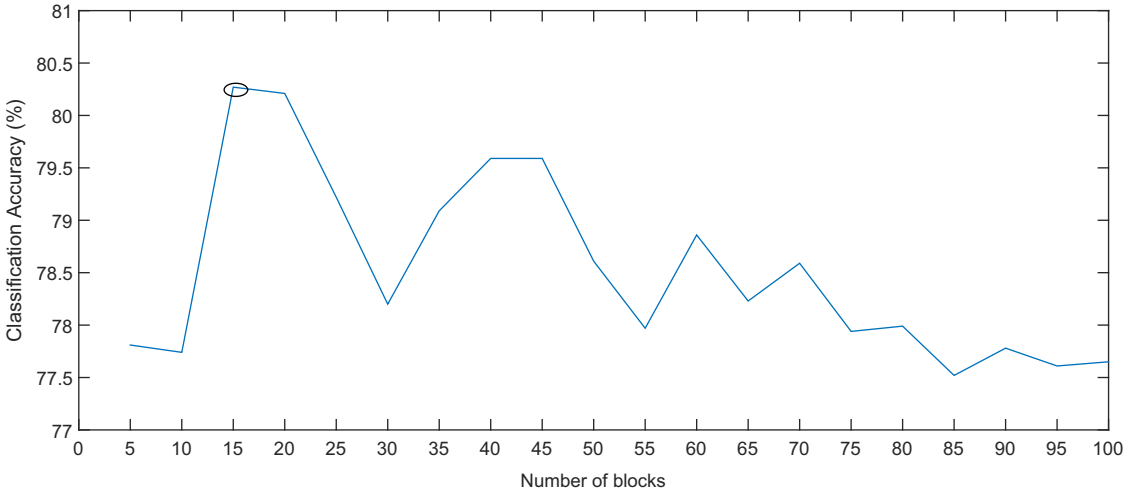


Fig. 3. Average Classification accuracies of MRTPDT + CART method across 5 to 100 blocks.

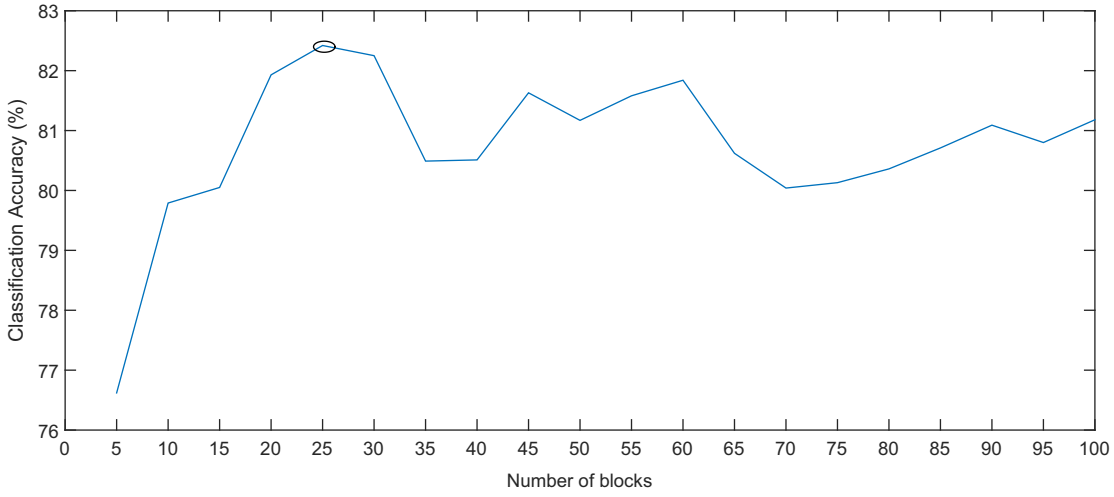


Fig. 4. Average Classification accuracies of MRTPDT + C4.5 method across 5 to 100 blocks.

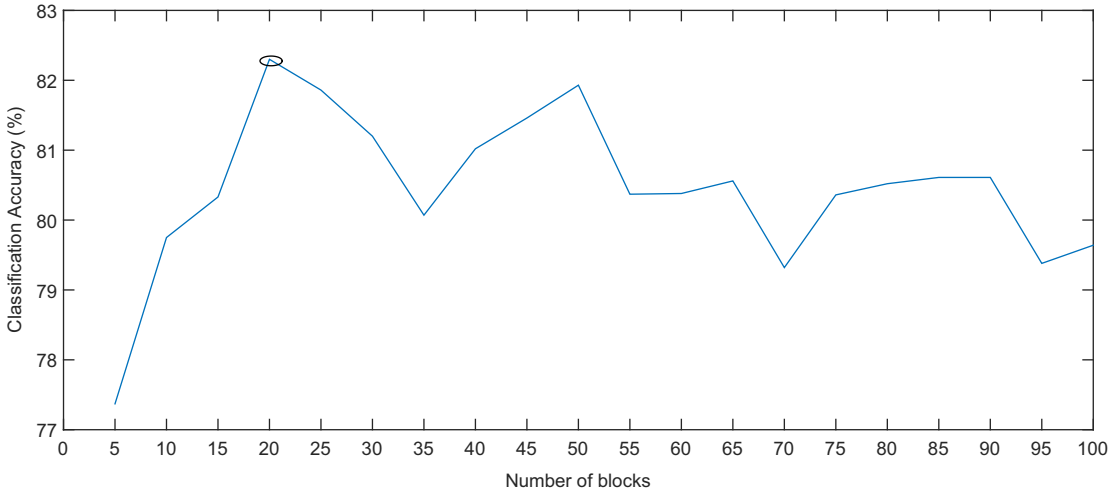


Fig. 5. Average Classification accuracies of MRTPDT + C5.0 method across 5 to 100 blocks.

