Analysis & Survey of Different Data Mining Techniques for Predicting Student's Performance

Bhoomi H. Parmar[1], Prof. Kruti Khalpada[2]

[1]*M.E. (Computer Engineering) Student,* [2] *Assistant Professor*
Dept. of Computer Engineering
Atmiya Institute of Technology and Science Rajkot, Gujarat, India

[1]`bhp890@gmail.com`
[2]`knkhaplada@aits.edu.in`

Abstract— **Data Mining has wide applicability due to wide ease of use of large amount of data and requirement of storage as per the need. Data mining techniques are widely useful in educational data mining for analysis of student data. In educational area data mining different data mining techniques like classification, clustering, association rule mining, decision tree method have been used to analyze student's learning manners, their mindset, forecasting their result, group them, and for finding out different patterns. Educational data mining helps for improving student's performance, for managing the student database and for managing the institute. This paper focuses on different data mining techniques that are useful for predicting student performance.**

Keywords— **Data Mining, Association Rules, Classification, Clustering, Decision Tree, Neural Networks**

I. INTRODUCTION

Data mining refers to extracting or "mining" knowledge from large amounts of data [1]. Data mining is a method of extracting required data from the large amount of data, where data can be stored in huge databases. Data mining is process of analysing data to gather the information from large amount of data. The main goal of data mining process is to extract information from large amount of data and to translate raw data into meaning full information. Data Mining helps to identify patterns among data. Data mining techniques are widely useful in different fields like games, business, medical diagnosis, science & engineering and many more. Data mining task includes different techniques from different disciplines such as statistics, machine learning, pattern recognition and many more. Different functionalities of data mining includes discovery of patterns, association, correlations, prediction, classification, clustering, trend analysis and many more. Data mining task also includes decision support system for more accurate prediction. Data mining technology uses different data mining tools and methods for decision making and mining the data.[1].

II. EDUCATIONAL DATA MINING

Educational data mining includes the application of data mining, machine learning and statistics for analysing and prediction of data gathered from educational institutes. Educational data mining refers to techniques or a tool that extracts patterns from large amount of student related data. In educational area data mining different techniques helps to analyse student's learning behaviour and for finding out different patterns.
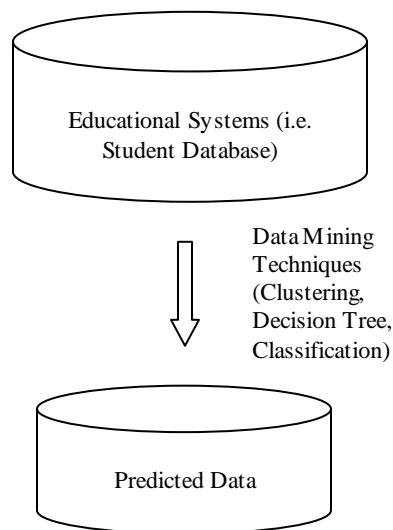

Fig 1: Educational data mining process

Educational Data Mining is concerned about both kinds of data that how the student learns and what they learn that is about predicting their result. Educational data mining helps for organization student performance and their databases. The Educational Data Mining process converts raw data coming from educational systems into useful information that could help in educational research and finding student learning behaviour.

Here the main goal is to translate raw data into meaningful or knowledgeable information for analysis and for finding appropriate decisions about learning habit of students, for forecasting their performance and managing student database. Thus, Educational Data Mining generally consists of four different phases which are as follow: [2][3]

1. In the first phase of the EDM process (not counting pre-processing) different similarities among the observations has been find out. This includes searching through a educational database with the objective of finding

National Conference on Emerging Trends in Computer, Electrical & Electronics (ETCEE-2015)
International Journal of Advance Engineering and Research Development (IJAERD)
e-ISSN: 2348 - 4470 , print-ISSN:2348-6406,Impact Factor:3.134

consistent relationships or similarities between records. For this different algorithms like classification, regression, clustering, neural networks etc. are used.

2. Examine the generated relationships.
3. These relationships will help in prediction and decision making process.
4. Predictions help to support decision-making processes and analysing the data.

