IJCAT International Journal of Computing and Technology, Volume 1, Issue 2, March 2014
ISSN : 2348 - 6090
www.IJCAT.org

Performance Analysis of Tree Classifiers using on Engineering Student's Recruitment Dataset

[1] **Samrat Singh,** [2] **Dr. Vikesh Kumar**

[1] Ph.d Research Scholar, School of Computer Science & IT, Singhania University
Rajasthan (India)

[2] Professor & Director, Neelkanth Institute of Technology
Meerut (India)

Abstract - Mining in education environment is called Educational Data Mining. Today, Data Mining is a powerful tool for academic intervention. Tree Classifiers have main role to developing new methods to discover knowledge from educational database and can used for decision making in educational system. In this paper, we collected the student's data from engineering institute that have different information about their previous and current academics records like student's Name, Branch, 10th, 12th, B.Tech passing percentage and final grade & then apply different J48, NBTree, REPTree, DecisionStump, RandomForest and RandomTree (Tree Classifiers) algorithm using Data Mining tools (WEKA) for analysis the students academics performances. This paper deals with a comparative study of various Tree Classifiers data mining algorithms for the performance analysis of the student's academic records and check J48 & RandomTree algorithms or Tree Classifiers are optimal for classifying students' based on their final grade. Before analysis we are classifies the Students' performance into Excellent, Good and Average categories.

Keywords **– Data Mining, Discover knowledge, Technical Education, Educational Data, Mining, Classification, WEKA, Tree Classifiers.**

1. Introduction

Data Mining is a process of extracting previously unknown, valid, potentional useful and hidden patterns from large data sets (Connolly, 1999). As the amount of data stored in educational databases is increasing rapidly. In order to get required benefits from such large data and to find hidden relationships between variables using different data mining techniques developed and used (Han and Kamber, 2006). There are increasing research interests in using data mining in education. This new emerging field, called Educational Data Mining, concerns with developing methods that discover knowledge from data come from educational environments [1]. The data can be collected form historical and operational data reside in the databases of educational institutes. The student data can be personal or academic. Also it can be collected from e-learning systems which have a vast amount of information used by most institutes [2][3]. Educational data mining used many techniques such as decision trees, neural networks, k-nearest Neighbor, Naïve Bayes, support vector machines and many others. Using these methods many kinds of knowledge can be discovered such as association rules, classifications and clustering. The discovered knowledge can be used to better understand students' behavior, to assist instructors, to improve teaching, to evaluate and improve e-learning systems , to improve curriculums and many other benefits [4] [1].

Performance monitoring involves assessments which serve a vital role in providing information that is geared to help students, teachers, administrators, and policy makers take decisions.[5] The changing factors in contemporary education has led to the quest to effectively and efficiently monitor student performance in educational institutions, which is now moving away from the traditional measurement & evaluation techniques to the use of DMT which employs various intrusive data penetration and investigation methods to isolate vital implicit or hidden information.

Due to the fact that several new technologies have contributed and generated huge explicit knowledge, causing implicit knowledge to be unobserved and stacked away within huge amounts of data.

The main attribute of data mining is that it subsumes Knowledge Discovery (KD) which according to [6] is a nontrivial process of identifying valid, novel, potentially useful and ultimately understandable Patterns in data processes, thereby contributing to

predicting trends of outcomes by profiling performance attributes that supports effective decisions making. This paper deploys theory and practice of data mining as it relates to student's performance in their qualifications.

The main objective of this paper is to use data mining methodologies to study student's performance in their qualifications. Data mining provides many tasks that could be used to study the student performance. In this research, the classification tasks are used to evaluate student's performance and as there are many approach that are used for data classification. Information's like student's course Branch, passing % of 10^{th}, passing % of 12^{th} and passing % of B.Tech were collected from the student's database, to classify the performance grade. This paper also investigates the accuracy of different applied Tree classifiers for classification of student performance (Final Grade).

2. Educational Data Mining

Education is an essential element for the betterment and progress of a country. It enables the people of a country civilized and well mannered. Educational Data Mining is an emerging discipline concerned with developing methods for exploring the unique types of data that come from educational database. Mining in educational environment is called Educational Data Mining, concern with developing new methods to discover knowledge from educational databases (Galit, 2007) (Erdogan and Timor 2005) ,in order to analyze students trends and behaviors toward education(Alaa el-Halees , 2009). Lack of deep and enough knowledge in higher educational system may prevent system management to achieve quality objectives, data mining methodology can help bridging this knowledge gaps in higher education system.

