ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

Analysis of Influencing Factors in Predicting Students Performance Using MLP - A Comparative Study

Jai Ruby[1], Dr. K. David[2]

Research Scholar, Research & Development Centre, Bharathiar University, Tamilnadu, India[1]

Associate Professor, Dept. of Computer Science and Engg, Rover Engg College, Perambalur, Tamilnadu, India[2]

ABSTRACT: In recent years, Educational Data Mining has put on a massive appreciation within the research realm and it has become a vital need for the academic institutions to improve the quality of education. The quality of education is measured by the academic performance of students and the results produced. In higher education institutions a substantial amount of knowledge is hidden and need to be extracted using Knowledge Discovery process. Data mining helps to extract the knowledge from available dataset and should be created as knowledge intelligence for the benefit of the institution. Many factors influence the academic performance of the students. The study model is mainly focused on analyzing the prediction accuracy of the academic performance of the students using only influencing factors by Multi Layer Perceptron algorithm and to compare it with the prediction accuracy of the academic performance of the students using a dataset that comprises of all academic, personal and economic factors of the students using Multi Layer Perceptron algorithm.

KEY WORDS: Educational Data Mining, Academic Performance, Higher Education, Prediction, Classification, Multi Layer Perceptron

I. INTRODUCTION

Educational data has become a vital resource in this modern era, contributing much to the welfare of the society. Educational institutions are becoming more competitive because of the number of institutions growing rapidly. To stay afloat, these institutions are focusing more on improving various aspects and one important factor among them is quality learning. For providing quality education and to face new challenges, the institutions need to know about their potentials which are explicitly seen and which are hidden. The truths behind today's educational institutions are a substantial amount of knowledge is hidden. To be competitive, the institutions should identify their own potentials hidden and implement a technique to bring it out. In recent years, Educational Data Mining has put on a mammoth recognition within the research realm as it has become a vital need for the academic institutions to improve the quality of education.

The higher education institutions has potential knowledge such as academic performance of students, administrative accounts, potential knowledge of the faculty, demographic details of the students and many other information in a hidden form. The technique behind the extraction of the hidden knowledge is Knowledge Discovery process. Recently Data mining is widely used on educational dataset. Educational Data mining (EDM) has become a very useful research area [1]. Data mining helps to extract the knowledge from available dataset and should be created as knowledge intelligence for the benefit of the institution. Higher education does categorize the students by their academic performance. Many factors influence the academic performance of the student. The study model [2] is mainly focused on exploring various indicators that have an effect on the academic performance of the students. The extracted information that describes student performance can be stored as intelligent knowledge for decision making to improve the quality of education in institutions. The knowledge stored is used for predicting the students performance in advance.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

EDM can be considered as one of the learning sciences, as well as an area of data mining [3]. The researcher applied the educational data mining concerns with developing methods for discovering knowledge from data that come from educational domain and used to analyze learning behavior. Some of the benefits of data mining in education sector are identifying students' preferences towards course choices, their selection of specialization and predicting students' knowledge, grades, and final results [4]. Institutions of Higher Learning (IHL) are similar to knowledge businesses, in that both are involved in knowledge creation, dissemination, and learning [5]. However, people in business world are concerned with the profit they could gain by exploiting knowledge through the implementation of KMS whereas IHL consider that KMS could improve the quality of service deliveries and sustained competitive advantages in the academic world [6].

This paper makes a novel attempt to look into the higher educational domain of data mining focused on analyzing the prediction accuracy of the academic performance of the students using only influencing factors by Multi Layer Perceptron algorithm and to compare it with the prediction accuracy of the academic performance of the students using a dataset that comprises of all academic, personal and economic factors of the students using Multi Layer Perceptron algorithm. Section 2 gives the overview of data mining techniques available to extract the hidden information. Section 3 provides the general account of the data under study and details the neural network and various functions related to it. Section 4 predicts results by multi layer perceptron algorithm using all possible factors of student dataset and by the identified high influencing factors. Conclusion and a discussion on future work are in the final section.