Thus educational data mining is widely useful for different applications like analysis and visualization of data, providing feedback, recommendations for students, predicting student's performance student modelling, student grouping, detecting undesirable student behaviours and many more [3].

III. PREDICTION TECHNIQUES

This section provides overview of different data mining techniques that are used for prediction purpose.

A. Association Rule Learning

Association Rule Learning is a technique that finds different similar pattern from given database and finds strong rules from the database (i.e. which observation frequently occurs). Association Rule Learning helps into decision making process. Thus association rule learning is widely valuable in market basket analysis, web usage mining, medical diagnosis and many more. An association rule is an expression X → Y where X and Y are sets of different items. Association rules uses a user specified minimum support value and a user specified minimum confidence value.

1. First, minimum support is applied to find all frequent item sets in a database.
2. Second, these frequent item sets and the minimum confidence values are used to form rules. [1]

Advantages [4]:

- Association Rule Learning helps to find sequence of patterns.
- Association Rule Learning uses different methods like, acquisition, integration and integrity checks.

Limitations [4]:

- The technique does not show logical patterns with dependent variables.
- In association rule learning value of support and confidence is prerequisite.

Apriori Algorithm is an algorithm for association rule learning from the database. It finds the frequent data sets and generates association rules. Apriori algorithm uses breadth first search strategy to count support for item sets. Apriori algorithm is also known as level- wise algorithm and provides bottom up search. [5]

FP- growth algorithm in this algorithm FP means frequent pattern and the algorithm finds occurrence of the items and than builds tree for the items. In this algorithm, items are arranged in such a way that item at the top must contain highest frequent from its lower items. The algorithm uses threshold value to build the tree and generates the association rules. [6]

B. Classification

Classification is a technique that finds the rules that divides data into different groups. Classification process makes set of similar observation from the large database. An algorithm that implements classification is known as classifier, which is some mathematical function that implements the classification algorithm. The main task of this technique is to recognize the similar observation from large dataset and arrange them into a set. Classification techniques help to find different patterns among large dataset.

Advantages [4][7]:

- Efficiency is good.
- Handles the noisy data.
- Well suited for multimodal classes
- Requires short computational times

Limitations [4][7]:

- Sensitive to local structure of data
- Requires large memory.

Naïve bayes classifier combines approach of Baysian Classification and decision tree learning. Once a tree is grown this classifier is generated for each leaf using the data of that leaf and assigns class label to that leaf. Naive Bayes tree achives noticeable accuracy than either a naïve bayse classifier or decision tree learner. Advantages of naïve bayes classifier is that its easy computation.[8]

ADTree is Alternating Decision Tree. An alternating decision tree (ADTree) is a machine learning method for classification which generalizes decision trees. An alternating decision tree contains two nodes. In which a decision node denotes predicate condition. Prediction nodes denote single number. ADTree provides high accuracy compare to other techniques. [8]

Random Forest algorithm has good accuracy. It runs efficiently on large databases. It is able to handle thousands of input variables without variable deletion. It has an effective method for handling missing data and maintains accuracy when large parts of the data are

National Conference on Emerging Trends in Computer, Electrical & Electronics (ETCEE-2015)
International Journal of Advance Engineering and Research Development (IJAERD)
e-ISSN: 2348 - 4470 , print-ISSN:2348-6406,Impact Factor:3.134

missing. It provides a method for detecting variable interactions. [9]

K nearest neighbor is the one of the oldest technique of prediction**.** The method is that in order to predict what a prediction value is in one record look for records with similar predictor values in the past database and use the prediction value from the record that it "nearest" to the unclassified record.

C. Cluster Analysis

Cluster analysis or clustering is a technique for grouping the objects in a same group which finds similar characteristics. It is a method of grouping data. Cluster analysis is widely used in many fields like machine learning, pattern recognition, bioinformatics etc. Clustering is a technique like segmentation, which groups similar observations. Using association rule learning different patterns are generated from observations and than clustering technique groups them according to that pattern. In this approach same type data are store in same groups and these groups are known as clusters. Different techniques used in cluster analysis are artificial neural network and nearest neighbor search. Clustering helps in sequence analysis, grouping of items. The objectives of clustering are as follows:
- To discover natural groupings
- To produce hypothesis from data
- To find trustworthy organization of data.