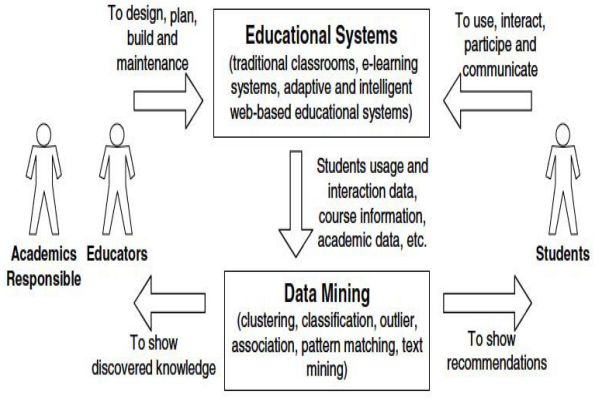

Figure 1: The Cycle of Data Mining in Educational Systems [15].

3. Data Mining Definitions

Data mining, also popularly known as Knowledge Discovery in Database, refers to extracting or "mining" knowledge from large amounts of data. Data mining techniques are used to operate on large volumes of data to discover hidden patterns and relationships helpful in decision making. While data mining and knowledge discovery in database are frequently treated as synonyms, data mining is actually part of the knowledge discovery process.

3.1 Classification

Classification is the most commonly applied data mining technique, which employs a set of pre-classified attributes to develop a model that can classify the population of records at large. This approach frequently employs decision tree or neural network-based classification algorithms. The data classification process involves learning and classification .In learning the training data are analyzed by classification algorithm. In classification test data are used to estimate the accuracy of the classification rules. If the accuracy is acceptable the rules can be applied to the new data tuples. The classifier-training algorithm uses these pre-classified attributes to determine the set of parameters required for proper discrimination. The algorithm then encodes these parameters into a model called a classifier. The Classification methods used for the comparative study are discussed in brief.

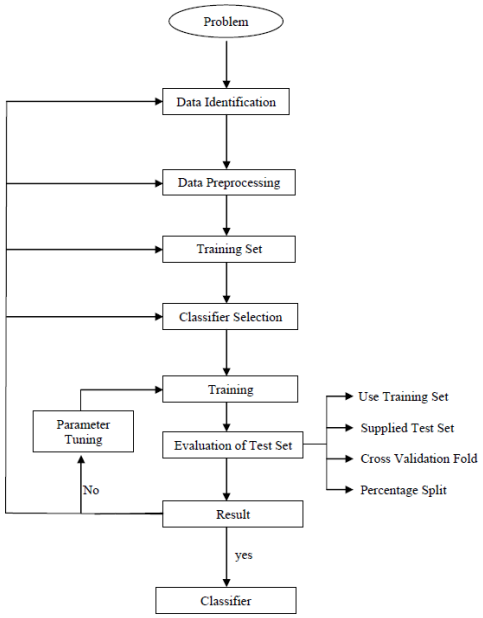

Figure -2: Process of Supervised Classification.

3.2 Tree Classifiers

Suppose we have a set of classes Y and a set of objects X such that each x belongs to a single y. For each x in X, we have n pieces of data {x1, …., xn}. Based on these pieces of information, we want to recreate the true mapping from X to Y.

Tree classifiers do this by performing iterated binary separations of the data space that try to isolate subsets of the data space that are close to homogeneous in class.

Each node in a tree classifier represents a subset of the data space. At non-leaf nodes, a decision process called a split is applied that partitions the node into two subsets that become new child nodes. This process is then potentially reapplied to each of the resulting children.

The resulting set of leaf nodes forms a partition of the whole data space. Each leaf node is assigned a class. The classifier can then be applied to a data point by beginning at the root and moving down the tree by applying each split to the data point until arriving at a root node and assigning the root node's prescribed classification.

To construct a tree classifier, we need to define three processes: how to select splits at each node, how to decide when to stop splitting a node, and how to decide which class to assign to each leaf node.

Example of Classifiers- *BFTree, DecisionsStump, J48 NBTree, SimpleCart, RandomForest, FT, RandomTree, LADTree, REPTree.*

4. Related Work

Data mining in higher education is a recent research field and this area of research is gaining popularity because of its potentials to educational institutes.

Data Mining can be used in educational field to enhance our understanding of learning process to focus on identifying, extracting and evaluating variables related to the learning process of students as described by Alaa el-Halees [7]. Mining in educational environment is called Educational Data Mining.