A. Related Work

Zlatko J. Kovacic and John Steven Green, have predicted student's academic performance using various attributes like gender, parent education, economic background etc.[7]. M.N. Quadri & N.V. Kalyankar explained that the previous academic result plays a key task to predict the students who are a threat to be unsuccessful in the exam[8]. Bhardwaj & Pal performed a study on the student performance among 300 students. By means of Bayesian classification method on 17 attributes, it is noted that the influencing factors like grade in senior secondary exam, medium of teaching, living location, mother's qualification, other habit, income and family status plays a vital role in the student academic performance[9]. The researchers performed a study on the student performance on 600 students[10]. The attributes like category, language and background qualification were used to predict student performance. Hijazid and Naqvi, performed a study on the student performance by selecting a sample of 300 students (225 males, 75 females). The study reveals that "Student's attendance, hours spent in study, family income and mother's education are significantly related with student performance" using linear regression[11].

The researchers used data mining classification techniques to enhance the quality of the higher educational system by evaluating students' data that may affect the students' performance in courses[12]. They used three different classification methods ID3, C4.5 and the NaiveBayes. The results indicated that the decision tree model had better prediction accuracy than the other models. Z. J. Kovacic in 2010 presented a study on educational data mining to identify up to what extent the enrolment data can be used to predict student's success [13]. The algorithms CHAID and CART were used. The researchers applied a Classification Technique in Data Mining to enhance the student's performance by extracting the discovery of knowledge from the end semester mark [14].

Mohammed M. Abu Tair and Alaa M. El-Halees applied the data mining for discovering knowledge from data that come from educational environment [15]. Student's data has been collected from the college of Science and Technology for a period of 15 years [1993-2007]. The collected data was preprocessed and data mining techniques are applied to improve graduate students' performance, and overcome the problem of low grades of graduate students. Muslihah W. et.al, have compared Artificial Neural Network and the combination of clustering and decision tree classification techniques for predicting and classifying student's academic performance. Students' data were collected from the data of the National Defense University of Malaysia (NDUM)[16]. H. W. Ian and F. Eibe gave a case study that used educational data mining to identify behavior of failing students who are at risk before the final exam [17].

Nguyen N et.al, compared the accuracy of decision tree and Bayesian network algorithms for predicting the academic performance of students of Under Graduate and Post Graduate students. The decision tree classifier provided better accuracy than the Bayesian network classifier[18]. Ramaswami M. and Bhaskaran R, have constructed a predictive

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

model using 772 students' records with 7-class response variables by using highly influencing predictive variables obtained through feature selection [19]. The accuracy of the present model was compared with other models and it has been found to be satisfactory. The researchers have conducted a study on a data set of size 50 MCA students for mining educational data to analyze students' performance. Decision tree method was used for classification and to predict the performance of the students [20]. Different measures that are not taken into consideration were economic background, technology exposure etc. Bengio Y. et.al, discussed that neural networks are suitable in data-rich environments and are typically used for extracting embedded knowledge in the form of rules, quantitative evaluation of these rules, clustering, self-organization, classification and regression[21]. Neural networks have an advantage, over other types of machine learning algorithms, for scaling.

II. DATA MINING

Data mining also termed as Knowledge Discovery in Databases (KDD) refers to extracting or "mining" knowledge from large amount of data. Han & M. Kamber in 2001 referred that Knowledge Discovery process involve various steps like Data cleaning, transformation, data mining, pattern evaluation in extracting knowledge from data [22]. Knowledge Discovery is involved in a multitude of tasks such as association, clustering, classification, prediction, etc. Classification and prediction are functions which are used to create models that are constructed by analyzing data and then used for assessing other data. Clustering is a way of identifying similar classes of objects. Association is mainly used to relate frequent item set among large data sets.

Two steps are involved in classification. In the first step, a model that describes a predetermined set of classes or concepts is made by examining a set of training dataset. The learning is known as supervised learning as the class labels of all the records of the dataset are known. The models are usually in the form of classification rules or decision tree. In the second step, the model is put to test using a different data set that is used to estimate the predictive accuracy of the model. Various methods like holdout, random sub sampling, k-fold cross validation, stratified cross validation, bootstrapping are used to estimate the accuracy of the model. If the accuracy of the model is considered acceptable, the model can be used to classify the dataset for which the class label is not known in advance [22].

