International Journal of Engineering & Technology, 7 (2) (2018) 948-955

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i2.10714
Research paper

Performance analysis of students debugging skills with trait emotional intelligence using decision tree based algorithms

V. Maria Antoniate Martinr [1] *, Dr. K. David [2]

[1] *Research Scholar, Department of Computer Science, Research and Development Centre, Bharathiar University, Coimbatore, Tamil Nadu, India*
[2] *Dr. K. David , Assistant Professor, Department of Computer Science, H.H. The Rajah's College, Pudukkottai, Tamil Nadu, India*
**Corresponding author E-mail: maria.anto82@gmail.com*

Abstract

Data mining is the extraction of knowledge, meaningful patterns, trends and relationships from huge amount of data stored in repositories. This research focuses on assessing the decision tree techniques like FT tree, J48 graft pruned tree, Random, NB and LAD. A study was conducted on students from post graduation. The questionnaire was designed to test the students debugging skills and Trait Emotional Intelligence. The TEI skills are broadly classified as wellbeing, self-control, emotionality, sociability and global trait EI. The decision trees techniques were applied to the factors like locality, gender, academic performance and students debugging skills. The performances of all the decision tree techniques are compared based on the error measures to find the best suited technique.

Keywords: *Data Mining; Trait Emotional Intelligence; FT Tree; J48Graftt, LAD; NB and Random.*

1. Introduction

Data mining discover interesting patterns from huge amount of data [8]. Data mining or KDD is the automatic extraction of interesting patterns from large data collections. It can be used to learn the model for the learning process or student modeling also evaluate the e-learning system [29]. The concept of TEI was popularized in 1995 by Daniel Goleman's book, Emotional Intelligence. EI can be "as powerful and at times more powerful, than IQ" in predicting how successful one is in life [16]. This paper focuses on finding out the students debugging skills with Trait Emotional Intelligence (TEI) using decision tree based algorithms. The test was conducted for the students. The questionnaire consisted of student demographic details which included gender, locality, academic details and debugging skills. The TEI questionnaire contained 30 questions. The TEI is divided into five factors; they are well being, self control, emotionality, sociability and global TEI. This data set was applied into data mining techniques. The decision tree techniques are FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. The based on the outcome and error rates the best technique is found. The performance in debugging is influenced by their TEI skills, students have varied TEI skills. Finally, it is observed that those who have low debugging marks and suggestion is given to the students to improve their low TEI factors.

2. Related works

Data mining is used to identify hidden patterns in a large data set. E. Chandra and K. Nandhini presents data mining in education environment that identifies students failure patterns [3]. The Trait Emotional Intelligence refers to a constellation of emotional self perceptions located at the lower levels of personality hierarchies [1]. The work done by C.Deepa, et al. presents that tree based algorithms outperform the traditional data driven techniques like multilayer perceptron and linear regression in predicting the strength of concrete mix [4]. Jai Ruby, et al. has tested the various classification algorithms like ID3, REPTree, Simple cart, J48, NB Tree, BF Tree, Decision Table, MLP and Bayesnet [7].
V. Maria Antoniate Martin et al. has found out the student performance using data mining techniques, association between trait emotional intelligence and personality trait, and to analyze the students debugging skills with TEI [12-15]. Olaiya Folorunsho has proved that Decision Tree algorithms like REP tree and LAD Tree outperform artificial neural network techniques like Multilayer perceptron and RBF in predicting the patient's health condition [17]. Sharon Dominick et al. have presented a work on the use of clustering techniques in predicting the rise of sea levels which is an effect of global warming [31]. Sherine Dominick et al. have observed the effects of the various classifiers in analyzing the factors affecting the academic performance of the students [32]. Payal Dhakate et al. have developed an ensemble combining classifiers like bagging, Adaboost, c4.5, J48 and Ad tree and Preprocessing and classification techniques were applied them on a medical dataset [19, 20]. Purva Sewaiwar et al., has done the work by comparative study of various decision tree classification algorithm like J48, LMT, Random forest and Random tree[24]. Trilok Chand Sharma et al., has done the work by comparative study of classification algorithm like J48, LAD tree, AD tree and FT tree [36]. Jafar Tanha et al. presented semi-supervised learning task from both labeled and unlabeled instances decision tree classifiers [6]. Koushal kumar has presented knowledge extracted from trained neural networks. Neural networks combined with decision trees to fetch the knowledge from training process [9]. Rashedul

M. Rahman have proved comparison various classification techniques using different data mining tools using medical data set [26]. K. Wisaeng has presented comparison of different classification techniques which consists of decision tree method and machine learning for bank data set [37]. Ashokkumar Vijaysinh Solanki, has showed the use of J48 and Random Tree for the classification of Sickle Cell disease cells in which Random tree produced the depth precision [2]. Dorina kabakchieva presents predicting student performance using classification [5]. V. Ramesh focus on student performance analysis predicted for placement using naïve bayes simple, multi layer perception, SMO, J48 and REP tree [25].