Han and Kamber [8] describes data mining software that allow the users to analyze data from different dimensions, categorize it and summarize the relationships which are identified during the mining process.

Pandey and Pal [9] conducted study on the student performance based by selecting 600 students from different colleges of Dr. R.M.L. Awadh University, Faizabad, India. By means of Bayes Classification on category, language and background qualification, it was found that whether new comer students will performer or not.

Al-Radaideh, et al [10] applied a decision tree model to predict the final grade of students who studied the C++ course in Yarmouk University, Jordan in the year 2005. Three different classification methods namely ID3, C4.5 and the Naïve Bayes were used. The outcome of their results indicated that Decision Tree model had better prediction than other models.

Data mining applications in higher education given in [11], they concluded with that the Data mining is a powerful analytical tool that enables educational institutions to better allocate resources and staff to proactively manage student outcomes and improve the effectiveness of alumni development.

Varsha, Anuj, Divakar, R.C Jain [13] applied four classification methods on student academic data i.e Decision tree (ID3), Multilayers perceptron, Decision table & Naïve Bayes classification method.

Brijesh kumar & Saurabh Pal [14] study the data set of 50 students from VBS Purvanchal University, Jaunpur (U.P). As there are many approaches that are used for data Classification, the decision tree method is used here. Information's like Attendance, Class test, Seminar and Assignment marks were collected from the student's previous database, to predict the performance at the end of the semester.

Sunita B Aher, Mr. LOBO L.M.R.J [16] defines the Data Mining in Educational System Using WEKA. They classify the student performance using WEKA (Data Mining Tool).

Tongshan Chang, & Ed.D [17] introduces a real project to assist higher education institutions in achieving enrollment goals using data mining techniques Furthermore, the results also provide evidence that data mining is an effective technology for college recruitment. It can help higher education institutions mange enrollment more effectively.

R. R. Kabra & R. S. Bichkar [18] shows that students past academic performance can be used to create the model using decision tree algorithm that can be used for

prediction of student's performance in First Year of engineering exam.

Neelam Naik & Seema Purohit [19] created classification trees for MCA result prediction and placement prediction of students are tested for validation data.

P.K.Srimani & Annapurna S Kamath [20] define a comparative study of the application of various data mining algorithms for the performance analysis of the learning model.

Performance analysis is the analysis of the data stored by the learning model in the mathematical pathway database which is used to track the progress of each student. The analysis classifies the performance of a student into average, below average and above average categories.

Samrat Singh & Dr. Vikesh Kumar [21] introduce the use of data mining process in a student's database using classification data mining techniques (decision tree method etc). The information generated after the analysis of data mining techniques on student's data base is helpful for executives for training & placement department of engineering colleges. This work classifies the categories of student's performance in their academic qualifications.

Table 1: Values of Final Grade

Attribute	Description	Possible Values
Student_ Name	Student's name in B.Tech course.	{alphabets Characters}
Branch	Student's branch in B.Tech course.	{CS, IT, EC}
HighSchool_ Percentage (10^{th} %)	Percentage of marks obtained in 10^{th} class exam.	{ First ≥ 60% Second > 45 & < 60 % Third > 35 & < 45 % }
Intermediate _ Percentage (12^{th} %)	Percentage of marks obtained in 12^{th} class exam.	{ First ≥60% Second > 50 & < 60 % }
B.Tech_ Percentage (B.Tech %)	Percentage of marks obtained in B.Tech course.	{ First ≥ 60% Second > 50 & < 60 % }
Final_ Grade	Final Grade obtained after analysis the passing percentage of 10^{th} ,12^{th} , B.Tech	{ Excellent, Good, Average }

Sudheep Elayidom, Sumam Mary Idikkula & Joseph Alexander [22] proved that the technology named data mining can be very effectively applied to the domain called employment prediction, which helps the students to choose a good branch that may fetch them placement. A generalized framework for similar problems has been proposed.

V.Ramesh, P.Parkavi & P.Yasodha [23] defines an attempt to use classification algorithms for predicting the student performance and comparing the performance of NaiveBayes Simple, MultiLayer Perception, SMO, J48, and REP Tree.
S. Anupama Kumar and Dr. Vijayalakshmi M.N [24] classification techniques can be applied on educational data for predicting the student's outcome and improve their results. The efficiency of various decision tree algorithms can be analyzed based on their accuracy and time taken to derive the tree. The predictions obtained from the system have helped the tutor to identify the weak students and improve their Performance.