Advantages [4]:

- Clustering helps end user by providing a high level view of what is on going on in the database.
- Very efficient technique.

Limitations[4]:

- For some clustering technique, it is fact that once a merge or split is committed it cannot be undone or refined.

K means clustering partitions n observations into *k* clusters in which each observation belongs to the cluster with the nearest mean. In centroid-based clustering, clusters are denoted by a central vector, which may not essentially be a member of the data set. In k means clustering no. of clusters are fixed and pre specified. K means clustering is one of machine learning technique. The limitation of this technique is the clusters are of smaller size leads to cut borders between clusters and we have to decide no. of clusters in advance.

Hierarchical clustering, in this technique the clusters are defined by the data and not as in k means clustering that no. of clusters are predetermined by user. And that the number of clusters can be increased or decreased by simple moving up and down the hierarchy. Objects are more related to their near objects than far

objects. These algorithms connect "objects" to form "clusters" based on their distance It is also known as connectivity based clustering technique. These technique will produce a hierachy of clusters. [11]

Density-based clustering in this technique, clusters are defined as areas of higher density than the remainder of the data set. [12]

Distribution-based clustering the clustering model is based on distribution models. A excellent property of this approach is that the way data sets are generated: by sampling random objects from a distribution. Distribution-based clustering produces complex models for clusters.

D. Decision Tree

A decision tree is a model that is use for prediction and can be viewed as a tree. Each branch of a tree denotes the condition and leaves denotes the result if the condition is satisfied. Each branch has been reported with a question or a condition and which have two or more than two answers. Each answer may leads toward another question or condition. Decision tree classifies the data according to the condition without losing any data and helps in decision making. Decision tree can work as a decision support tool. A decision tree is a flowchart-like arrangement in which each internal node represents a "test" on an attribute. The paths from root to leaf denotes classification rules. A decision tree is a classification method that generates a tree and a set of rules [5].

Advantages of decision tree [11]:

- Decision trees are simple to understand and interpret.
- **Requires little data preparation.** Other methods require data normalization, blank value removal etc.
- **Can handle both numerical and categorical data.** Other techniques like neural network only handle numerical data.
- **Robust.** Performs well even if record contains missing values.
- **Performs well with huge datasets.** Large amounts of data can be analysed in reasonable time.

Limitations of decision tree [11]:

- Difficulty in calculation, if many values are uncertain and/or if many results are linked.
- Decision-tree learners may create complex trees that do not generalize well from the training data. Techniques such as pruning are helpful to avoid this problem.

ID3 refers to Iterative Dichotomizer 3 is the algorithm for generating the decision trees from data. In ID3, each node corresponds to splitting attribute and each arc is a possible value of that attribute. At each node, splitting

attribute is the most informative attribute. Entropy is used to measure how informative is node. This algorithm uses information gain to determine splitting node. The attribute with greatest information gain is taken as the splitting attribute, and that node splits the data. [5]

C4.5 is extension of ID3 algorithm that helps to prune decision tree and generate rules. C4.5 produces decision trees with variable branches per node.[5]

CART stands for Classification and Regression Trees and is a data exploration and prediction algorithm. CART builds a binary decision tree by splitting records at each node, according to a function. CART use gini index for determining the best split.

CHAID is Chi-Square Automatic Interaction Detector. CHAID is similar to CART in that it builds a decision tree but it differs in the way that it chooses its splits. Instead of the entropy or Gini metrics for choosing optimal splits the technique depends on the chi square test[10]. CHAID tries to stop growing the tree before over fitting occurs. [5]

E. Neural Networks

Neural network are also used for detecting different patterns, and making predictions. A set of input is given to neural network and is used to predict one or more outputs. Neural network gives numerical output. Neural networks are widely useful in fraud detection, customer response prediction, image understanding and many more. Neural network also help in outlier analysis or in clustering. [5]

Advantages [10]

- Good learning ability.
- Speed is good.