Rithika et al. has research focused on data mining classification like decision tree, neural networks, genetic programming and statistical algorithms [27]. She has predicted the blood donors population using classification technique [28]. Shomona Gracia Jacob found the accuracy of eight clustering algorithms in detecting clusters of patient records and sixteen classification algorithms on the lymphography data set [33]. Xindong wu et al. surveyed top 10 algorithms in data mining like c4.5, k-means, svm, apriori, em, page rank, adaboost, kNN, navie bayes and CART [38]. Lakshmi Devasena C has compared the LAD tree and REP tree classifiers for credit risk forecast [10]. Payal P. Dhakate has proved preprocessing and classification using different classifiers like LAD, J48, random tree, REP tree and simple cart [18]. Manuel J. Fonseca et al. proposed an efficient indexing structure for mapping multidimensional data points to a one dimensional line by computing the Euclidean Norm and storing the data points in a B+tree[11]. Pooja Sharma et al. has work done by the student analysis performance using decision tree algorithm like NB tree [23]. Yogendra kumar et al. studied four machine learning algorithms like J48, BayesNet, OneR and NB for the detecting intrusions and compares their relative performances [40].

Natalie L. Shipley has examined the academic performance and relationship between the emotional intelligence measured by the Trait Emotional Intelligence Questionnaire Short Form [16]. K.V. Petrides et al. have focused the role of trait emotional intelligence in academic performance. He also examined Trait Emotional Intelligence in children's peer relations at school [21,22]. M. J. Sanchez ruiz et al. investigated the relationship between trait emotional intelligence and two creativity indicators like divergent thinking and creative personality [30]. Stella Mavroveli et. al. investigated the relationship between trait EI and four distinct socio emotional criteria on adolescents. They also investigated the construct of trait emotional intelligence in children [34, 35].

3. Methodology

The methodology diagram for this research is shown in figure 1. The demographic details collected from students. The debugging test and Trait Emotional Intelligence test were conducted to the students. The demographic details include name, gender, locality and academic details. The results from debugging questionnaires in object oriented programming languages are also included in the dataset.

Debugging performance marks is graded, as follows, for 0 -39 marks 'F Grade', for 40 -49 marks 'C Grade', for 50 -59 marks 'B Grade', for 60 -74 marks 'A Grade', for 75 -79 marks 'D Grade', for 80 -100 marks 'O Grade'.

3.1. Trait emotional intelligence (TEI)

K.V. Petrides's questionnaire which is based on factors the TEI factors [16] like well-being, self-control, emotionality, sociability and global trait EI is used for the analysis with reference to 7 point Likert scale. Each factor has values ranging from completely disagree to completely agree.

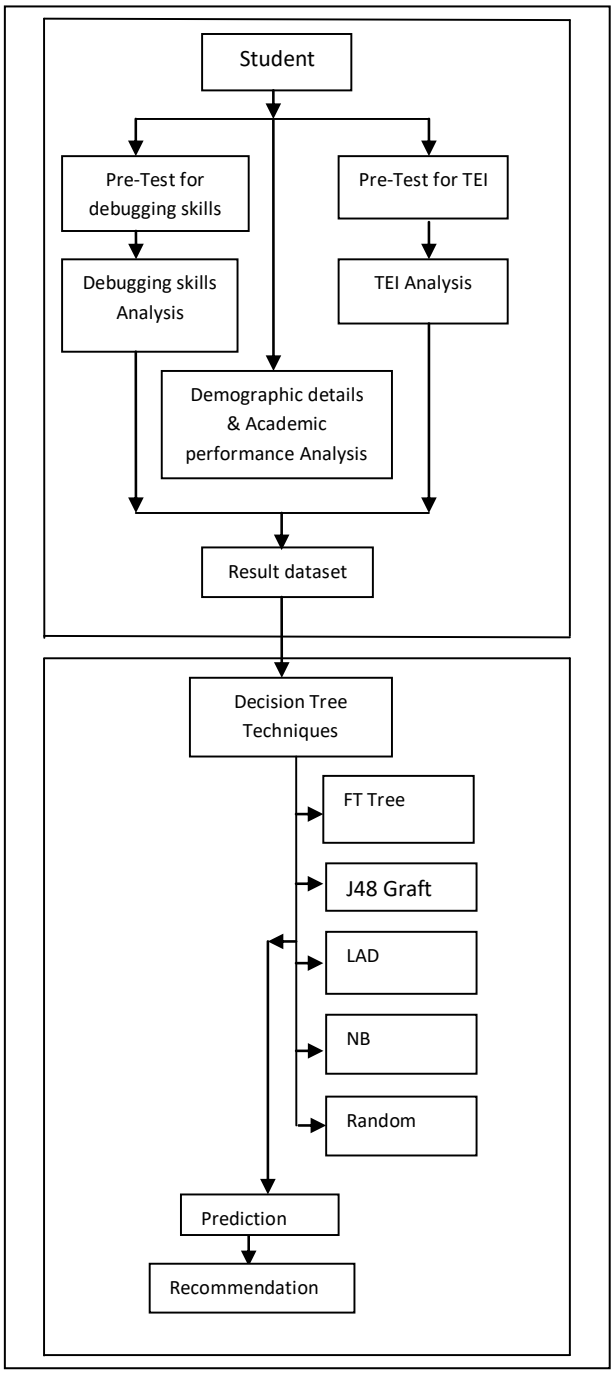

Fig. 1: Methodology Diagram.

Well-being attribute highlights the abilities like self confidence, cheerfulness, satisfaction and positive attitude. Self Control reflects on the ability to control emotions, withstand pressure, regulate stress, being less reflective and control urges. Emotionality refers to the constraint relating to ones feeling, ability to communicate the feeling, fulfilling relationships and empathy. Sociability influences good social skills, which creates in others feelings, frank and ready to fight for rights. Global Trait EI is about being flexible, adapting new environments, and ready to face and withstand adversity.