5. Proposed Work

5.1 Data Collection & Preparations

The data set used in this study was obtained from the different branches of Engineering College. Initially size of the data is 540. The data sets have six attributes like student's Name, Branch, passing percentage (%) of 10^{th} class, passing percentage (%) of 12^{th} class and passing percentage (%) and Final Grade for analysis.

We discretized the numerical attributes to categorical ones. For example, variable X (X = x_0, x_1, x_2 Where $x_{0=}10^{th}$ %, $x_{1=}12^{th}$ %, $x_{2=}$B.Tech %) is common variable of student's passing percentage (%) in 10^{th}, 12^{th} & B.Tech. We grouped all grades into three groups Excellent, Good, Average as described in table below.

Table 2: The Symbolic Attribute Description

Final_Percentage	Final_Grade
X ≥ 60%	Excellent
X ≥ 45%	Good
X ≥ 35%	Average

In the same way, we descretized other attributes such as student's course Branch, passing % of 10^{th}, passing % of 12^{th}, passing % of B.Tech. Finally the most significant attributes presented in following table:-

The domain values for some of the variables were defined for the present investigation as follows:

- **Branch –** Student's branch in they are enrolled in B.Tech Course. Branch split in four classes: *CS, IT, EC*.

- **HighSchool_Percentage (10th %) --** Student's passing Percentage (%) in 10th class. 10th % is split into three classes: *First- ≥60% Second - >45% and <60%, Third - >35% and < 45%.*

- **Intermediate_Percentage (12th %) --**Student's passing Percentage (%) in 12th class. For admission in B.Tech course minimum 50% marks are needed in 12th class.
  So 12th % is split into two classes:*First- ≥60% Second>50% and <60%.*

- **B.Tech_Percentage (B.Tech %) --**Student's passing percentage (%) in B.Tech Course. In B.Tech course Minimum 50 marks is compulsory for passing. So B.Tech % is split into two classes: *First- >60% Second - >50% and <60%.*

- **Final_Grade –**The value of final grade (X) will be finding after analysis of rule sets of Student's passing percentage (%) in 10th (x_0), 12th (x_1), B.Tech (x_2). The final grade is divided into three categories: *Excellent, Good, Average.*

6. WEKA (Data Mining Tool)

The Weka workbench contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality. It is freely available software. It is portable & platform independent because it is fully implemented in the Java programming language and thus runs on almost any modern computing platform. Weka has several standard data mining tasks, data preprocessing, clustering, classification, association, visualization, and feature selection. The WEKA GUI chooser launches the WEKA's graphical environment which has six buttons: Simple CLI, Explorer, Experimenter, Knowledge Flow, ARFF-Viewer, & Log.

The *Explorer* interface has several panels that give access to the main components of the workbench:

1. The *Preprocess* panel imports the data from a database, a CSV file, ARFF etc., and preprocesses this data using *filtering* algorithm which can be used to transform the data from one format to other e.g. numeric attributes into discrete ones. It is also possible to delete instances and attributes according to specific criteria on the preprocess screen. It is also possible to view the graph for particular attribute.

2. The *Classify* panel allows the user to apply classification and regression algorithms (e.g. NaiveBays algorithm, ADTree, ID3 Tree , J48 Tree, ZeroR rules etc) to the dataset estimate the accuracy of the resulting model. It is also possible to visualize erroneous predictions, ROC curves, etc. Result of classification can be seen in classifier output area .

3. The *Cluster* panel is used to access the clustering techniques in Weka, e.g., the simple k-means, EM, DBScan, XMeans algorithm. Sometimes it is necessary to ignore some attribute while using the clustering algorithm, so it is possible with Ignore Attribute button.

4. The *Associate* panel gives access to association rule e.g. Apriori, PredictiveApriori algorithm. Once the appropriate parameter for association rule is chosen then result list allows the result set to viewed or saved.

5. The *Select attributes* panel allows to search among all possible combination of attribute in dataset, which subset of attribute is best for making prediction.

6. The *Visualize* panel visualizes 2D plots of current relation.

7. Result & Discussion

The data set of 540 students used in this study was obtained from the Engineering College of B.Tech course.