Basic techniques for classification are decision tree induction, Bayesian classification and neural networks. Other approaches like genetic algorithms, rough sets, fuzzy logic, case based reasoning can also be used for classification. Decision Tree classifier is a powerful and popular classification and prediction technique (Chaudhuri, 1998). Some of the decision tree classifiers are J48, NBTree, ID3, CART, REPTree, Simplecart, BFTree and others. A Decision Tree is a flow-chart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test and leaf nodes represent cases or class distributions [22].

Bayesian classifiers are statistical classifier. The Naive Bayes algorithm is a simple probabilistic classifier that calculates a set of probabilities by counting the frequency and combinations of values in a given data set. In an NBTree, a local naive Bayes is deployed on each leaf of a traditional decision tree, and an instance is classified using the local naive Bayes on the leaf into which it falls. After a tree is grown, a naive Bayes is constructed for each leaf using the data associated with that leaf. An NBTree classifies an example by sorting it to a leaf and applying the naive Bayes in that leaf to assign a class label to it.

III. ARTIFICIAL NEURAL NETWORK - ANN

A neural network is a biologically simulated computational model. Neural networks have been used in a large number of applications and have proven to be effective in performing complex functions in a variety of fields. These include pattern recognition, classification, vision, control systems, and prediction[23]. A neural network has two primary components called processing elements and the connections between them. The processing elements are said as neurons. The connections between the neurons are termed as links. Each link has its own weight parameter associated with it. The ANN consists of an input layer, an output layer and at least one layer of nonlinear processing elements, known as the hidden layer. The input values to the network are received from the input layer through the hidden layer to the output layer. The processing of input values is done within the individual nodes of the input layer and then the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

output values are forwarded to the nodes in the hidden layer. The values obtained as inputs by the hidden layer nodes are processed within them and forwarded to either the nodes of the next hidden layer or nodes of the output layer.

A neural network is a network that has the ability to learn from its background and improve its performance through learning. The learning process can be classified as supervised and unsupervised. Supervised learning is training the neural network with a training set and the network parameters are adjusted iteratively so that the network gets trained to produce the desired output for any given input. In unsupervised learning there is no training set to supervise the learning process. Here the network is required to learn by itself and should produce desired output for any given input data and creates new classes automatically. Single Layer Perceptron, Multi Layer Perceptron(MLP), ADALINE, Support Vector Machine (SVM), K-Nearest Neighbour (KNN) are some of the well known supervised learning algorithms and K-means, Self Organising Map (SOM), Adaptive Resonance Theory (ART) are some of the unsupervised learning algorithms.

MLP – Multi Layer Perceptron algorithm is one of the most widely used and common supervised neural network method. Multilayer Perceptron is a feed forward artificial neural network model trained with the standard back propagation algorithm that maps sets of input data onto a collection of acceptable output. An MLP consists of multiple layers of nodes in a directed graph, with every layer totally connected to the consequent one. These are supervised networks so they require a desired response to be trained. They learn how to transform input data into a desired response, so they are widely used for pattern classification and prediction.

IV. METHODOLOGY

The dataset used for this study for performance analysis was taken from PG Computer Application course offered by an Arts and Science College between 2007 and 2012. The data of 165 students were collected. Student personal and academic details along with their attendance were collected from the student information system. The collected information was integrated into a distinct table. Student dataset contains various attributes like Theory Scores, Laboratory scores, Medium of study, UG course, Family Income, Parental Education, First Generation Learner, Stay, Extracurricular activities etc. Among the different attributes initially present using feature selection techniques like chi square, info gain, gain ratio, correlation and regression it was found that the high impact attributes that contribute for the performance of the students are Theory, Medium of Study, Previous Course studied, UG Percentage, Stay, Extra Curricular Activities and Family Income [2]. The influencing attributes are selected and are used to classify and predict the student performance using weka data mining tool.