3.2. Data mining techniques

The five data mining decision tree techniques were applied to the dataset. They are following, (i) FT Tree (ii) J48 Graft (iii) LAD Tree (iv) NB Tree (v) Random Tree.

3.2.1. Confusion matrix

Confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known.

- Precision is defined as the number of true positive from the positively assigned document which is illustrated,

Precision = tp/ (tp+fp)

- Recall is denoted as the number of true positive out of the actual positive document which is illustrated,

Recall =tp/ (tp+fn)

- F-measure t is defined as a weighted method of precision and recall, and it which is computed as,

F - Measure = 2*precision*recall / precision + recall

- ROC Area means Receiver Operating Characteristic curve which is the true positive rate (Sensitivity) and is plotted in function of the false positive rate (100- Specificity) for different cut-off points. Each point on The ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold.

3.3. Error measures in classification

Various decision tree techniques results are tested using measures like kappa statistics, Mean Absolute Error, Root Mean Squared Error and Relative Absolute Error. These error measures denote the presence of outliers, values that possibly may not belong to the classification model, which indeed disturbs the entire process. Loss functions measure the error between yi and the predicted value, yi. The most common loss functions are

- Kappa Statistics is used to measure the agreement between the classification items and the categories. The kappa statistic value is denoted by k.

$$K=\frac{\Pr(a)-\Pr(e)}{1-\Pr(e)}$$

Pr (a) denotes the agreement among the raters and Pr (e) denotes the probability of chance of agreement. In case of complete agreement k=1, else k=0 for no agreement.

- Mean Absolute Error, the generalization error which is the average loss over the test is denoted by,

$$\text{Mean Absolute Error} = \frac{\sum_{i=1}^{d} | y_i - y_i ' |}{d}$$

- Root Mean Squared Error It is used in data mining as it combines concepts from,

$$\text{Mean Squared Error} = \frac{\sum_{i=1}^{d} (y_i - y_i ')^2}{d}$$

$$\frac{\sum_{i=1}^{d} | y_i - y_i ' |}{\sum_{i=1}^{d} | y_i - \bar{y} |}$$

- Relative Absolute Error =
- Root Relative Squared Error

This relative squared error takes the total squared error and normalizes it by dividing by the total squared error of the simple predictor. By taking the square root of the relative squared error it reduces the error to the same dimensions as the quantity being predicted. It is the square root of the Relative Squared Error.

$$\text{Relative Squared Error} = \frac{\sum_{i=1}^{d} (y_i - y_i ')^2}{\sum_{i=1}^{d} (y_i - \bar{y})^2}$$

4. Result and discussion

A comparison is made between different Decision Tree Techniques of FT Tree, J48 Graft, LAD, NB and Random using weka tool. The decision tree techniques FT Tree, J48 Graft, LAD, NB and Random are applied to the data set and accuracy is determined using factors like TP Rate, FP Rate, precision, Recall, F-Measure and ROC area.

The performance of the classification function is noted with respect to some error measures like kappa statistic, mean absolute error, Root mean squared error, Relative absolute error and Root relative squared error.

Table 1 shows the accuracy measures for gender where ‘A and B’ denotes the Male and Female attribute of the Confusion Matrix respectively. The decision tree classification results with respect to Table 1 are (77.2277, 81.1881, 77.2277, 73.2673, 59.4059) % of correctly classified instances and (22.7723, 18.8119, 22.7723, 26.7327, 40.5941)% of incorrectly classified instances of the gender attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the J48 Graft tree provides better results with high weighted average of TP Rate as 0.812, FP Rate as 0.194, Precision as 0.814, recall as 0.812, F-Measure 0.811 and ROC area as 0.79.

Table 2 shows the accuracy measures for locality where ‘A and B’ denotes the rural and urban attributes of the Confusion Matrix respectively. The decision tree classification results with respect to Table 2 are (52.4752, 59.4059, 51.4851, 49.505, 57.4257) % of correctly classified instances and (47.5248, 40.5941, 48.5149, 50.495, 42.5743)% of incorrectly classified instances of the locality attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the J48 Graft tree provides better results with high weighted average of TP Rate as 0.594 FP Rate as 0.433, Precision as 0.593, recall as 0.594, F-Measure 0.593 and ROC area as 0.597.

Table 3 shows the accuracy measures for academic performance where ‘A, B, C, and D’ denotes the ‘A, O, D, and B’ GRADE of the Confusion Matrix. The decision tree classification results with respect to Table 3 are (52.4752, 59.4059, 51.4851, 49.505, 57.4257) % of correctly classified instances and (47.5248, 40.5941, 48.5149, 50.495, 42.5743)% of incorrectly classified instances of the locality attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the J48 Graft tree provides better results with high weighted average of TP Rate as 0.594, FP Rate as 0.433, Precision as 0.593, recall as 0.594, F-Measure 0.593 and ROC area as 0.597.