Table 3: Rule Set for Student's Final Grade

IF 10th % ="First" AND 12th % ="First" AND B.Tech % = "First" THEN Final_Grade = "Excellent"
IF 10th % ="Second" AND 12th % ="First" AND B.Tech % = "First" THEN Final_Grade = "Good"
IF 10th % ="Third" AND 12th % ="First" AND B.Tech % = "First" THEN Final_Grade = "Average"
IF 10th % ="First" AND 12th % ="Second" AND B.Tech % = "First" THEN Final_Grade = "Good"
IF 10th % ="Second" AND 12th % ="Second"AND B.Tech % = "First" THEN Final_Grade = "Average"
IF 10th % ="Third" AND 12th % ="Second" AND B.Tech % = "First" THEN Final_Grade = "Average"
IF 10th % ="First" AND 12th % ="First" AND B.Tech % = "Second" THEN Final_Grade = "Average"
IF 10th % ="Second" AND 12th % ="First" AND B.Tech % = "Second" THEN Final_Grade = "Average"
IF 10th % ="Third" AND 12th % ="First" AND B.Tech % = "Second" THEN Final_Grade = "Average"
IF 10th % ="First" AND 12th % ="Second" AND B.Tech % = "Second" THEN Final_Grade = "Average"
IF 10th % ="Second" AND 12th % ="Second"AND B.Tech % = "Second" THEN Final_Grade = "Average"
IF 10th % ="Third" AND 12th % ="Second" AND B.Tech % = "Second" THEN Final_Grade = "Average"

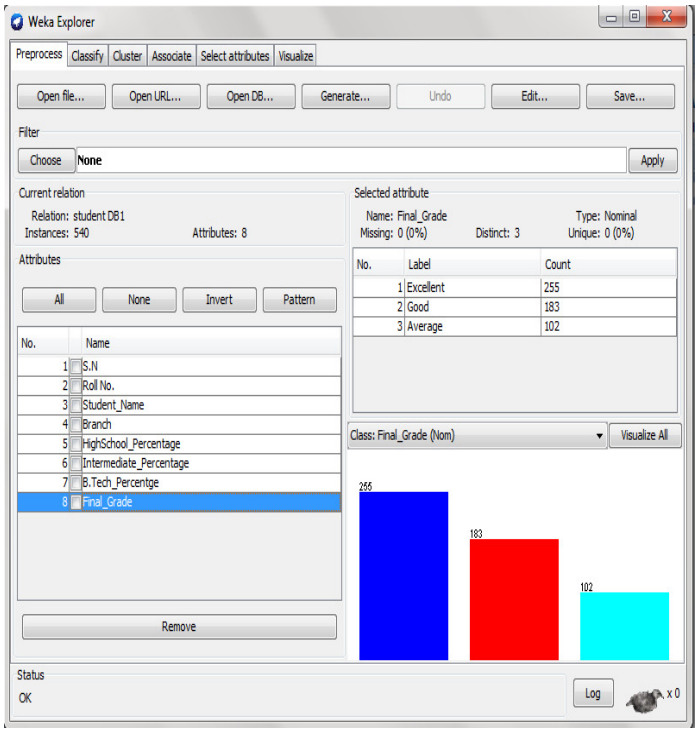

Figure 3: Weka 3.6.8 with Explorer window open with Student's Dataset.

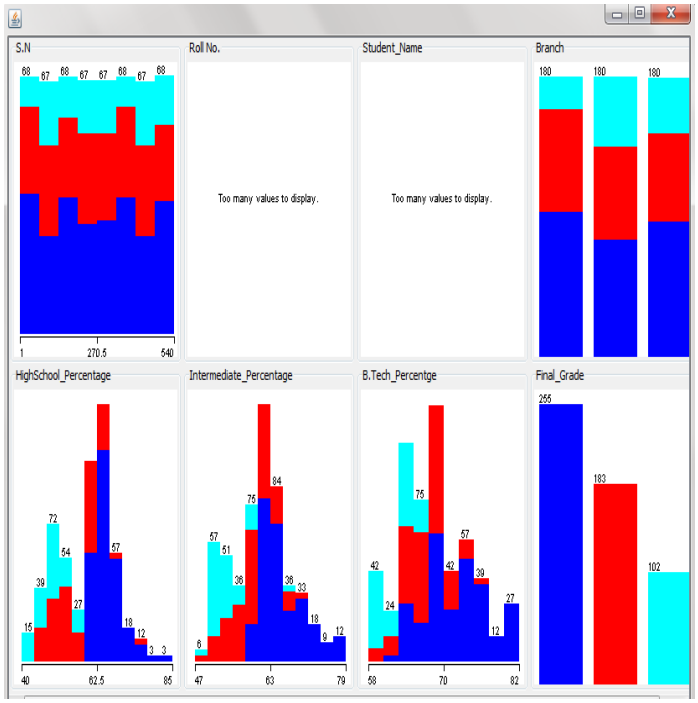

Figure 4: Visualized Result of Student's Dataset from Weka.