Limitations [10]

- Handles only numeric data. So we have to translate each data into numeric form.
- Due to learning, problem of local optimal may occur.

Feed forward neural network is most simple type of artificial neural network. In this network the data moves in only one direction i.e. forwarding direction. From the input nodes data passes through the hidden nodes (if any) and to the output nodes. Network does not contain any cycles or loops.

Radial basis function (RBF) network have only one hidden layer. RBF networks have the advantage of not suffering from local minima in the same way as Multi-Layer Perceptions.

Bayesian network is a high-level representation of a probability distribution over a set of variables It uses probability values for generating the result. Bayesian networks explicitly represent the dependence and independence of various attributes. In a Bayesian network each node represents features from problem domain, and the arcs between the nodes represent the direct dependencies between the equivalent variables. In addition, each node contains a table of conditional probabilities. [13]

F. Support Vector Machine

Support vector machines are a supervised learning technique that uses different learning algorithms for analyzing data and finding patterns. This technique is widely used in classification and regression analysis. Support vector machines technique is used in medical diagnosis and can also recognize hand written characters. Support vector machines is based on statistical learning theory and helps in risk minimizing and minimize the classification error. [5]

Advantages of Support Vector Machine Techniques [14]:

- Categorize text and hyper text. Application reduces the need for labelled training instances.
- Helps in image classification.
- Useful in recognizing hand-written characters.

Limitations of Support Vector Machine Techniques [15]:

- The speed of this technique law for both training and testing.
- In support vector machine there is a possibility of presence of discrete data.
- Support vector machine have high algorithmic complexity.
- Requires large memory.

G. Regression Analysis

Regression Analysis finds similar observations from different patterns from the set of observations. It analyse relation between two different variables and finds that how the value of dependent variable changes with the change in value of independent variable. It checks the condition for the independent variable. This technique is used in prediction. Regression analysis uses a function for the independent variable that is known as regression function. Regression analysis is a Type of structured data analysis.

Advantages [16][17]:

- Help to correct errors
- Helpful in prediction
- Supports in decision making

- Indicates relationship between dependent variable and independent variables

Limitations[18]:

- Parameter instability.
- Public distribution of relationships.

H. Sequential Pattern Mining

A sequence is an ordered list of transactions. Sequential pattern mining as the name implies finds different patterns from the datasets. The technique recognize the repeatedly occur patterns. Sequential pattern mining helps to identify current trends or regular occurrence of similar events. Two types of sequential pattern mining that is string mining and item set mining. In item set mining the algorithm recognizes set of observations that frequently occurs using different rules. Sequential pattern mining is widely used in business application. Sequential pattern mining widely uses apriori algorithm. Sequential pattern mining discovers subsequences that are common to more than minimum support sequences in a database, where minimum support is set by the user.

Advantages [10]

- Helps for predictions which are based on sequence of events.
- Simple learning mechanism

Limitations [10]

- Sometimes requires large computation time.
- Limited prediction capabilities

The following table shows the comparison of all data mining techniques

Table 1 : Comparison

Data Mining Technique	Features	Advantages	Limitations
Association Rule Learning	• Finds similar patterns from data and produces rules. • Generates association relationship from data. • Helps in decision making process. • Uses min support value and min confidence support value.	• Helps to find sequential patterns • Uses acquisition, integration and integrity checks methods..	• Do not show reasonable patterns with dependent variables. • Value of support and confidence is prerequisite.
Classification	• Technique that finds the rules that partitions data into different groups. • Identify the similar observation from large dataset and arrange them into a set.	• Efficiency is good. • Handles the noisy data. • Well suited for multimodal classes • Requires short computational times.	• Sensitive to local structure of data • Requires large memory.