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

tests and level of significance, $\alpha = 0.05$ in this analysis. If $p - value \leqslant \alpha$, the null hypothesis is rejected. The classification accuracies obtained by MRTPDT and other methods, as shown in Tables 2–4, are used to conduct statistical tests. The $p - values$ obtained by various tests are shown in Tables 5–7. It is evident that the proposed method shows *significant improvement statistically* in classification over other methods because $p < \alpha$ for most of the cases (Tables 5–7).

5. Experimental analysis related to the stability of decision tree

In this section, we used standard deviation and misclassification rate measures to describe the structural stability of the proposed technique. If we say a method is more stable when it shows smaller standard deviation and misclassification rate values.

5.1. Standard deviation

Standard deviation metric is used to assess the structural stability of the proposed MRTPDT method. Fig. 6 presents standard deviation of classification accuracies for 15 datasets (Tables 2–4) obtained by MRTPDT + CART, MRTPDT + C4.5, MRTPDT + C5.0, and other methods using boxplots. It is noted that the mean and other summary values of standard deviation, related to the MRTPDT method are less than the summary values related to other benchmark techniques. Therefore, MRTPDT method is structurally stable.

Tables 8–10 present $p - values$ computed based on the standard deviation values of the proposed and other methods. It is clear that $p < 0.05$ for most of the statistical tests. In other words, the proposed method shows statistically significant values of standard deviation over other methods.

5.2. Misclassification rate

Fig. 7 demonstrates an average misclassification rates of MRTPDT and other techniques for 15 datasets. The average misclassification rate of MRTPDT + CART method is 17.4%, which is less than CART (i.e. 29.2%) decision tree. Similarly, MRTPDT + C4.5 and MRTPDT + C5.0 show much lower misclassification rate as compared to other benchmark methods. Fig. 7 gives an evidence of 11.8%, 13.8% and 12.7% decrease in misclassification rate for MRTPDT + CART, MRTPDT + C4.5 and MRTPDT + C5.0 respectively as compared to classical decision trees (CART, C4.5, C5.0). Therefore, the proposed method is more stable as it shows lower misclassification rate.

6. Analysis of experimental results for class imbalanced datasets

F-score measure, a harmonic mean, is used to evaluate the performance of class imbalanced datasets. MRTPDT technique uses Synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002) to increase the samples of minority class, then applies 10-fold cross validation to classify the test data.

Table 11 represents characteristics of class imbalanced datasets considered for experimental analysis with class distribution and class imbalance ratio. Fig. 8 shows distribution of average F-score

Table 5
Various Statistical tests for MRTPDT + CART and other methods using average classification accuracies of 15 datasets.

		p-value					
Sl. No.	Tests	CART	Random Forest	Bagging	Adaboost	Ensemble Technique	Vertical Partitioning Technique
1	Friedman Test	0.0001	**0.070**	0.0045	0.00010	0.0045	0.00010
2	Wilcoxon Sign Test	0.0006	**0.060**	0.0012	0.0006	0.0006	0.0006
3	ANOVA	3.56e−07	1.10e−07	6.18e−07	3.91e−07	3.4e−06	**0.597**
4	Spearmans Test	0.1.71e−09	6.823e−10	2.73e−08	0.005704	6.23e−07	**0.766**
5	Pearson Test	1.756e−08	6.85e−07	2.77e−07	5.704e−07	6.25e−07	**0.766**

Table 6
Various Statical tests for MRTPDT + C4.5 and other methods using average classification accuracies of 15 datasets.

		p-value					
Sl. No.	Tests	C4.5	Random Forest	Bagging	Adaboost	Ensemble Technique	Vertical Partitioning Technique
1	Friedman Test	0.00010	0.00010	0.00010	0.00010	0.00078	0.00010
2	Wilcoxon Sign Test	0.027	**0.2539**	0.0464	0.011	**0.1465**	4.788e−05
3	ANOVA	9.54e−06	3.24e−09	3.8e−07	0.0074	3.15e−06	**0.705**
4	Spearmans Test	3.308e−06	4.481e−10	3.045e−07	5.64e−07	5.4e−06	**0.8593**
5	Pearson Test	9.545e−05	3.245e−09	3.796e−07	0.0073	3.15e−07	**0.7047**

Table 7
Various Statistical tests for MRTPDT + C5.0 and other methods using average classification accuracies of 15 datasets.

		p-value					
Sl. No.	Tests	C5.0	Random Forest	Bagging	Adaboost	Ensemble Technique	Vertical Partitioning Technique
1	Friedman Test	0.0013	0.00010	0.00010	0.00010	0.00078	0.00010
2	Wilcoxon Sign Test	**0.0618**	**0.2997**	**0.088**	0.020	**0.1465**	5.72e−05
3	ANOVA	4.77e−06	5.47e−09	7.48e−07	0.0086	0.0001	**0.643**
4	Spearmans Test	1.073e−05	1.513e−08	8.537e−08	0.00815	0.00048	**0.737**
5	Pearson Test	4.766e−06	5.469e−09	7.478e−07	0.0086	0.00010	**0.6434**