Weka (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka is a free software available under the GNU General Public License. The Weka workbench contains a collection of visualization tools and algorithms for data analysis and predictive modelling, together with graphical user interfaces for easy access to this functionality. Weka tool contains many packages which include Filters, Classifiers, Clusters, Associations, and Attribute Selection. The Visualization tool in weka allows datasets and the predictions of Classifiers in a pictorial form. Weka is a collection of machine learning algorithms for solving real-world data mining problems. It is written in Java and runs on almost any platform. The algorithms can either be applied directly to a dataset or called from own Java code. In Weka datasets should be formatted to the ARFF format. The initial dataset of 165 records was split up into two sets. Two thirds of the data are allocated to the training set and the remaining one third is allocated to the test set. The training set help in building the model and it is used for classification. For estimating classifier accuracy k-fold cross-validation is used. Training and testing is performed k-times. The accuracy estimate is the overall number of correct classifications from the k iterations divided by the total number of samples in the initial data[22].

The classify panel in the weka tool facilitates to apply classification algorithms and to estimate the accuracy of the predictive model. Among the different classifiers of ID3, J48, NBTree, RepTree, Multi Layer Perceptron (MLP), SimpleCart and Decision table the study model [24] show that MLP learning algorithm proved to be the best. The aim of this study is to justify that the found out high impact attributes that contribute for the performance of the students using feature selection [2] are true and to justify that MLP classifier is the best among the other alternate classification

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

algorithms[24]. This study is carried out by comparing the predicted results of the students using MLP with all the 12 attributes family income, parent education, previous course studied, UG percentage, first generation learner, stay, urban, medium of study, attendance, theory marks, extracurricular activities, lab marks and the student performance predicted by the identified 7 high influence factors using the best learning algorithm MLP. The study model is focused on analyzing the prediction accuracy of the academic performance of the students using a dataset with only influencing factors by Multi Layer Perceptron algorithm and to compare it with the prediction accuracy of the academic performance of the students using a dataset that comprises of all academic, personal and economic factors of the students.

V. RESULTS AND DISCUSSION

The student data set of 165 records with all the 12 attributes that includes the personal, academic and economic (family income, parent education, previous course studied, UG percentage, first generation learner, stay, urban, medium of study, attendance, theory marks, extra curricular activities, lab marks) was split into two sets consisting of two-third as training set and one-third as testing set. The training set is used to build a model and the test set is used to estimate the accuracy of the classifier and if it is acceptable then it is used for the prediction of data for which the class label is unknown. MLP was the data mining classification algorithm chosen for the study via weka. The dataset set was divided into 5 sets (train) of distinct two-third records and 5 sets (test) of distinct one-third records. These sets were used in Run1 through Run5 respectively. In each run, 3 new set of data whose class labels are unknown was given for prediction. Since we use three different sets of new data for prediction, the average of the 3 results was considered for each run. Since we have 5 training data set and 5 test data set we get 5 results.