Table 4: Result From Different Classifier Using Weka

Table 5 : Best Classifiers of Different Measurements

Classifier Name→	J48	NBTree	REPTree	Decision Stump	Random Forest	Randam Tree
Total No.of Instances	540	540	540	540	540	540
Correctly Classified Instances	540 (100%)	540 (100%)	255 (47.2%)	360 (66.7%)	540 (100%)	540 (100%)
Incorrectly Classified Instances	0 (0%)	0 (0%)	285 (52.8%)	180 (33.3%)	0 (0%)	0 (0%)
Time Taken to build the Model	0.01 Second	0.42 Second	0.01 Second	0.01 Second	0.05 Second	0.01 Second
Kappa Statistic	1	1	0	.4242	1	1
Mean Absolute Error	0	0.0019	0.4172	.2752	0.09	0
Root Mean Squared Error	0	0.0038	0.457	0.371	0.12	0
Relative Absolute Error	0%	.4449%	99.96%	65.87%	23.40%	0%
Root Relative Squared Error	0%	.8242%	99.99%	81.17%	27.04%	0%
TP Rate	1	1	.472	.667	1	1
FP Rate	0	0	.472	.226	0	0
Precision	1	1	.223	.534	1	1
Recall	1	1	.472	.667	1	1
F-Measure	1	1	.303	.592	1	1
ROC Value	1	1	.5	.792	1	1
Confusion Matrix	255 0 0 0 183 0 0 0 102	255 0 0 0 183 0 0 0 102	255 0 0 0 183 0 0 0 102	255 0 0 183 0 0 102 0 0	255 0 0 78 105 0 0 102 0	255 0 0 0 183 0 0 0 102

Classifier Name→	J48	NBTree	REPTree	Decision Stump	Random Forest	Randam Tree
Total No.of Instances	540	540	540	540	540	540
Correctly Classified Instances	J48	NBTree	--	--	Random Forest	Randam Tree
Time Taken to build the Model	J48	--	REPTree	Decision Stump	--	Randam Tree
Kappa Statistic	J48	NBTree	--	--	Random Forest	Randam Tree
Mean Absolute Error	J48	--	--	--	--	Randam Tree
Root Mean Squared Error	J48	--	--	--	--	Randam Tree
Relative Absolute Error	J48	--	--	--	--	Randam Tree
Root Relative Squared Error	J48	--	--	--	--	Randam Tree
TP Rate	J48	NBTree	--	--	Random Forest	Randam Tree
FP Rate	J48	NBTree	--	--	Random Forest	Randam Tree
Precision	J48	NBTree	--	--	Random Forest	Randam Tree
Recall	J48	NBTree	--	--	Random Forest	Randam Tree
F-Measure	J48	NBTree	--	--	Random Forest	Randam Tree
ROC Value	J48	NBTree	--	--	Random Forest	Randam Tree
Confusion Matrix	J48	NBTree	--	--	Random Forest	Randam Tree
Total points (14) →	14	09	01	01	09	14

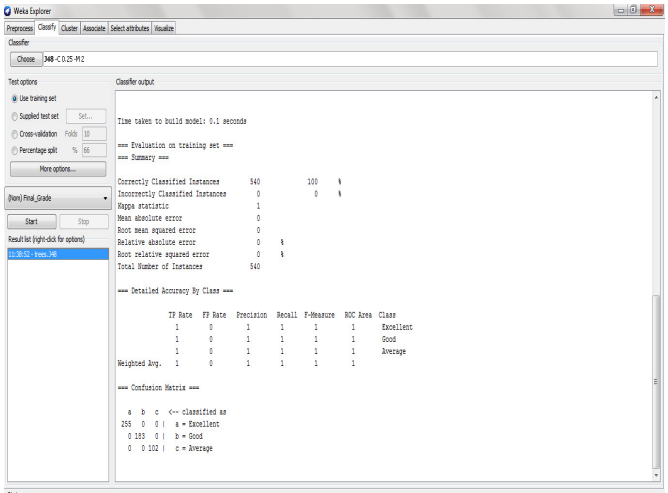

Figure 5: J48 Classifier Result of Student's Dataset from Weka.