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

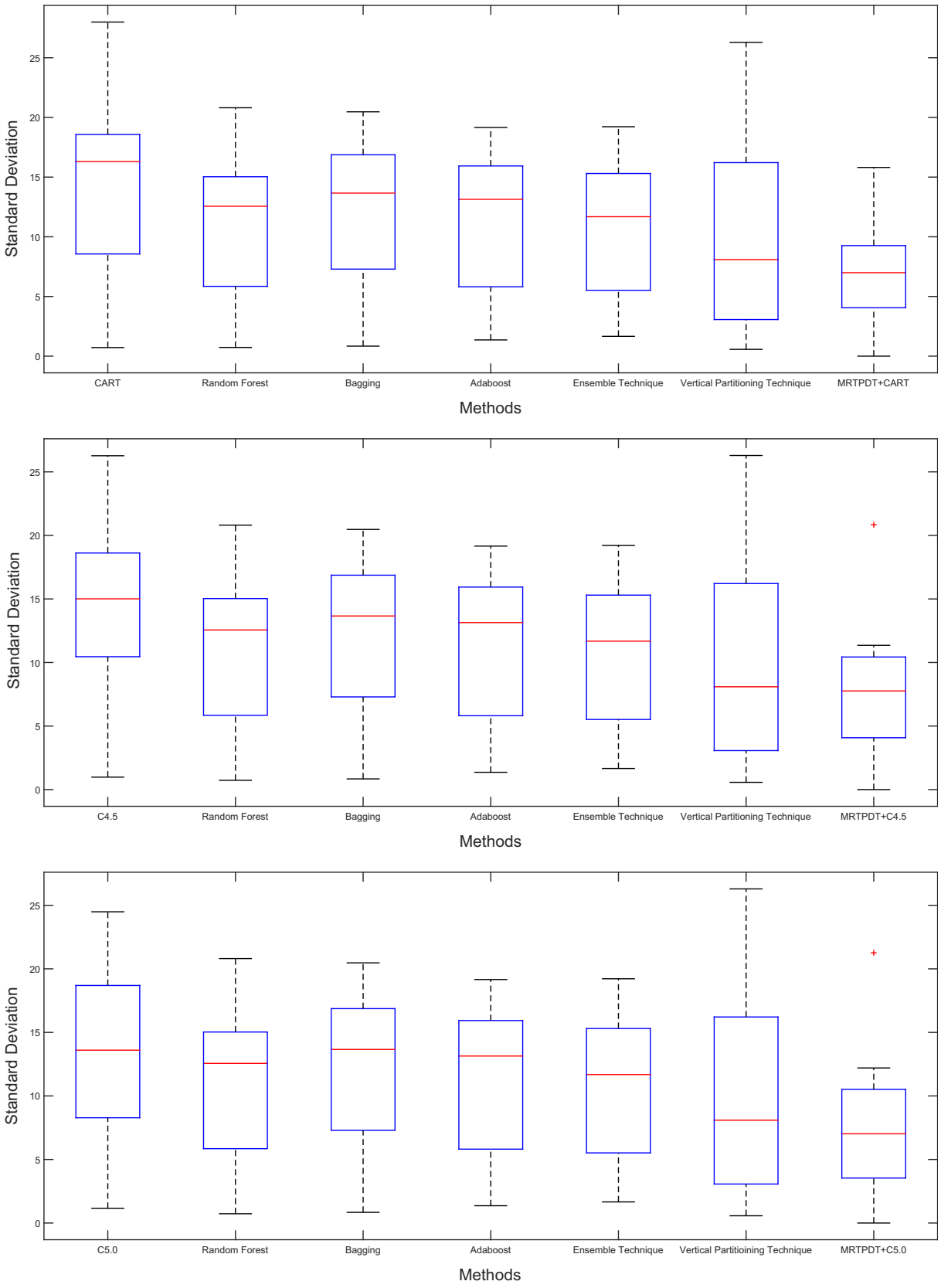


Fig. 6. Standard deviation of classification accuracies obtained by MRTPDT and other methods.

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

Table 8
Various statistical tests for MRTPDT + CART and other methods using standard deviation values.

		p-value					
Sl. No.	Tests	CART	Random Forest	Bagging	Adaboost	Ensemble Technique	Vertical Partitioning Technique
1	Friedman Test	0.00078	0.00078	0.00010	0.020	0.00010	**0.07**
2	Wilcoxon Sign Test	0.0042	0.03	0.014	0.031	**0.056**	**0.25**
3	ANOVA	0.004	6.63e−07	1.282e−06	0.0257	8.9e−07	0.03
4	Spearmans Test	0.0010	3.98e−06	1.06e−06	0.012	3.68e−07	0.0031
5	Pearson Test	0.0040	6.357e−08	0.00028	0.025	8.94e−08	0.030

Table 9
Various Statical tests for MRTPDT + C4.5 and other methods using standard deviation values.

		p-value					
Sl. No.	Tests	C4.5	Random Forest	Bagging	Adaboost	Ensemble Technique	Vertical Partitioning Technique
1	Friedman Test	0.00010	0.0007	0.00010	0.020	0.00010	0.07
2	Wilcoxon Sign Test	0.010	**0.07**	0.042	**0.07**	0.105	0.0042
3	ANOVA	0.0021	5.34e−06	7.8e−05	0.037	1.3e−06	0.024
4	Spearmans test	9.87e−06	1.09e−05	7.292e−07	0.0087	4.92e−07	9.51e−06
5	Pearson test	0.0021	5.39e−07	7.8e−06	0.037	1.38e−05	0.024

Table 10
Various Statistical tests for MRTPDT + C5.0 and other methods using standard deviation values.

		p-value					
Sl. No.	Tests	C5.0	Random Forest	Bagging	Adaboost	Ensemble Technique	Vertical Partitioning Technique
1	Friedman Test	0.0013	0.00078	0.0007	0.0075	0.00078	**0.7963**
2	Wilcoxon Sign Test	0.013	**0.081**	0.042	**0.071**	0.081	**0.29**
3	ANOVA	0.002	2.67e−05	3.71e−06	0.035	3.71e−05	0.027
4	Spearmans Test	0.0011	3.27e−06	0.0019	0.007	2.26e−06	0.010
5	Pearson Test	0.0024	2.66e−05	3.15e−06	0.035	3.71e−05	0.027

values across different blocks for MRTPDT + SMOTE + CART, MRTPDT + SMOTE + C4.5 and MRTPDT + SMOTE + C5.0 decision trees. It is observed that MRTPDT + SMOTE + CART, MRTPDT + S MOTE + C45 and MRTPDT + SMOTE + C5.0 show highest F-score for 20 blocks, 45 blocks and 15 blocks respectively. Fig. 8, shows the approximate number of blocks suitable for the proposed method can be obtained.