Cluster Analysis	• Technique for assemblage the objects in a group which finds similar characteristics.
	• The objectives of the technique are as follows:
	  o To discover natural groupings
	  o To produce hypothesis from data
	  o To find reliable organization of data.
		• Provides end user high level view of what is going in database.
		• Very efficient technique
			• Once a merge or split is committed it cannot be undone or refined.

Decision Tree	• A model that is use for prediction and can be viewed as a tree. • Flow chart like structure • Each branch of a tree represents the condition and leaves represent the result if the condition is satisfied. • Decision tree divides the data according to the condition and helps in decision making.	• Simple to understand and interpret. • Able to handle both numerical and categorical data. • Robust. • Performs well with large datasets.	• Sometimes calculations are complex.. • Sometimes suffers from over fitting problem.

National Conference on Emerging Trends in Computer, Electrical & Electronics (ETCEE-2015)
International Journal of Advance Engineering and Research Development (IJAERD)
e-ISSN: 2348 - 4470 , print-ISSN:2348-6406,Impact Factor:3.134

	• Decision tree can work as a decision support tool.		
Neural Network	• Used for detecting different patterns, and making predictions. • Generates numerical output. • Widely useful in fraud detection, customer response prediction, image understanding and many more.	• Good learning ability. • Speed is good.	• Handles only numeric data. So we have to translate each data into numeric form. • Due to learning, problem of local optimal may occur.
Support Vector Machine	• Supervised Learning Technique • Helps in risk minimizing and minimize the classification error. • Part of linear classifiers and can be extension of perception technique.	• Produces very accurate classifiers • Less over fitting and handles noise • Memory intensive • Useful in recognizing hand-written characters.	• Speed is law for both training and testing. • Possibility of presence of discrete data. • High algorithmic complexity.
Regression Analysis	• Analyse relation between two different variables. • Finds that how the value of dependent variable changes with the change in value of independent variable. • Uses a function for the independent variable that is known as regression function.	• Helps in correcting errors • Helpful in prediction • Supports in decision making • Indicates relationship between dependent variable and independent variables.	• Parameter instability • Public distribution of relationships
Sequential Pattern Mining	• Finds different patterns from database. • Finds current trends or regular occurrence of similar events • Uses apriori algorithm	• Helps for predictions based on sequence of events • Simple learning mechanism	• Sometimes requires large computation time. • Limited prediction capabilities.

IV. APPLICATIONS OF PREDICTION TECHNIQUES

Educational data mining helps to find out different patterns fr4om student database and make predictions that characterize student's behaviour. Educational data mining provides support for educational institutes for improving student's performance. [19].

Different goals of educational data mining are as follow [20].

1. Predicting and analyzing student's learning behavior: This can be achieved by creating student groups according to their learning characteristics including their past data, class performance and many more.
2. Improving domain models: Using educational data mining discovery of new improvement in existing model of institute is done. Analyzing the learning content to connect students and to support student's learning routine. And also helps to increase their performance.
3. Helps to provide recommendations for students.
4. Helps to predict student's performance.
5. Generates groups on the basis of student's learning habit.

6. Using this technique, institute can detect undesirable student behavior.

V. EXISTING APPROACHES FOR PREDICTING TECHNIQUES ON STUDENT DATA

[21] presents K means clustering algorithm by using data mining tool WEKA that monitors the progress of student's academic performance in higher education. They use cluster analysis to group students as per their characteristics using their class internal marks, GPA, exam mark, assignment and lab work. They evaluate and predict student's learning activity, which may improve student's academic performance and reduce the failing ratio. Using K means clustering and WEKA interface they classified students according to their performance. They also present student's detailed analysis of student performance.

[22] apply the decision tree in educational data mining. Decision trees algorithm are applied on student's past performance. They used ID3 method of decision tree generation to study student's performance in end General

appreciation. They used different attributes like high school degree, mid term marks, lab test grade, seminar performance, assignment, attendance, homework etc. from student database to analyze student's performance. Here the study helps to evaluate student's performance and helps to improve their study.