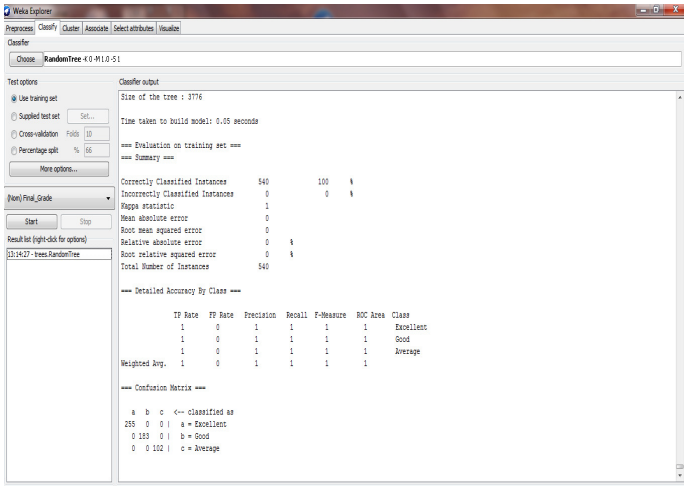

Figure 6: RandomTree Classifier Result of Student's Dataset from Weka

The results obtained from the various data mining algorithms viz, J48, NBTree, REPTree, DecisionStump, RandomForest and RandomTree on the data set for different branches of students are tabulated and the performance analyzed Comparison table gives the total no. of instances, Correctly classified and Incorrectly classified instances, Time taken to build a model, Confusion matrix, Kappa statistics and ROC value. The interpretations of the results based on these parameters are as follows:

In figure 4, WEKA generates the performance of student's Final Grade are given. According to different branches have 255 (47%) Excellent students, 183 (34%) Good students and 102 (19%) Average students.

In Table 4, reveals that (i) the classifiers J48, NBTree, RandomForest and RandomTree algorithm are found to be very efficient and accurate. In this case the correctly classified instances are 100%, (ii) The Time taken by J48, REPTree, DecisionStump and RandomTree are 0.01 Second to build the model but IB1 Algorithm is found 100% correctly classified instances. (iii) The classifiers J48, NBTree, RandomForest and RandomTree algorithm are found the accurate diagonal elements of the confusion matrix predicts the correctly classified instance. (iv) with respect to Kappa Statistic the classifiers J48, NBTree, RandomForest and RandomTree algorithm are preferred. (v) Classifiers J48, RandomTree have not any error value of MAE, RMSE, RAE, RRSE in their results. (vi) Classifiers J48, NBTree, RandomForest and RandomTree are found the accurate value of the area (TP, FP, Precision, Recall, F-Measure, ROC Value). So according to these Tree classifiers performance analysis of J48 & RandomTree secure total 14 points out of 14 points and generates the efficient and accurate results on this type of data set.

8. Conclusion & Future Work

This work is an attempt to use Data Mining techniques to analyze students' academic data and to enhance the quality of technical educational system. In this work we applied six classification (Tree Classifiers) methods on student data i.e. J48, NBTree, REPTree, DecisionStump, RandomForest and RandomTree Classification method .We notice that according to experimental result J48 & RandomTree Classifiers is most suitable method for this type of student dataset. The Higher management's executives for training & placement department of engineering colleges or Company Executives can use such classification model to measures or visualized the students' performance according to the extracted knowledge.

For future work, this study will be helpful for improve the classification accuracy of other tree classifiers that could be given the 100% accuracy of the given instances. We can be generating the information after implementing the others data mining techniques like clustering, Predication and Association rules etc with help of Data Mining tools on different eligibility criteria of industry recruitment process.

References

[1] Romero, C. , Ventura, S. and Garcia, E., "Data mining in course management systems: Moodle case study and Tutorial". Computers & Education,Vol. 51, No. 1. pp.368- 384. 2008.

[2] Machado, L. and Becker, K. "Distance Education: A Web Usage Mining Case Study for the Evaluation of Learning Sites". Third IEEE InternationalConference on Advanced Learning Technologies , 2003.

[3] Mostow,J and Beck , J., "Some useful tactics to modify, map and mine data from intelligent tutors". Natural Language Engineering 12 (2), 195- 208. 2006.

[4] Romero,C. and Ventura, S. ,"Educational data Mining: A Survey from 1995 to 2005".Expert Systems with Applications (33) 135-146. 2007.

[5] Council N. "Knowing What Student Knows. The Science and Design of Educational Assessment ".National Academic Press. Washington, D.C. 2001

[6] Frawley, W.J., Piatetsky-Shapiro, G and Matheus, "Knowledge Discovery databases: An overview In": Piattetsky-Shapiro and Frawley, W. J. (eds) Knowledge Discovery in Databases, AAAI/MI .1991.pp 1-27.