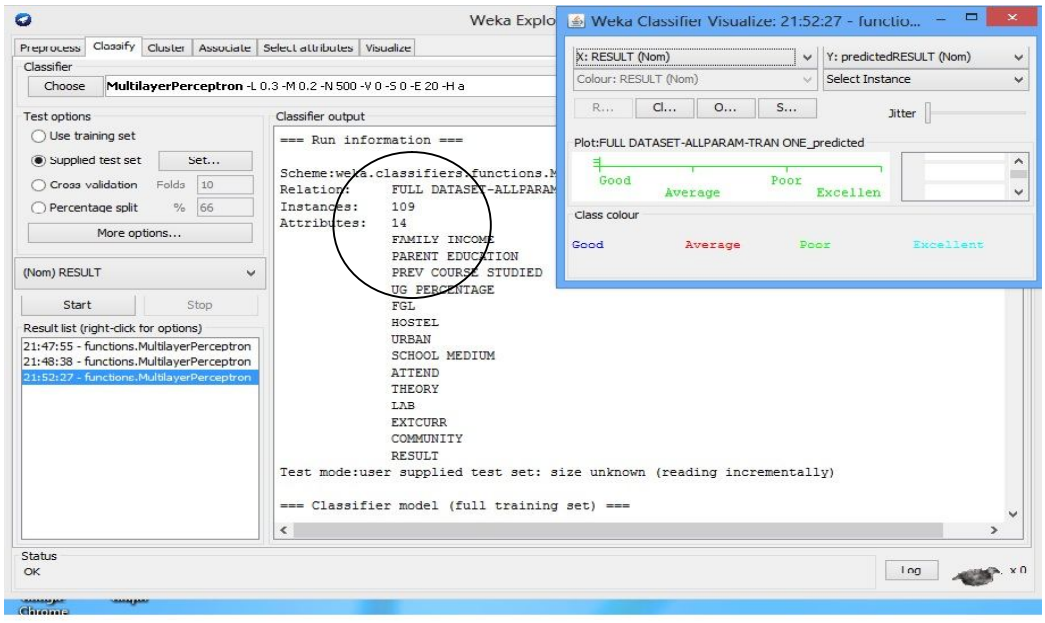

Fig.1 shows a sample of the prediction result of dataset with all 14 attributes using MLP via weka. The experiment was carried out using 5 different train sets of data with 5 different test sets using MLP classification algorithm. In each run, 3 new set of data whose class labels are unknown was given for prediction and the results are tabulated below in Table 1.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

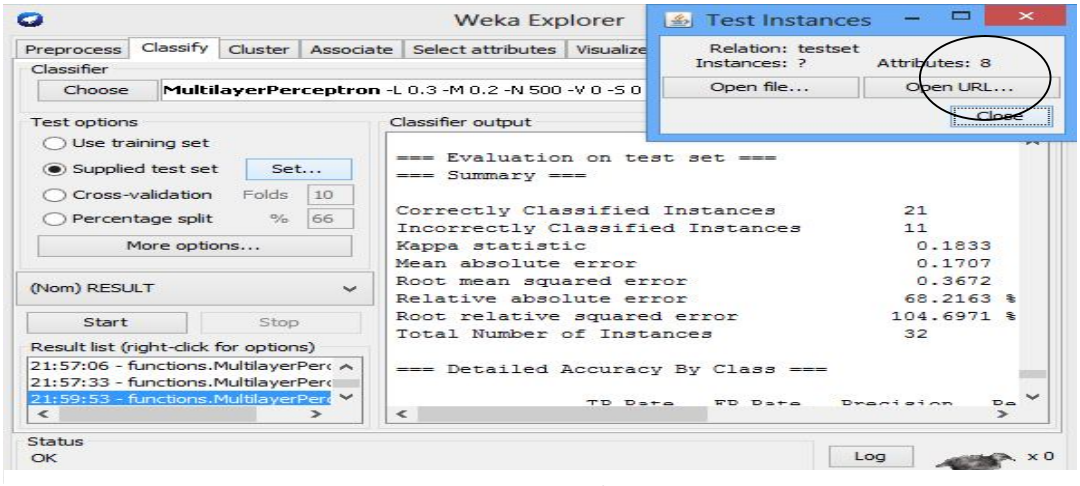

Fig. 2: Prediction using dataset with high influence attributes applying MLP

Fig.2 shows the prediction result of dataset with the high influencing attributes using MLP via weka. The experiment was repeated using the above said procedure for a dataset using only 7 attributes (family income, previous course studied, UG percentage, stay, medium of study, theory marks, extra curricular activity). The 8th attribute in Fig.2 represents the unknown 'Result' attribute that is to be predicted by the algorithm. The experiment was carried out using 5 different train sets of data with 5 different test sets using MLP classification algorithm. In each run, 3 new set of data whose class labels are unknown was given for prediction and the results are tabulated below in Table 1.

DATA SET	Run 1	Run 2	Run 3	Run 4	Run 5
All Attributes	81.6	32	57	44.3	83.6
High Influence Attributes	73.7	60	57.3	49	82.6

Table 1 : Prediction Accuracy of the MLP classification algorithm in percentage for 5 different dataset

Table-1 shows prediction accuracy of the MLP classification algorithm in percentage for a different datasets, ie with all attributes and only with the attributes having high influence. In each run, 3 new set of data whose class labels are unknown was given for prediction and the average of the 3 results was considered in terms of percentage. Table 2 shows the prediction percentage by computing the average of the 5 runs using all the attributes and only with high influence attributes.