Fig. 9 indicates average F-score values obtained by various methods for 3 different datasets mentioned in Table 11. It is evident that the performance of MRTPDT method is better than SMOTE + classical decision tree and classical decision tree methods. For each dataset, we compute average of F-scores obtained by the proposed method by varying the number of blocks from 5 to 100 blocks with a step of 5. For NIC60 data, it is observed from Fig. 9a that the MRTPDT + SMOTE + CART method shows an improvement of 39% and 30% in F-score over CART and SMOTE + CART respectively; similarly, it shows 34% and 12%, 8% and 15% F-score improvement for Colon tumor dataset and Central nervous system datasets respectively. It is noted from Fig. 9b that the MRTPDT + SMOTE + C4.5 method shows an improvement of 45% and 39% in F-score over CART and SMOTE + CART respectively; similarly, it shows 7% and 8%, 2% and 16% F-score improvement for Colon tumor dataset and Central nervous system datasets respectively. It is observed from Fig. 9c that the MRTPDT + SMOTE + C5. 0 method shows an improvement of 44% and 36% in F-score over CART and SMOTE + CART respectively; similarly, it shows 8% and 7%, 2% and 6% F-score improvement for Colon tumor dataset and Central nervous system datasets respectively.

Fig. 10 indicates the average time taken by proposed MRTPDT method as compared to traditional classifiers. It shows an average time improvement by 8.2% 20.9% and 35.6% milliseconds to MRTPDT + CART, MRTPDT + C4.5 and MRTPDT + C5.0 respectively. Therefore, the proposed MRTPDT method is superior as compared to traditional decision trees in terms of time complexity (run time).

7. Discussion

The proposed methods – MRTPDT + CART, MRTPDT + C4.5 and MRTPDT + C5.0 – show highest classification accuracy for 15, 25 and 20 blocks respectively, among all the blocks (Figs. 3–5). According to the definition mentioned in the work (Vipin and Sonajharia, 2016), the partition with highest classification accuracy suggests an approximate number of blocks to be created for a dataset. Therefore, Figs. 3–5 give a direction to choose right number of blocks for the proposed method. Experimental analysis demonstrates that the blocks having the combination of features with high, moderate and low correlation values are helpful in the improvement of classification accuracy of the proposed method (Tables 2,3,4, and Figs. 6, 7 and 9). The music rhythm tree enables us to create such combinations of features in each block. Our experimental results demonstrate that the proposed method is also scalable for high dimensional datasets. The proposed method is more efficient as compared to other benchmark techniques for the following reasons: Ordering of the features based on the strength of correlation values and music rhythm tree is helpful to produce blocks having features with a combination of high, moderate and low correlation values.

The proposed algorithm has some drawbacks: It is slower than other ensemble techniques in terms of computational complexity and tree depth size of the MRTPDT is comparatively higher than other state-of-the-art ensemble techniques.

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

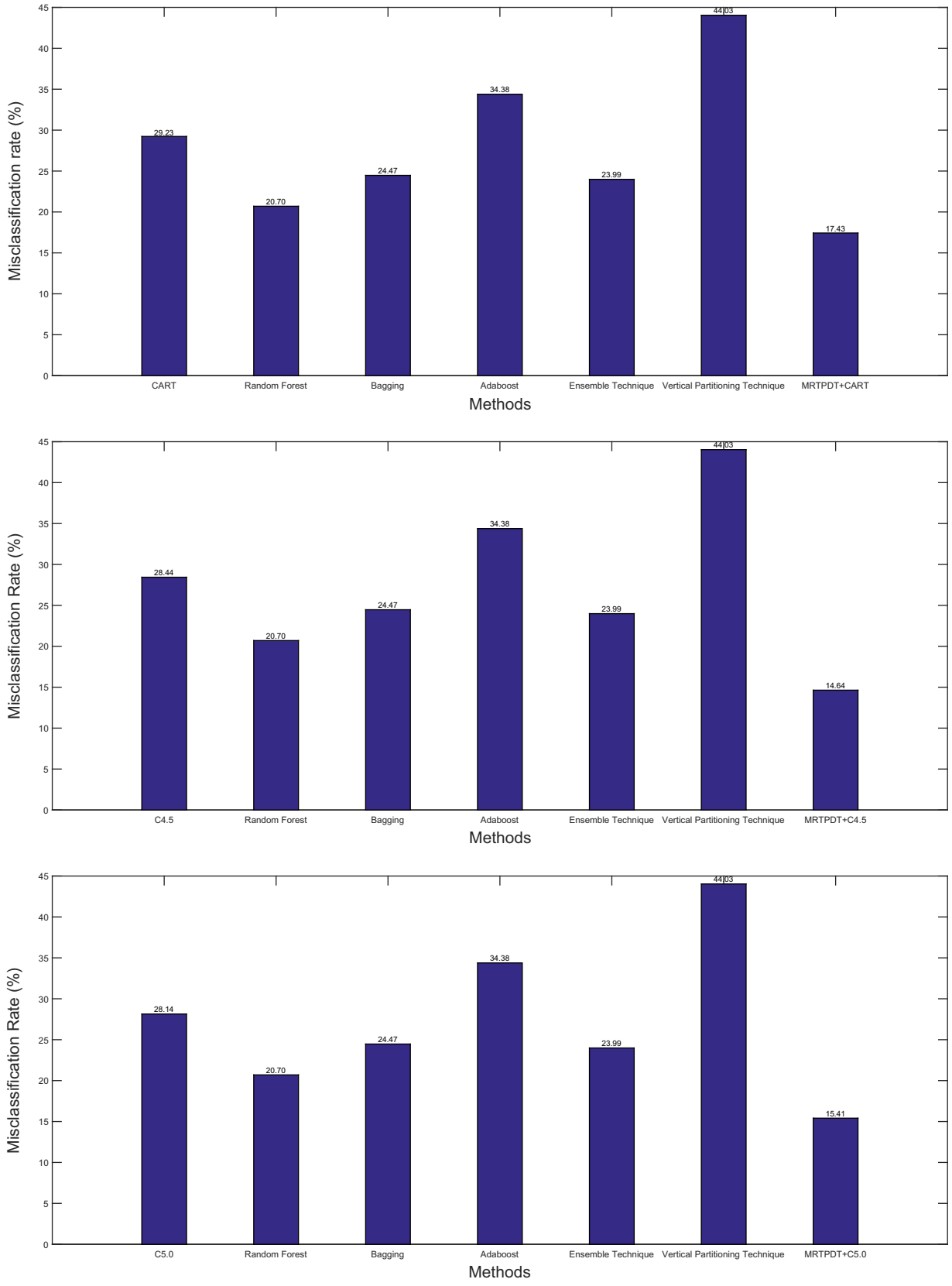
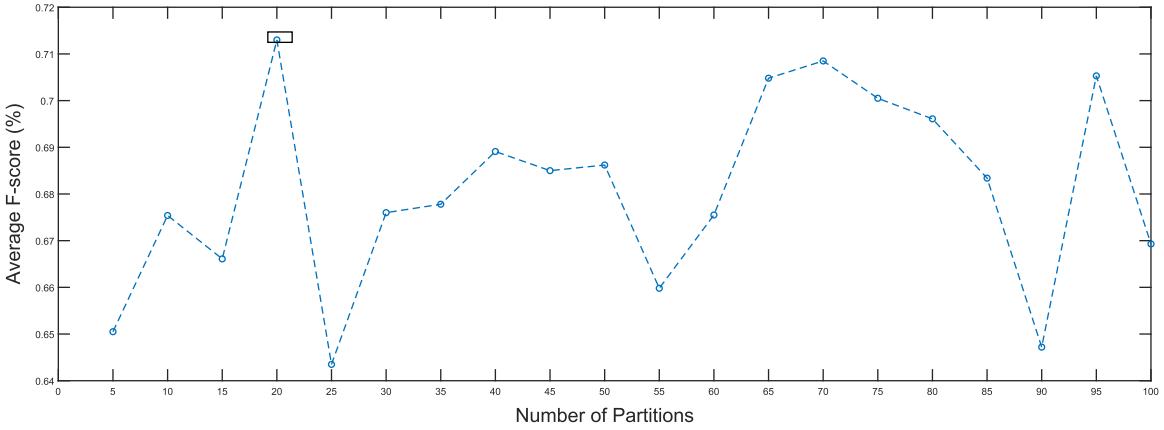