Table 1: Accuracy Measures for Gender

Classifier	FT Tree		J48 Graft		LAD		NB		RANDOM	
	A	B	A	B	A	B	A	B	A	B
TP Rate	0.83	0.708	0.868	0.75	0.792	0.75	0.736	0.729	0.755	0.417
FP Rate	0.292	0.17	0.25	0.132	0.25	0.208	0.271	0.264	0.583	0.245
Precision	0.759	0.791	0.793	0.837	0.778	0.766	0.75	0.714	0.588	0.606
Recall	0.83	0.708	0.868	0.75	0.792	0.75	0.736	0.729	0.755	0.417
F-Measure	0.793	0.747	0.829	0.791	0.785	0.758	0.743	0.722	0.661	0.494
ROC Area	0.795	0.795	0.79	0.79	0.79	0.79	0.82	0.82	0.656	0.656

Table 2: Accuracy Measures for Locality

Classifier	FT Tree		J48 Graft		LAD		NB		RANDOM	
	A	B	A	B	A	B	A	B	A	B
TP Rate	0.475	0.595	0.661	0.5	0.542	0.476	0.559	0.405	0.763	0.31
FP Rate	0.405	0.525	0.5	0.339	0.524	0.458	0.595	0.441	0.69	0.237
Precision	0.622	0.446	0.65	0.512	0.593	0.426	0.569	0.395	0.608	0.481
Recall	0.475	0.595	0.661	0.5	0.542	0.476	0.559	0.405	0.763	0.31
F-Measure	0.538	0.51	0.655	0.506	0.566	0.449	0.564	0.4	0.677	0.377
ROC Area	0.524	0.524	0.597	0.597	0.538	0.538	0.454	0.454	0.54	0.54

Table 3: Accuracy Measures for Academic Performance

Classifier	FT Tree				J48 Graft				LAD				NB				RANDOM			
	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
TP Rate	0.771	0.067	0.429	0	0.771	0.067	0.393	0	0.75	0.133	0.25	0.1	0.646	0.2	0.357	0	0.792	0.067	0.179	0
FP Rate	0.453	0.058	0.274	0.022	0.528	0.058	0.26	0	0.415	0.105	0.288	0.033	0.377	0.105	0.274	0.088	0.642	0.093	0.192	0.011
Precision	0.607	0.167	0.375	0	0.569	0.167	0.367	0	0.621	0.182	0.25	0.25	0.608	0.25	0.333	0	0.528	0.111	0.263	0
Recall	0.771	0.067	0.429	0	0.771	0.067	0.393	0	0.75	0.133	0.25	0.1	0.646	0.2	0.357	0	0.792	0.067	0.179	0
F-Measure	0.679	0.095	0.4	0	0.655	0.095	0.379	0	0.679	0.154	0.25	0.143	0.626	0.222	0.345	0	0.633	0.083	0.213	0
ROC Area	0.709	0.667	0.678	0.637	0.591	0.527	0.447	0.73	0.648	0.664	0.559	0.619	0.61	0.683	0.575	0.697	0.585	0.5	0.466	0.601

Table 4: Accuracy Measures for Debugging Skills

Classifier	FT Tree						J48 Graft						LAD						NB						RANDOM					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
TP Rate	0.167	0.588	0.516	0	0	0	0.278	0.529	0.484	0	0	0	0.278	0.412	0.484	0	0	0	0.333	0.441	0.387	0.167	0	0	0.056	0.882	0.129	0	0	0
FP Rate	0.096	0.373	0.414	0	0	0	0.12	0.373	0.314	0.042	0.011	0.01	0.096	0.388	0.357	0.032	0.022	0.031	0.06	0.343	0.4	0.084	0.011	0.021	0.012	0.881	0.086	0	0	0
Precision	0.273	0.440	0.356	0	0	0	0.333	0.419	0.405	0	0	0	0.385	0.35	0.375	0	0	0	0.545	0.395	0.3	0.110	0	0	0.5	0.337	0.4	0	0	0
Recall	0.167	0.588	0.516	0	0	0	0.278	0.529	0.484	0	0	0	0.278	0.412	0.484	0	0	0	0.333	0.441	0.387	0.167	0	0	0.056	0.882	0.129	0	0	0
F-Measure	0.207	0.506	0.421	0	0	0	0.303	0.468	0.441	0	0	0	0.323	0.378	0.423	0	0	0	0.414	0.417	0.338	0.133	0	0	0.1	0.488	0.195	0	0	0
ROC Area	0.592	0.619	0.56	0.605	0.61	0.348	0.673	0.503	0.583	0.735	0.554	0.371	0.687	0.505	0.51	0.831	0.618	0.398	0.821	0.567	0.56	0.868	0.491	0.559	0.481	0.505	0.549	0.422	0.591	0.254

Table 4 shows the accuracy measures for debugging skills performance where 'A, B, C, D, E and F' denotes the 'F, A, B, D, O and C' GRADE attribute respectively of the Confusion Matrix. The decision tree classification results with respect to Table 4 are (38.6139, 37.6238, 33.6634, 33.6634, 34.6535) % of correctly classified instances and (61.3861, 62.3762, 66.3366, 66.3366, 65.3465)% of incorrectly classified instances of the locality attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB

tree and Random tree. When comparing all the results, the FT tree provides better results with high weighted average of TP Rate as 0.386, FP Rate as 0.27, Precision as 0.307, recall as 0.386, F-Measure 0.337 and ROC area as 0.584.

Table 5 shows the accuracy measures for well being where A, B, and C' denotes the 'WBN, WBA, and WBD' attributes respectively, which means well being neutral, agree and disagree of the Confusion Matrix. The decision tree classification results with respect to Table 5 are (81.1881, 81.1881, 72.2772, 76.2376, 80.198) % of correctly classified instances and (18.8119, 18.8119, 27.7228, 23.7624, 19.802)% of incorrectly classified instances of the well being attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the Random tree provides better results with high weighted average of TP Rate as 0.802, FP Rate as 0.772, Precision as 0.709, recall as 0.802, F-Measure 0.738 and ROC area as 0.535.