Dataset	Prediction Percentage
All Attributes	**59.7**
High Influence Attributes	**64.5**

Table 2 : Average Prediction Percentage of 5 runs using dataset with different number of attributes

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

The results shows that dataset with high influence attributes show a better prediction percentage than dataset with all attributes. This shows that the attributes are really high influence attributes in the original dataset.

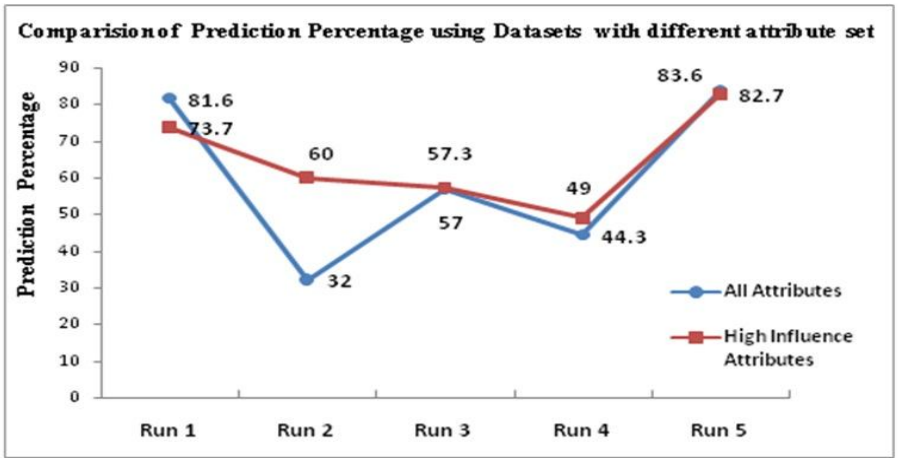

Fig. 3: Comparison of Prediction Accuracy using Datasets with different attributes using MLP classification Algorithm

Fig.3 shows the comparison of prediction accuracy using datasets with different set of attributes using MLP classification data mining algorithm. The results show that prediction percentage of dataset with high influence attributes behave alike whatever may be the training and test data. The variation of prediction percentage for different runs for the dataset with all attributes was around 52% and for the dataset with high influence attribute was around 33%. This shows that the attributes identified in the study are practically high influencing factors in predicting student performance.

VI. CONCLUSION

This model is mainly focused on analyzing the prediction accuracy of the academic performance of the students using only influencing factors by Multi Layer Perceptron algorithm and to compare it with the prediction accuracy of the academic performance of the students using a dataset that comprises of all academic, personal and economic factors of the students using Multi Layer Perceptron algorithm. The paper prove the attributes chosen from the original dataset are really high influence using MLP. This study paper helps the institution to know the academic status of the students in advance and can concentrate on weak students to improve their academic results. The study can be carried by combining two or more algorithms for better prediction or a new algorithm can be developed for better classification and prediction using the high influence attributes would be the future work.

REFERENCES

[1]. Baker R.S.J.D., & Yacef K, "The state of educational data mining in 2009:A review and future vision", Journal of Educational Data Mining, I, pg. 3-17,2009.
[2]. Jai Ruby & K. David, "A study model on the impact of various indicators in the performance of students in higher education", IJRET International Journal of Research in Engineering and Technology, Vol. 3, Issue 5, pp.750-755, May 2014.
[3]. Monika Goyal & Rajan Vohra, "Applications of Data Mining in Higher Education" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, pp.130-120, March 2012.
[4]. Mohd Maqsood Ali, "International Journal of Computer Science and Mobile Computing", Vol.2 Issue. 4, pg. 374-383, April- 2013.
[5]. Rowley, J.,"Is higher education ready for knowledge management?",International Journal of Educational Management, vol. 14(7), pp. 325–333, 2000.
[6]. Lubega, J. T., Omona, W., & Weide, T. V. D., "Knowledge management technologies and higher education processes_: approach to integration for performance improvement", International Journal of Computing and ICT Research, vol. 5(Special Issue), pp. 55–68, 2011.
[7]. Zlatko J. Kovacic & John Steven Green, "Predictive working tool for early identification of 'at risk' students", New Zealand, 2010.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 2, February 2015