Fig. 7. Average misclassification rate for MRTPDT and other methods.

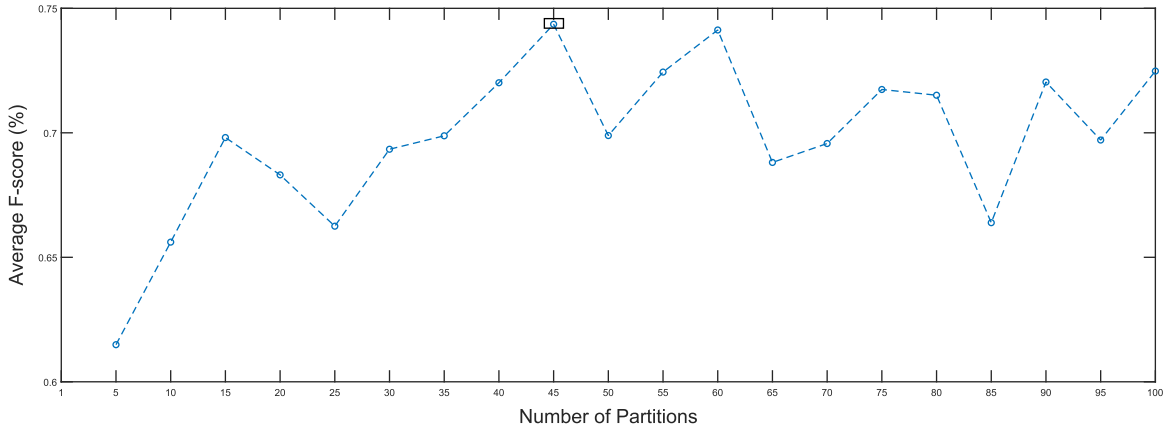
Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

Table 11
Characteristics of Class Imbalance Datasets.

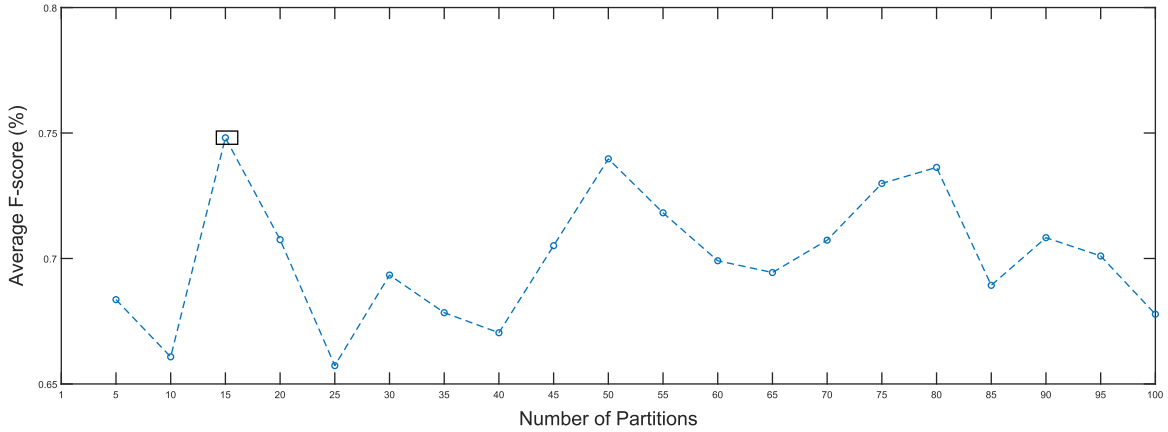
Sl. No.	Data set	Imbalance Ratio	Class distribution Ratio
1	Colon Tumor (Andres Cano and Andres, 2005)	0.55	0.645;0.354
2	NIC60 (Zhu et al., 2010)	0.66	0.11;0.13;0.098;0.11;0.14;0.065;0.098;0.098;0.13
3	Central Nervous System (Andres Cano and Andres, 2005)	0.54	0.65;0.35