[7] Alaa el-Halees "Mining students data to analyze e-Learning behavior: A Case Study", 2009.

[8] J. Han and M. Kamber, "Data Mining: Concepts and Techniques,"Morgan Kaufmann, 2000.

[9] U . K. Pandey, and S. Pal, "Data Mining: A prediction of performer or underperformer singclassification",(IJCSIT) International Journal of Computer Science and Information Technology, Vol. 2(2), pp.686-690 ISSN 0975- 9646, 2011.

[10] Q. A. AI-Radaideh, E. W. AI-Shawakfa, and M. I. AI-Najjar, "Mining student data using decision tree". ,International Arab Conference on Information Technology, Yarmouk University, Jordan,2006.

[11] Jing Luan "Data mining application in higher education" .Chief planning and Research Officer, Cabrillo College founder knowledge Discovery ,2006.

[12] J. R. Quinlan, "Introduction of decision tree: Machine learn", 1: pp.86-106, 1986.

[13] Varsha, Anuj, Divakar, R.C Jain, "Result analysis using Classification techniques", International Journal of

Computer Applications (0975-8887)Volume 1-No. 22, 2010.

[14] Brijesh kumar & Saurabh Pal , " Mining educational data to Analyze students performance ", International Journal of Advanced Computer Science & Applications Volume 2, No- 6, 2011.

[15] K. H. Rashan, Anushka Peiris, "Data Mining Applications in the Education Sector", MSIT, Carnegie Mellon University, retrieved on 2011.

[16] Sunita B Aher, Mr. LOBO L.M.R.J, "Data Mining in Education System Using WEKA" International Conference on Emerging Technology Trends (ICETT) 2011 Proceedings published by International Journal of Computer Applications® (IJCA).

[17] Tongshan Chang, Ed.D., "Data Mining: A Magic Technology for College Recruitment" Paper of Overseas Chinese Association for Institutional Research (www.ocair.org) in 2008.

[18] R. R. Kabra & R. S. Bichkar "Performance Prediction of Engineering Students using Decision Trees" International Journal of Computer Applications (0975 – 8887) Volume 36– No.11, December 2011.

[19] Neelam Naik & Seema Purohit "Prediction of Final Result and Placement of Students using Classification Algorithm"International Journal of Computer Applications (0975 – 8887) Volume 56– No.12, October 2012.

[20] P.K.Srimani & Annapurna S Kamath "Data Mining Techniques for the Performance Analysis of Learning Model – A Case Study International Journal of Computer Applications (0975 – 8887) Volume 53– No.5, September 2012

[21] Samrat Singh & Dr. Vikesh Kumar "Classification of Student's data Using Data Mining Techniques for Training & Placement Department in Technical Education" International Journal of Computer Science and Network (IJCSN) Volume 1, Issue 4, August 2012 www.ijcsn.org ISSN 2277-5420

[22] Sudheep Elayidom , Sumam Mary Idikkula & Joseph Alexander "A Generalized Data mining Framework for Placement Chance Prediction Problems" International Journal of Computer Applications (0975– 8887) Volume 31– No.3, October 2011.

[23] V.Ramesh, P.Parkavi & P.Yasodha "Performance Analysis of Data Mining Techniques for Placement Chance Prediction" International Journal of Scientific & Engineering Research Volume 2, Issue 8, August-2011, ISSN 2229-5518

[24] S. Anupama Kumar and Dr. Vijayalakshmi M.N "Implication of classification Techniques in Predicting Student's Recital" International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.1, No.5, September 2011.

First Author

Samrat Singh is Ph.D Research Scholar in School of Computer Science & IT of Singhania University, Rajasthan (India) .His area of specialization in Educational Data Mining. He did complete his master degrees MCA from UPTU, Lucknow (India), M.phil from Alagappa University, Tamilnadu (India) and M.Tech from KSOU, Mysore (India). He published many research papers in reputed conferences and journals on different issues.

Second Author

Dr. Vikesh Kumar is working as Professor & Director in NIT, Meerut (India). He did complete his doctorate degree from Gurukul kangri Vishwavidyalaya, Haridwar (India). He has more than 60 research papers in reputed international & national journals. He got completed more than 10 candidates of M.Phil and M.Tech degree under his supervision. Currently he also guides to many candidates of Ph.d program under his supervision.