Table 5: Accuracy Measures for Well Being

Classifier	FT Tree			J48 Graft			LAD			NB			RANDOM		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
TP Rate	0.217	0.87	0	0	0.971	0	0.304	0.841	0	0.391	0.812	0	0.087	0.899	0.111
FP Rate	0.103	0.844	0.011	0.026	1	0	0.115	0.75	0.033	0.141	0.656	0.043	0.09	0.875	0.011
Precision	0.385	0.69	0	0	0.677	0	0.438	0.707	0	0.45	0.727	0	0.222	0.689	0.5
Recall	0.217	0.87	0	0	0.971	0	0.304	0.841	0	0.391	0.812	0	0.087	0.899	0.111
F-Measure	0.278	0.769	0	0	0.798	0	0.359	0.768	0	0.419	0.767	0	0.125	0.78	0.182
ROC Area	0.604	0.55	0.493	0.435	0.448	0.42	0.534	0.49	0.531	0.61	0.577	0.607	0.495	0.48	0.583

Table 6: Accuracy Measures for Self Control

Classifier	FT Tree			J48 Graft			LAD			NB			RANDOM		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
TP Rate	0	1	0	0	1	0	0	0.89	0	0.143	0.915	0	0.071	0.976	0
FP Rate	0	1	0	0	1	0	0.08	1	0.021	0.057	0.789	0.042	0.023	0.947	0
Precision	0	0.812	0	0	0.812	0	0	0.793	0	0.286	0.833	0	0.333	0.816	0
Recall	0	1	0	0	1	0	0	0.89	0	0.143	0.915	0	0.071	0.976	0
F-Measure	0	0.896	0	0	0.896	0	0	0.839	0	0.19	0.872	0	0.118	0.889	0
ROC Area	0.487	0.535	0.451	0.396	0.466	0.235	0.501	0.515	0.265	0.617	0.626	0.627	0.459	0.545	0.581

Table 7: Accuracy Measures for Emotionality

Classifier	FT Tree			J48 Graft			LAD			NB			RANDOM		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
TP Rate	0	0.919	0	0.059	0.968	0	0.059	0.806	0.045	0.118	0.694	0.136	0	0.968	0.045
FP Rate	0.012	0.974	0.063	0.012	0.949	0.025	0.036	0.923	0.127	0.107	0.718	0.203	0	0.974	0.025
Precision	0	0.6	0	0.5	0.619	0	0.25	0.581	0.091	0.182	0.606	0.158	0	0.612	0.333
Recall	0	0.919	0	0.059	0.968	0	0.059	0.806	0.045	0.118	0.694	0.136	0	0.968	0.045
F-Measure	0	0.726	0	0.105	0.755	0	0.095	0.676	0.061	0.143	0.647	0.146	0	0.75	0.08
ROC Area	0.492	0.461	0.47	0.466	0.448	0.393	0.571	0.416	0.431	0.535	0.453	0.413	0.563	0.507	0.541

Table 8: Accuracy Measures for Sociability

Classifier	FT Tree			J48 Graft			LAD			NB			RANDOM		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
TP Rate	0.625	0.182	0.129	0.729	0.364	0.355	0.542	0.182	0.323	0.542	0.136	0.484	0.875	0.136	0.065
FP Rate	0.66	0.215	0.157	0.566	0.063	0.171	0.472	0.127	0.371	0.547	0.139	0.243	0.868	0.063	0.043
Precision	0.462	0.19	0.267	0.538	0.615	0.478	0.51	0.286	0.278	0.473	0.214	0.469	0.477	0.375	0.4
Recall	0.625	0.182	0.129	0.729	0.364	0.355	0.542	0.182	0.323	0.542	0.136	0.484	0.875	0.136	0.065
F-Measure	0.531	0.186	0.174	0.619	0.457	0.407	0.525	0.222	0.299	0.505	0.167	0.476	0.618	0.2	0.111
ROC Area	0.518	0.563	0.476	0.577	0.57	0.581	0.531	0.527	0.483	0.469	0.549	0.578	0.446	0.522	0.527

Table 6 shows the accuracy measures for self control where 'A , B, C' denotes the 'SCN, SCA and SCD' attributes respectively, which means self control neutral, agree and disagree of the Confusion Matrix. The decision tree classification results with respect to Table 6 are (64.3564, 66.3366, 64.3564, 64.3564, 64.3564) % of correctly classified instances and (35.6436, 33.6634, 35.6436, 35.6436, 35.6436)% of incorrectly classified instances of the self control attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the NB tree provides better results with high weighted average of TP Rate as 0.644, FP Rate as 0.484, Precision as 0.599, recall as 0.644, F-Measure 0.619 and ROC area as 0.587.

Table 7 shows the accuracy measures for emotionality where 'A, B, C' denotes the 'EN, EA, ED' attributes which means emotionality neutral, agree and disagree of the Confusion Matrix. The decision tree classification results with respect to Table 7 are with respect to Table 7 are (56.4356, 60.396, 51.4851, 47.5248, 60.396) % of correctly classified instances and (43.5644, 39.604, 48.5149, 52.4752, 39.604) % of incorrectly classified instances of the emotionality attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the J48 Graft tree provides better results with high weighted average of TP Rate as 0.604, FP Rate as 0.59, Precision as0.464, recall as 0.604, F-Measure 0.481 and ROC area as 0.439.