(a) Average F-score for MRTPDT+SMOTE+CART



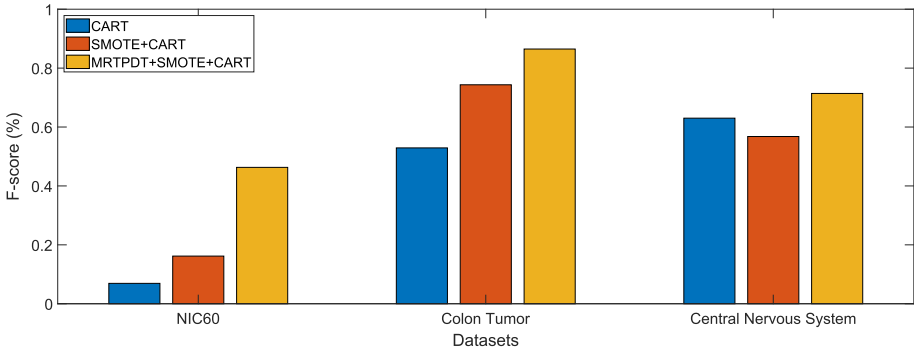
(b) Average F-score for MRTPDT+SMOTE+C4.5



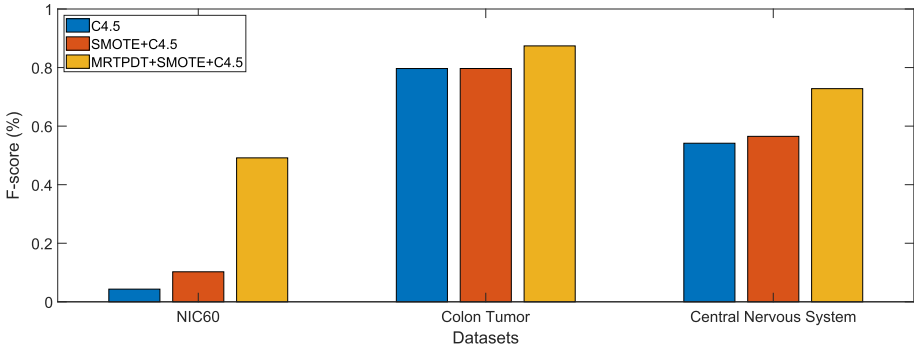
(c) Average F-score for MRTPDT+SMOTE+C5.0

Fig. 8. Average F-score for MRTPDT method across the partitions.

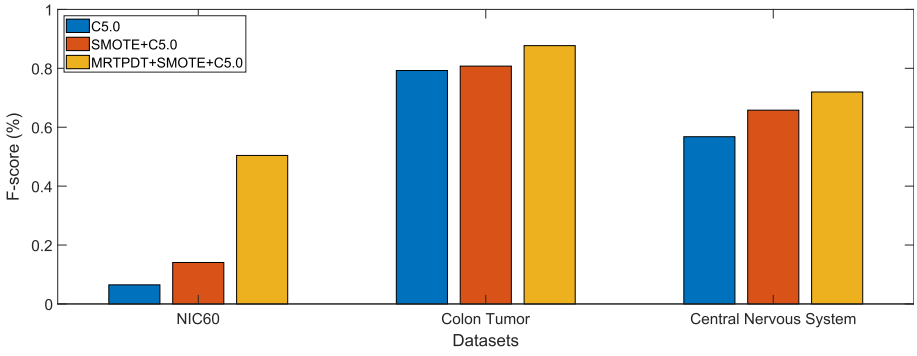
Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015



(a) Average F-score values for MRTPDT+SMOTE+CART



(b) Average F-score values for MRTPDT+SMOTE+C4.5



(c) Average F-score values for MRTPDT+SMOTE+C5.0

Fig. 9. Average F-score values for datasets using MRTPDT and other methods.

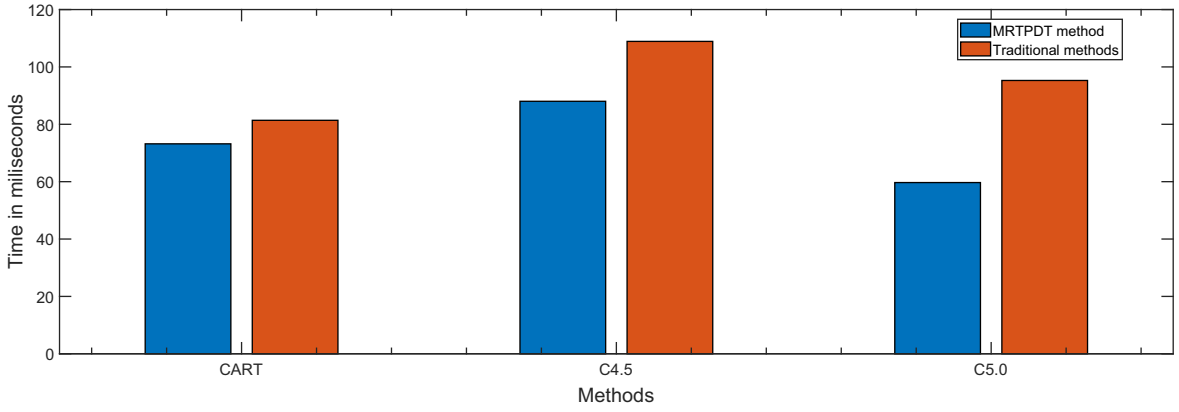


Fig. 10. Average time complexity.