[23] provides framework for mining educational data. Using classification methods they generate different rules for evaluation of student data. The main object of their research is to apply data mining technique s in the field of higher education and to identify which data mining application is suited for what kind of application. They use predictive classification to enhance the quality of higher education system and to evaluate student data to study that main attribute that may affect the student enrolment factors to plan for institute resources by applying different data mining techniques like association analysis, classification and clustering.

[24] performs analysis and prediction over student's record for helping students to make subject choice, to improve their result for warning students about barriers in their performance using faculty feedback analysis and the knowledge gained factor of students based on their past mark records. They use R statistical tool for the prediction. They build predictive model to forecast student performance and the model helps to find out teaching pattern, judgment of performance and helps to motivate students. They suggested that the result can be enhanced by adding some social factors to predict student's performance and this will help to get more accurate result.

[25] suggested a approach using Rapid minor tool to suggest carrier options to high school students and to predict their potentially violent

[26] used a hybrid approach based on Association rule mining and Rule induction to get accurate result in less processing time. They used association rule mining for generating the item sets. They can also reduce error rate and processing time.

[27] used Bayesian Network to forecast student's data. They used past attendance record to forecast student's performance. They used information gain to reduce the number of variables. They used both grade data and time recording data retrieved from smart cards and provided a bayesian network to predict performance of students.

[28] suggested a data mining approach to find out risky students and evaluate them to reduce drop out rate in higher education and to analyze student retention..

[29] developed a model to predict secondary education placement test course, using different techniques like decision tree algorithm, support vector machines, artificial neural network, and multinomial logistic regression and

find out that the accuracy of decision tree algorithm is higher compare to other techniques. They used k-Fold cross validation to minimize the bias associated with random sampling.

[30] suggested a data mining approach for analysing the risk of undergraduate student retention and aim to prevent student retention using different data mining techniques like logistic regression, rules induction and compare their efficiency.

VI.   CONCLUSION

This paper provides that how different data mining techniques are useful in predicting student's performance in educational data mining. Thus different data mining techniques helps to analyze student's learning behavior predicting their result, grouping them, and for finding out similar patterns. Thus it can be conclude that educational data mining helps for improving student's performance, for managing the student database..

REFERENCES

[1] Jiawei Han & Micheline Kamber: "Data Mining Concepts and Techniques" 2E. Morgan Kaufmann Publishers is imprint from Elsevier, San Francisco, CA 94111

[2] R. Baker (2010) Data Mining for Education. In McGaw, B., Peterson, P., Baker, E. (Eds.) International Encyclopedia of Education (3rd edition), vol. 7, pp. 112-118. Oxford, UK: Elsevier.

[3] C. Romero, S. Ventura. Educational Data Mining: A Review of the State-of-the-Art. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 40(6), 601-618, 2010.

[4] Moawia Elfaki Yahia, Mutada El-mukashfi El-taher "A new approach for evaluation of data mining techniques " ,IJCSI International Journal for Computer Science Issues vol 7, Issue 5, September 2012

[5] Data Mining Techniques : Arun K. Pujari, Universities Press 10 E.

[6] Witten, Frank, Hall: Data mining practical machine learning tools and techniques, 3rd edition

[7] S. Archna, Dr. K. Elangovan "Survey of Classification Techniques in Data Mining", International Journal of Computer Science and Mobile Applications vol 2, Issue 2,, February 2014, p.g. 65-71

[8] Yongheng Zhao and Yanxia Zhang : Comparison of decision tree methods for finding active objects, National Astronomical observations, CAS, 20A Datun Road, Chaoyang District Beijing 100012 China, arXiv:0708.4274v1 [astro-ph]

[9] Random Forest Algorithm : Data Science Control

[10] Smart Home Technologies : Data Mining & Prediction

[11] An Overview of Data Mining Techniques Excerpted from the book Building Data Mining Applications for CRM by Alex Berson, Stephen Smith, and Kurt Thearling

[12] Kriegel, Hans-Peter; Kröger, Peer; Sander, Jörg; Zimek, Arthur (2011). "Density-based Clustering". *WIREs Data Mining and Knowledge Discovery* 1 (3): 231–240.doi:10.1002/widm.30.