Table 8 shows the accuracy measures for sociability where 'A, B, C' denotes the 'SA, SN,SD' attribute which means sociability agree, neutral and disagree of the Confusion Matrix. The decision tree classification results with respect to Table 8 are (37.6238, 53.4653, 51.4851, 43.5644, 46.5347) % of correctly classified instances and (62.3762, 46.5347, 48.5149, 56.4356, 53.4653)% of incorrectly classified instances of the sociability attribute respectively using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the J48 Graft tree provides better results with high weighted average of TP Rate as

0.535, FP Rate as 0.335, Precision as 0.537, recall as 0.535, F-Measure 0.519 and ROC area as 0.577.

Table 9 shows the accuracy measures for Global TEI where 'A, B,C' denotes the GD attribute which means Global TEI disagree, neutral and agree of the Confusion Matrix. The decision tree classification results with respect to Table 9 are (61.3861, 62.3762, 57.4257, 47.5248, 61.3861) % of correctly classified instances and (38.6139, 37.6238, 42.5743, 52.4752, 38.6139)% of incorrectly classified instances of the self control attribute respectively.

Table 9: Accuracy Measures for Global TEI

Classifier	FT Tree			J48 Graft			LAD			NB			RANDOM		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
TP Rate	0	0	0.954	0	0	0.969	0.118	0.105	0.831	0.118	0.158	0.662	0	0.053	0.938
FP Rate	0	0.037	1	0.012	0.037	0.944	0.071	0.073	0.861	0.131	0.22	0.667	0.06	0	0.944
Precision	0	0	0.633	0	0	0.649	0.25	0.25	0.635	0.154	0.143	0.642	0	1	0.642
Recall	0	0	0.954	0	0	0.969	0.118	0.105	0.831	0.118	0.158	0.662	0	0.053	0.938
F-Measure	0	0	0.761	0	0	0.778	0.16	0.148	0.72	0.133	0.15	0.652	0	0.1	0.763
ROC Area	0.501	0.53	0.53	0.411	0.507	0.494	0.614	0.55	0.585	0.623	0.593	0.568	0.459	0.508	0.493

Table 10: Accuracy Measures for Global TEI

Classification Factors	Decision Tree Techniques	Kappa Statistics	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Root Relative Squared Error
Well-Being	FT Tree	0	0.1592	0.3445	72.1786 %	105.3501 %
	J48 Graft	0	0.2132	0.3269	96.6509 %	99.9518 %
	LAD	-0.1095	0.1997	0.3875	90.546 %	118.4992 %
	NB	0.1032	0.1923	0.3381	87.1607 %	103.3721 %
	RANDOM	0.0485	0.2076	0.3355	94.1027 %	102.582 %
Self-Control	FT Tree	0.0653	0.2438	0.4709	76.1714 %	118.4022 %
	J48 Graft	-0.0331	0.3199	0.4042	99.9595 %	101.6326 %
	LAD	0.1234	0.2801	0.4329	87.5089 %	108.8396 %
	NB	0.1721	0.2967	0.4298	92.7259 %	108.0567 %
	RANDOM	0.0345	0.3035	0.426	94.8395 %	107.1057 %
Emotionality	FT Tree	-0.0621	0.2998	0.5194	81.5094 %	121.4344 %
	J48 Graft	0.0168	0.3581	0.4411	97.3555 %	103.1296 %
	LAD	-0.0855	0.3859	0.4897	104.9151 %	114.4978 %
	NB	-0.0306	0.3904	0.4985	106.1304 %	116.5648 %
	RANDOM	0.0047	0.3549	0.4245	96.485 %	99.2654 %
Sociability	FT Tree	-0.034	0.4167	0.5908	98.5941 %	128.5776 %
	J48 Graft	0.2195	0.3852	0.4865	91.1441 %	105.8916 %
	LAD	0.0265	0.4224	0.5185	99.9473 %	112.8513 %
	NB	0.0805	0.4111	0.5126	97.2743 %	111.5707 %
	RANDOM	0.034	0.4251	0.4965	100.5769 %	108.0475 %
GTEI	FT Tree	-0.0437	0.2679	0.4836	76.2186 %	115.7738 %
	J48 Graft	-0.0042	0.352	0.4238	100.1431 %	101.4544 %
	LAD	0.0103	0.3196	0.4576	90.9161 %	109.5514 %
	NB	-0.0243	0.3419	0.4649	97.2726 %	111.2991 %
	RANDOM	-0.0043	0.3544	0.4418	100.8086 %	105.775 %
Academic Performance	FT Tree	0.1781	0.2641	0.4072	78.8732 %	99.7295 %
	J48 Graft	0.1462	0.3027	0.4317	90.4128 %	105.7299 %
	LAD	0.1358	0.2885	0.4396	86.1767 %	107.6774 %
	NB	0.1347	0.3111	0.438	92.916 %	107.2708 %
	RANDOM	0.0513	0.3248	0.4448	97.0098 %	108.9509 %
Debugging Skills	FT Tree	0.1153	0.2197	0.3906	87.2271 %	110.3332 %
	J48 Graft	0.1267	0.2379	0.3921	94.4496 %	110.7596 %
	LAD	0.0757	0.2413	0.3835	95.7987 %	108.3313 %
	NB	0.0857	0.2293	0.3717	91.0171 %	104.9974 %
	RANDOM	0.0239	0.2491	0.3581	98.8931 %	101.1405 %

Using FT Tree, J48 Graft tree, LAD tree, NB tree and Random tree. When comparing all the results, the LAD tree provides better results with high weighted average of TP Rate as 0.614, FP Rate as 0.618, Precision as 0.601, recall as 0.614, F-Measure 0.51 and ROC area as 0.49.