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

8. Conclusion and future work

In the present study, we explored a novel vertical partitioning approach by using the ideas from music rhythm tree. The MRTPDT method is compared with Random Forest, Bagging, AdaBoost, an ensemble technique and a vertical partitioning method with respect to classification accuracy, standard deviation, misclassification rate and F-score. We used high dimensional datasets to evaluate superiority of the proposed method and it is proved that our technique is unique and scalable for such datasets. Our study also indicates a direction to choose right number of partitions. The method outperforms other established ensemble techniques and classical decision trees in terms of classification accuracy, structural stability and handling class imbalanced datasets. In future, we would like to explore evolutionary techniques to do further improvement in the performance of MRTPDT method.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to Technical Education Quality Improvement Programme (TEQIP-III) of B.M.S. College of Engineering, Bengaluru for providing sufficient facilities to publish this work.

References

Aha, D., Murphy, P., 1994. UCI Repository of machine learning databases. http://www.ics.uci.edu/mlearn/MLRepository.html. Irvine, CA: University of California, Department of Information and Computer Science.

Alexander, S.C., Garner, D.K., Somoroff, M., Gramling, D.J., Norton, S.A., Gramling, R., 2015. Using music[al] knowledge to represent expressions of emotions. Patient Educ. Couns. 98 (11), 1339–1345.

Cano, Andres, Masegosa, Andres, Moral, Serafin, 2005. ELVIRA Biomedical Data Set Repository. http://leo.ugr.es/elvira/DBCRepository/index.html.

Balasubramanian, G., Kanagasabai, A., Mohan, J., Seshadri, N.G., 2018. Music induced emotion using wavelet packet decompositionan eeg study. Biomed. Signal Process. Control 42, 115–128.

Baranauskas, J.A., 2015. The number of classes as a source for instability of decision tree algorithms in high dimensional datasets. Artif. Intell. Rev., 301–311

Blum, A., Mitchell, T., 1998. Combining labeled and unlabeled data with cotraining. In: COLT: Proceedings of the Workshop on Computational Learning Theory. Morgan Kaufmann Publishers.

Boenn, G., 2018. The Farey Sequence as a Model for Musical Rhythm and Meter. Springer, Springer International PublishingSpringer International Publishing, pp. 83–112.

C5.0, 1993. See5: An informal tutorial. http://www.rulequest/see5-win.html..

Catal, C., Tufekci, S., Pirmit, E., Kocabag, G., 2015. On the use of ensemble of classifiers for accelerometer-based activity recognition. Appl. Soft Comput. 37, 1018–1022.

Cervantes, B., Monroy, R., Medina-Prez, M.A., Gonzalez-Mendoza, M., Ramirez-Marquez, J., 2018. Some features speak loud, but together they all speak louder: a study on the correlation between classification error and feature usage in decision-tree classification ensembles. Eng. Appl. Artif. Intell. 67, 270–282.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. Synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357.

Chen, Z., Lin, T., Xia, X., Xu, H., Ding, S., 2018. A synthetic neighborhood generation based ensemble learning for the imbalanced data classification. Appl. Intell. 48 (8), 2441–2457.

Chih-Chung, C., Chih-Jen, L., 2011. Libsvm: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2 (3), 1–27.

Das, R., Walia, E., 2019. Partition selection with sparse autoencoders for content based image classification. Neural Comput. Appl. 31 (3), 675–690.

Dostál, M., 2013. Evolutionary Music Composition. Springer, Berlin Heidelberg, pp. 935–964.

Erdem Günay, M., Türker, L., Alper Tapan, N., 2018. Decision tree analysis for efficient CO_2 utilization in electrochemical systems. J. CO_2 Util. 28, 83–95. https://doi.org/10.1016/j.jcou.2018.09.011.

Guggari, S., Kadappa, V., Umadevi, V., 2018. Non-sequential partitioning approaches to decision tree classifier. Future Comput. Inf. J., 1–11

Guyon, Isabelle, 2003. NIPS 2003 workshop on feature extraction. http://clopinet.com/isabelle/Projects/NIPS2003/..

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The weka data mining software: an update. ACM SIGKDD Explor. Newslett. 11 (1), 10–18.

Jacquemard, F., Donat-Bouillud, P., Bresson, J., 2015. In: Collins, T., Meredith, D., Volk, A. (Eds.), Mathematics and Computation in Music, pp. 3–15.

Kadappa, Vijayakumar, Guggari, Shankru, Negi, Atul, 2015. Decision tree classifier using theme based partitioning. In: IEEE International Conference on computing and network Communications (CoCoNet'15), pp. 546–552.

Kotz, S., Ravignani, A., Fitch, W., 2018. The evolution of rhythm processing. Trends Cogn. Sci. 22 (10), 896–910. special Issue: Time in the Brain.

Kumar, Vipin, Minz, Sonajharia, 2015. Multi-view ensemble learning: a supervised feature set partitioning for high dimensional data classification. In: Proceedings of the Third International Symposium on Women in Computing and Informatics, pp. 31–37.

Kumar, Vipin, Minz, Sonajharia, 2015. Multi-view ensemble learning using optimal feature set partitioning: an extended experiments and analysis in low dimensional scenario. In: Second International Symposium on Computer Vision and the Internet (VisionNet15), vol. 58. pp. 499–506.