[8]. Quadril.M.N & Dr. Kalyanka N.V.," Drop Out Feature of Student Data for Academic Performance Using Decision Tree", Global Journal of Computer Science and Technology Vol. 10 Issue 2 (Ver 1.0), April 2010.
[9]. Bharadwaj. B.K & S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.
[10]. Pandey U.K & S.Pal, "Data Mining A prediction of performer or underperformer Using classification", IJCSIT International Journal of Computer Science and Information Technology, Vol. 2(2),pp.686-690,ISSN:0975-9646,2011
[11]. Hijazi S. T., &. Naqvi R. S. M. M, "Factors affecting student's performance: A Case of Private Colleges", Bangladesh e-Journal of Sociology, Vol. 3, No. 1, 2006.
[12]. Al-Radaideh Q., Al-Shawakfa E., & AI-Najjar M., "Mining Student Data using Decision Trees", In Proceedings of the International Arab Conference on Information Technology (ACIT'2006), Yarmouk University, Jordan, 2006.
[13]. Kovacic Z. J., "Early prediction of student success: Mining student enrollment data", Proceedings of Informing Science & IT Education Conference 2010.
[14]. Shanmuga Priya. K, & Senthil Kumar A.V., "Improving the Student's Performance Using Educational Data Mining", 2013.
[15]. Mohammed M. Abu Tair & Alaa M. El-Halees, "Mining Educational Data to Improve Students' Performance: A Case Study", 2012
[16]. Muslihah W., Yuhanim Y., Norshahriah W., Mohd Rizal M., Nor Fatimah A., & Hoo Y. S., "Predicting NDUM Student's Academic Performance Using Data Mining Techniques", In Proceedings of the Second International Conference on Computer and Electrical Engineering, IEEE computer society, 2009.
[17]. Ian H. W. & Eibe F., "Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations," California: Morgan Kaufmann, 2005
[18]. Nguyen N., Paul J., & Peter H., "A Comparative Analysis of Techniques for Predicting Academic Performance", In Proceedings of the 37th ASEE/IEEE Frontiers in Education Conference. pp. 7-12, 2007.
[19]. Ramaswami M., & Bhaskaran R., "CHAID Based Performance Prediction Model in Educational Data Mining", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 1, No. 1, 2010.
[20]. Brijesh Kumar Baradwaj & Saurabh Pal, "Mining Educational Data to Analyze Students' Performance", IJACSA, Vol.2, No.6, 2011
[21]. Bengio Y., Buhmann J. M., Embrechts M., & Zurada J. M., "Introduction to the special issue on neural networks for data mining and knowledge discovery," IEEE Trans. Neural Networks, vol. 11, pp. 545-549, 2000.
[22]. Han. J & Kamber. M, "Data mining concepts and techniques", San Francisco, USA, Morgan Kaufmann,2001.
[23]. Fausett, L., "Fundamentals of Neural Networks", Prentice-Hall, Englewood Cliffs, NJ, 1994.
[24]. Jai Ruby & K. David, "Predicting The Performance Of Students In Higher Education Using Data Mining Classification Algorithms - A Case Study", IJRASET International Journal for Research in Applied Science & Engineering Technology, Vol. 2, Issue XI, ISSN No. 2321-9653, November 2014.

BIOGRAPHY

1. Jai Ruby is a Research Scholar in Research & Development Centre, Bharathiar University, Tamilnadu, India. She has 14 years experience in teaching field and research. Her current areas of research are Data Mining and Mobile Communications.

2. Dr. K. David is working as an Associate Professor and Head, Dept. of Computer Science and Engineering, Rover Engineering College, Perambalur, Tamilnadu, India. He has over 15 years of teaching experience and about 4.5 years of Industry experience. He has published scores of papers in peer reviewed journals of national and international repute and is currently guiding 6 Ph.D scholars. His research interests include, UML, OOAD, Knowledge Management, Web Services and Software Engineering.