Table 10 discusses the comparative analysis of TEI factors with respect to decision tree classification techniques and error measures. The various factors taken for categorization are academic performance, debugging skills, well being, Self control, emotionality, sociability and Global TEI. The various decision tree techniques like FT tree, J48 Graft, LAD tree, NB tree and Random tree are applied on the classification factors. Error measures like kappa statistics mean absolute error, root mean squared error, relative absolute error and root relative squared error, which are used for cross validation.

5. Conclusion

The FT tree technique provides the best result with the highest kappa statistics value for academic performance. The LAD tree technique provides the best result with the highest kappa statistics value for GTEI. The NB Tree technique provides the best result with the highest kappa statistics value for well being and self control attributes. The J48 Graft tree technique provides the best result with the highest kappa statistics value for Emotionality, Sociability and debugging skills.

Acknowledgement

I thank the Lord Almighty for his gracious blessings. I extend my heartfelt gratitude to my parents. I thank Dr. K. David for spending his time to guide me in research. Finally, I express my thanks to my wife for supporting me.

References

[1] Andrew Cooper. and Petrides K.V., "A Psychometric Analysis of the Trait Emotional Intelligence Questionnaire–Short Form (TEIQue–SF) Using Item Response Theory," Journal of Personality Assessment, Vol. 92(5). pp. 449–457, 2010. https://doi.org/10.1080/00223891.2010.497426.

[2] Ashokkumar Vijaysinh Solanki, "Data Mining Techniques Using WEKA classification for Sickle Cell Disease", International Journal of Computer Science and Information Technologies, Vol. 5(4), pp. 5857-5860, 2014.

[3] Chandra E., and Nandhini K.,"Knowledge Mining from Student Data," European Journal of Scientific Research, Vol.47, No. 1, pp. 156 – 163, 2010.

[4] Deepa C., SathiyaKumari E., and Pream Sudha V., "Prediction of the Compressive Strength of High Performance Concrete Mix Using Tree Based Modeling," International Journal of Computer Applications (0975 – 8887), Volume 6- No. 5, 2010.

[5] Dorina Kabakchieva, "Predicting Student Performance by Using Data Mining Methods for Classification", Cybernetics and Information Technologies, Volume 13, No. 1, pp.61-72, 2013. https://doi.org/10.2478/cait-2013-0006.

[6] Jafar Tanha., Maarten van Someren., and Hamideh Afsarmanesh., "Semi-supervised self-training for decision tree classifiers", Springer, January 2015.

[7] Jai Ruby, and David K., "A study model on the impact of various indicators in the performance of students in higher education," International Journal of Research in Engineering and Technology, Vol. 3, Issue 5. Pp.750-755, 2014.

[8] Jiawei Han, Micheline Kamber, and Jian Pei, "Data Mining Concepts and Techniques" Morgan Kaufmann Publishers, USA, 3rd edition, 2012.

[9] Koushal Kumar, "Knowledge Extraction from Trained Neural Networks" International Journal of Information and Network Security, Vol. 1, No. 4, pp. 282-293, October 2012. https://doi.org/10.11591/ijins.v1i4.792.

[10] Lakshmi Devasena C, "Proficiency comparison ofladtree and reptree classifiers for credit risk forecast", International Journal On Computational Sciences & Applications, Vol. 5, No. 1, February 2015.

[11] Manuel J. Fonseca and Joaquim A. Jorge., "Indexing High-Dimensional Data for Content-Based Retrieval in Large Databases", Proceedings of the Eighth International Conference on Database Systems for Advanced Applications, IEEE Computer Society, 2003. https://doi.org/10.1109/DASFAA.2003.1192391.

[12] Maria Antoniate Martin V., Edison M., and George Gabriel Richard Roy, "An Association between Trait Emotional Intelligence and Personality Trait based on Students Performance," CiiT International Journal of Software Engineering and Technology, ISSN: 0974 – 9632, 2012.

[13] Maria Antoniate Martin V., "Evaluation of Object Oriented Programming Skills of Students with respect to Trait Emotional Intelligence based on Students Performance," International Journal of Computer Applications (0975 – 888), Volume 48– No.10, 2012.

[14] Maria Antoniate Martin V., and Lucia Agnes Beena T., "Prediction of Association among Numerical Aptitude, Programming Skills, Trait Emotional Intelligence on Students Performance," International Journal on Computer Science and Engineering (0975–3397), Volume 4, Issue 9 – pp. 1639-1646, 2012.

[15] Maria Antoniate Martin V., and David K., "Finding the best suited classification function to analyze the association of students debugging skills with Trait Emotional Intelligence", International Journal of Applied Engineering Research, Vol. 10, No. 5, pp. 3979-3982, 2015.

[16] Natalie L. Shipley., Mary Jo Jackson., and Sharon Larisa Segrest, "The effects of emotional intelligence, age, work experience, and academic performance," Research in Higher Educational Journal, 2010.

[17] Olaiya Folorunsho, "Comparative Study of Different Data Mining Techniques Performance in knowledge Discovery from Medical Database," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3. ISSN: 2277 128X, 2013.