Kumar, S., Minz, Vipinand, 2017. An optimal multi-view ensemble learning for high dimensional data classification using constrained particle swarm optimization. Inf. Commun. Comput. Technol., 363–378

Kusiak, A., Larson, N., 1995. Decomposition and representation methods in mechanical design. Trans. ASME 117.

Last, M., Maimon, O., Minkov, E., 2002. Improving stability of decision trees. Int. J. Pattern Recogn. Artif. Intell. 16, 145–159.

Leo, B., 1996. Heuristics of instability and stabilization in model selection. Ann. Statist. 24, 2350–2383.

Leo, B., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Liu, S., Wang, Y., Chen, C., Xiang, Y., 2016. An ensemble learning approach for addressing the class imbalance problem in twitter spam detection. In: Liu, J.K., Steinfeld, R. (Eds.), Information Security and Privacy. Springer International Publishing, pp. 215–228.

Liu, Chang, Zhenhua, Hu., Li, Yan, Liu, Shaojun, 2017. Forecasting copper prices by decision tree learning. Resour. Policy 52, 427–434.

Mirzamomen, Z., Kangavari, M.R., 2016. A framework to induce more stable decision trees for pattern classification. Pattern Anal. Appl., 1–14

Mirzamomen, Z., Fekri, M.N., Kangavari, M., 2015. Cross split decision trees for pattern classification. In: International Conference on Computer and Knowledge Engineering (ICCKE), pp. 240–245.

Mohammed, A.M., Onieva, E., Woźniak, M., 2020. Vertical and horizontal data partitioning for classifier ensemble learning. Prog. Comput. Recogn. Syst., 86–97

Natek, Srecko, Zwilling, Moti, 2014. Student data mining solution knowledge management system related to higher education institutions. Comput. Ind. 41, 6400–6407.

Oliveira, Werbeston D., Vieira, Joo P.A., Bezerra, Ubiratan H., Martins, Daniel A., Benedito das, G. Rodrigues, 2017. Power system security assessment for multiple contingencies using multiway decision tree. Electr. Power Syst. Res. 148, 264–272.

Sebö, A., Waksman, Z., 1999. Optimal binary trees with order constraints. Discrete Appl. Math. 91 (1), 305–311. https://doi.org/10.1016/S0166-218X(98)00138-3.

Paul, J., Verleysen, M., Dupont, P., 2012. The stability of feature selection and class prediction from ensemble tree classifiers.

Rizo, D., Quereda, J.M.I., 2002. Tree-structured representation of melodies for comparison and retrieval. In: PRIS.

Rizo, D., Lemström, K., Iñesta, J.M., 2009. Tree representation in combined polyphonic music comparison. In: Computer Music Modeling and Retrieval. Genesis of Meaning in Sound and Music. pp. 177–195.

Rizo, D., Iesta, J.M., Moreno-seco, F., 2003. Tree-structured representation of musical information. In: First Iberian Conference Pattern Recognition and Image Analysis, Palma de Mallorca (Spain), Volume 2652 of LNCS. pp. 838–846.

Rokach, L., 2001. Decision tree. In: Book Chapter, vol. 24.

Rokach, Lior, 2006. Decomposition methodology for classification tasks – a meta decomposer framework. Pattern Anal. Appl. 9, 257–271.

Rokach, Lior, 2008. Genetic algorithm-based feature set partitioning for classification problems. J. Pattern Recogn. Soc. 41, 1676–1700.

Rokach, L., 2010. Ensemble-based classifiers. Artif. Intell. Rev. 33 (1), 1–39.

Rokach, Lior, Maimon, Oded, 2006. Data mining for improving the quality of manufacturing: a feature set decomposition approach. J. Intell. Manuf. 17, 285–299.

Seetha, Hari, Narasimha Murty, M., 2016. Classification by majority voting in feature partitions. Int. J. Inf. Decis. Sci. 8, 109–124.

Tao, X., Li, Q., Ren, C., Guo, W., Li, C., He, Q., Liu, R., Zou, J., 2019. Real-value negative selection over-sampling for imbalanced data set learning. Expert Syst. Appl. 129, 118–134.

R, 1993. The R project for statistical computing. http://www.r-project.org/..

Vipin, K., Sonajharia, M., 2016. Multi-view ensemble learning: an optimal feature set partitioning for high-dimensional data classification. Knowl. Inf. Syst. 49 (1), 1–59.

Wang, L., Li, Q., Yu, Y., Liu, J., 2018. Region compatibility based stability assessment for decision trees. Expert Syst. Appl. 105, 112–128.

Wang, Z., Chen, Z., Zhu, Y., Zhang, J., Du, W., Li, D., 2019. Multi-matrices entropy discriminant ensemble learning for imbalanced problem. Neural Comput. Appl., 1–20

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015

Xindong, Wu., Vipin Kumar, J., Quinlan, Ross, et al., 2008. Top 10 algorithms in data mining. Knowl. Inf. Syst. 14, 1–3.
Zhang, H., Liu, W., Shan, J., Liu, Q., 2018. Online active learning paired ensemble for concept drift and class imbalance. IEEE Access 6, 73815–73828.
Zhang, C., Bi, J., Xu, S., Ramentol, E., Fan, G., Qiao, B., Fujita, H., 2019. Multi-imbalance: an open-source software for multi-class imbalance learning. Knowl.-Based Syst. 174, 137–143.
Zhu, Z., Ong, Y., Zurada, J.M., 2010. Identification of full and partial class relevant genes. IEEE/ACM Trans. Comput. Biol. Bioinf. 7 (2), 263–277.
Zimmermann, A., 2008. Ensemble-trees: leveraging ensemble power inside decision trees. In: Discovery Science, pp. 76–87.

Please cite this article as: S. Guggari, V. Kadappa, V. Umadevi et al., Music rhythm tree based partitioning approach to decision tree classifier, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2020.03.015