[13] Advantages of Bayesian Networks in Data Mining and Knowledge Discovery By Petri Myllymäki, Ph.D., Academy Research Fellow, Complex Systems Computation Group, Helsinki Institute for Information Technology

[14] Meyer, D.; Leisch, F.; Hornik, K. (2003). "The support vector machine under test". *Neurocomputing* 55: 169. doi:10.1016/S0925-2312(03)00431-4. Edit

[15] www.svms.org/disadvantages.html

[16] smallbusiness.chron.com/advantages-regression-analysis-forecasting-61800.html

National Conference on Emerging Trends in Computer, Electrical & Electronics (ETCEE-2015)
International Journal of Advance Engineering and Research Development (IJAERD)
e-ISSN: 2348 - 4470 , print-ISSN:2348-6406,Impact Factor:3.134

[17]   A concise guide to market research. Chapter -7 Regression Analysis

[18]   Investopedia- Quantitative methods Regression Analysis

[19]   Review Educational data mining: A survey and a data mining-based analysis of recent works Alejandro Pena-Ayala WOLNM & ESIME Zacatenco, Instituto Politécnico Nacional, U. Profesional Adolfo López Mateos, Edificio Z-4, 2do piso, cubiculo 6, Miguel Othón de Mendizábal S/N, La Escalera, Gustavo A. Madero, D.F., C.P. 07320, Mexico

[20]   Baker, R.S.; Yacef, K (2009). "The state of educational data mining in 2009: A review and future visions". JEDM-Journal of Educational Data Mining 1 (1): 201

[21]   Mahesh Singh, Anita Rani, and Ritu Sharma**: "**An optimized approach for student's academic performance by k means clustering algorithm using WEKA interface." International Journal of Advanced Computational Engineering and Networking vol-2, Issue-7, July 2014.

[22]   D. Bhulakshmi, S. Arundhati amd Dr. Jagdish : "Data Mining : A prediction for student's performance using decision tree ID3 method". International Journal of Science & Engineering Research vol 5, issue 7, July 2014

[23]   P. Veeramuthu and Dr. R. Periasamy : "Application of higher education system for predicting student using data mining techniques". International Journal of Innovative Research in Advanced Engineering vol 1, issue 5, June 2014

[24]   Ananya Chandraker and G. Vadivu : "Analytics and prediction over student's record". International Journal of Advances in Science Engineering and Technology vol -2, Issue-2, April 2014.

[25]   Elakia, Gayathri, Asrthi, Naren J : "Application of Data Mining in Educational Database for Predicting Behavioural Patterns of the Students" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4649-4652

[26]   Er. Amarjeet Kaur, Er. Kumar Saurabh and Er. Gurpreet Singh : "A Combined Approach of Data Mining Algorithms Based on Association Rule Mining and Rule Induction." International Journal of Soft Computing and Engineering. ISSN : 2231-2307, Vol-3, Issue -5, November 2013

[27]   Hirotaka Itoh, Yuma Itoh, Kenji Funahashi, Daisuke Yamamoto, Shoichi Saito, Ichi Takumi and Hiroshi Matsuo : "Forecasting Student's Future Academic Records Using Past Attendance Recording Data and Grade Data" . ELSEVIER 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems- KES2013Procedia Computer Science 22 (2013) 921-927

[28]   [Ghazi Mohammed Zafaruddin and Heena Jadhav : "Data Mining Approach to Student Retention" International Monthly Refereed Journal of Research In Management & Technology Volume II, March'13

[29]   Baha Sen, Emine Ucar, Dursun Delen : " Predicting and analyzing secondary education placement test scores : A date mining approach" . ELSEVIER Expert System with Applications 39 (2012) 9468-9476.

[30]   Hadautho Roberto Barros da Silva and Paulo Jorge Leitão Adeodato : "A Data Mining Approach for Preventing Undergraduate Students Retention" WCCI 2012 IEEE World Congress on Computational Intelligence June, 10-15, 2012 - Brisbane, Australia