[18] Payal P. Dhakate., Suvarna Patil., Rajeswari K., and Deepa Abin., "Preprocessing and Classification in WEKA Using Different Classifiers", Int. Journal of Engineering Research and Applications, Vol. 4, Issue 8, pp. 91-93, August 2014.

[19] Payal Dhakate., Rajeswari K., and Deepa Abin., "Analysis of Different Classifiers for Medical Dataset using Various Measures", International Journal of Computer Applications,Volume 111 – No 5, pp.20-24, February 2015.

[20] Payal P. Dhakate., Suvarna Patil., Rajeswari K., and Deepa Abink, "Preprocessing and Classification in WEKA Using Different Classifiers", Int. Journal of Engineering Research and Applications, Vol. 4, Issue 8 (Version 5), pp. 91-93, August 2014.

[21] Petrides K.V., Norah Fredericksonb, and Adrian Furnhamb, "The role of trait emotional intelligence in academic performance and deviant behavior at school," Personality and Individual Differences, Vol. 36, pp. 277–293, 2004. https://doi.org/10.1016/S0191-8869(03)00084-9.

[22] Petrides K.V., Yolanda Sangareau, Adrian Furnham, and Norah Frederickson, "Trait Emotional Intelligence and Children's Peer Relations at School" Blackwell Publishing Ltd. pp. 537 – 547, 2006.

[23] Pooja Sharma and Rupali Bhartiya, "Implementation of Decision Tree Algorithm to Analysis the Performance", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 1, Issue 10, pp. 861- 864, December 2012.

[24] Purva Sewaiwar and Kamal Kant Verma., Comparative Study of Various Decision Tree Classification Algorithm Using WEKA", International Journal of Emerging Research in Management and Technology, Volume-4, Issue-10, pp. 87-91, October 2015.

[25] Ramesh V., Parkavi P., and Yasodha P., "Performance Analysis of Data Mining Techniques for Placement Chance Prediction", International Journal of Scientific & Engineering Research, Volume2, Issue 8, pp. 1-7, August-2011.

[26] Rashedur M., Rahman., and Farhana Afroz., "Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis", Journal of Software Engineering and Applications, Vol. 6, pp. 85-97, 2013. https://doi.org/10.4236/jsea.2013.63013.

[27] Ritika, "Research on Data Mining Classification", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 4, pp. 329-332, April 2014.

[28] Ritika and Aman Paul., " Prediction of Blood Donors" Population using Data Mining Classification Technique", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 6, pp. 634-638, June 2014.

[29] Romero C., and Ventra S., "Educational data mining: A survey from 1995 to 2005", Expert Systems with Applications 33. Elsevier. pp. 135–146, 2007. https://doi.org/10.1016/j.eswa.2006.04.005.

[30] Sanchez-Ruiz M.J., Hernández-Torrano D., Pérez-González J.C., Batey M., and Petrides K.V., "The Relationship between Trait Emotional Intelligence and Creativity across Subject Domains," Springer, 2011.

[31] Sharon Dominick, and Abdul Razak T., "Improving the Cluster Efficiency on Sea Level Rise Dataset using Data Discretization," International Journal of Computer Applications, Volume 102, pp.13, 2014. https://doi.org/10.5120/17875-8856.

[32] Sherine Dominick, and Abdul Razak T., "Analyzing the Student Performance using Classification Techniques to find the better Suited Classifier," International Journal of Computer Applications, Volume 104, pp. 4, 2014. https://doi.org/10.5120/18187-9092.

[33] Shomona Gracia Jacob and Geetha Ramani R., "Evolving efficient clustering and classification patterns in lymphography Data through data mining techniques", International Journal on Soft Computing, Vol.3, No.3, pp.119-132, August 2012. https://doi.org/10.5121/ijsc.2012.3309.

[34] Stella Mavroveli., Petrides K.V., Carolien Rieffe, and Femke Bakker, "Trait emotional intelligence, psychological well-being and peer-rated social competence in adolescence," British Journal of Developmental Psychology, Vol. 25. pp. 263–275, 2007. https://doi.org/10.1348/026151006X118577.

[35] Stella Mavroveli, Petrides K.V., Chloe Shove, and Amanda Whitehead, "Investigation of the construct of trait emotional intelligence in children," Eur Child Adolesc Psychiatry, Vol. 17. pp. 516–526, 2008. https://doi.org/10.1007/s00787-008-0696-6.

[36] Trilok Chand Sharma and Manoj Jain., "WEKA Approach for Comparative Study of Classification Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 4, April 2013.

[37] Wisaeng K., "A Comparison of Different Classification Techniques for Bank Direct Marketing", International Journal of Soft Computing and Engineering, Vol. 3, Issue-4, pp. 116-119, September 2013.

[38] XindongWu et al., "Top 10 algorithms in data mining", Springer-Verlag London Limited, 2007.

[39] Yasodha P. and Ananthanarayanan N.R., "Comparative Study of Diabetic Patient Data's Using Classification Algorithm in WEKA Tool", International Journal of Computer Applications Technology and Research, Vol. 3, Issue – 9, pp. 554-558, 2014.

[40] Yogendra Kumar Jain and Upendra, "An efficient Intrusion Detection Based on Decision Tree Classifier Using Feature Reduction", International Journal of Scientific and Research Publicaiton, Volume 2, Issue 1, pp. 1 – 6, January 2012.