

# Recent Methodologies for Improving and Evaluating Academic Performance

Ajay Varma<sup>1\*</sup> and Y. S. Chouhan<sup>2</sup>

<sup>1\*,2</sup> Christian Eminent College, Indore

Available online at [www.isroset.org](http://www.isroset.org)

**Abstract-** In the real world, predicting the performance of the students is a challenging task. Many of the well known technical colleges are successful as they have meritorious students and faculty with them and a foolproof system working for them to grow continuously. The primary goal of Data mining in practice tends to be Prediction and Description. For educational institutions, the success of creation of human capital is the subject of a continuous analysis. To date, higher educational organizations are placed in a very high competitive environment and to remain competitive, organizations need better assessment, evaluation, planning, and decision making. As such, classification modeling for academic performance for the graduates could provide some insight to the university in order to take necessary information for improving the students' academic performance. Hence, the aim of this study is to provide the review of different data mining techniques that have been used in educational field with regard to evaluation of students' academic performance. Academic Data Mining used many techniques such as Decision Trees, Neural Networks, Naïve Bayes, K- Nearest neighbor, and many others. Using these techniques many kinds of knowledge can be discovered such as association rules, classifications and clustering. The discovered knowledge can be used for prediction and analysis purposes of student patterns.

**Keywords:** Classification, Data Mining, Bayesian Network, Neural Network

## INTRODUCTION

The topic of explanation and prediction of academic performance is widely researched. Data Mining Techniques is the promising methodology to extract valuable information in this objective. The data collected from different applications require proper method of extracting knowledge from large repositories for better decision making. Knowledge discovery in databases (KDD), often called data mining, aims at the discovery of useful information from large collections of data. In this perspective, Data Mining can analyze relevant information results and produce different Perspectives to understand more about the students' activities so as to customize the course for student learning. Data mining task is used in computer science and information technology aspects such as online learning and collaborative learning to facilitate students learning. Results are satisfactory because the existing technology aids and addresses the aspects of automated learning, practicing and evaluations of an academic cycle. They facilitate to understand/monitor students performance based on that 'moment' scores. At that time, there is no perfect usage of data mining techniques to facilitate Students Learning. So a better system is required to monitor and analyze student's performance based on a knowledge base constructed from automated learning, practicing and evaluations of the academic cycle.

For optimally analyzing the student performance, the classification task is used on student database to predict the students division on the basis of previous database. Classification is a data mining (DM) technique used to predict group membership for data instances(Lee & Mangasarian, 2001). These techniques have been applied in a great number of fields including bioinformatics, retail sales, counter-terrorism, stock market, real estate, cus-tomer relationship management, engineering, medicine, web mining and others. Each tech-nique differs in terms of complexity and power, and provides different model for different use. It applies modern and statistical computational technologies in its quest to expose useful pattern hidden within the large databases (Siraj&Abdoulha, 2011).An example of classification method in DM is illustrated in Figure 1.

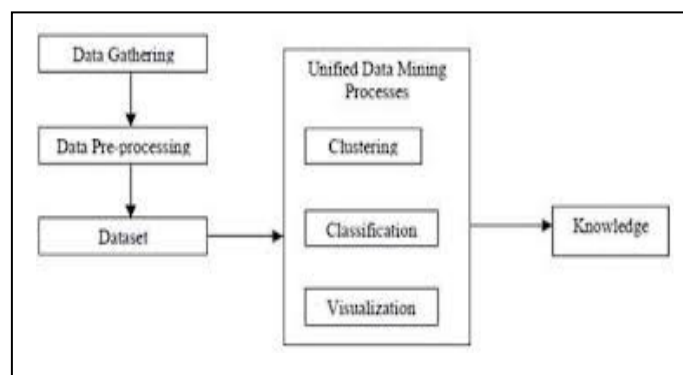


Figure 1: classification method in DM

Early work on classification in Statistics focuses on discriminant analysis, which constructs a set of discriminant functions, such as linear functions of the predictor variables, based on a set of training examples to be discriminated among the groups defined by the class variable. In recent years, the worldwide emerging challenges of the new technological revolution pay attention to education, so that higher education and future talent cultivation may be reformed (Maoyan, 2010). Classification is one of the DM techniques that are applicable in higher education environment; it is more complex and more powerful than regression, and provides more than just statistical model to better represent the data.

Educational DM is an emerging field that deals with the use of DM techniques on the data related to the field of education. For higher educational institutions, the success of creation of human capital is the subject of a continuous analysis (Osmanbegovic&Suljic, 2012). These institutions contain enormous amount of data which needs to be collected, integrated and utilized. The data can be mined too, for example, predict the academic outcome of students or predict student's behaviour or interest towards a subject. In such cases, classification techniques can be effectively applied on educational data to achieve these goals.

To date, higher educational organizations are placed in a very high competitive environment and to remain competitive, organizations need better assessment, evaluation, planning, and decision making. As such, classification modelling for academic performance for the graduates could provide some insight to the university in order to take necessary action for improving the students' academic performance. Academic performance is measured as the most recent cumulative grade point average earned based on a 4.0 grade point system, or the equivalent (Siraj&Haris, 2011). Abilities such as intelligence, aptitude and Cumulated Grade Point Average Score (CGPA) were generally believed to be the deciding factors of academic performance and it has been a standard practice worldwide. The academic performance of the students becomes an important criterion by which the management of the educational institutions and performance are judged (Mayston, 2003). Consequently, classification methods for predicting the academic performance also become important.

## CLASSIFICATION TECHNIQUES IN EDUCATION

Classification is the action of assigning an object to a category according to the characteristics of the object (An, 2005). Classification technology is the basic customized service for solving problems (Qu & Wang, 2008) which can be applied in both business and educational organization. Its procedure is an integral part in the scholastic environment as a means of reducing data to concise information for purposes of facilitating better hypothetical decision making (Shreve, Schneider, & Soysal, 2011) in improving academic performance. The technique is growing in popularity due to its capability in processing a wider variety of data than regression (Thair, 2009). Academic achievement is one of the most concerning issues for students, teachers, and parents. To understand the factors that contribute to academic achievement and to learn new methods of improving academic achievement, classification technique is relevant to be utilized in such environment (Pai, Lyu & Wang, 2010).

### Decision Tree

Decision trees are trees that classify instances by sorting them based on feature values (Thair, 2009; Ann, 2005). It is one of the most popular classification algorithms which induce a decision tree from data. A decision tree is a tree structured prediction model where each internal node denotes a test on an attribute, each outgoing branch represents an outcome of the test, and each leaf node is labeled with a class or class distribution. An object is then classified by following a path from the root to a leaf, taking the edges corresponding to the values of the attributes in the object (An, 2005).

Decision rules are a set of if-then rules. A typical decision tree learning algorithm adopts a topdown recursive divide-and conquers strategy to construct a decision tree. An example of decision rules is "if  $X < 1$  and  $Y = B$ , then the example belongs to Class 2". The algorithms differ in the search method (e.g. general-to specific, specific-to-general, or two-way search), the search heuristics that control the search, and the pruning method used.

Some applications of decision tree technique including the use of such technique to predict students' academic performance, and helped the tutors to identify the weak students and create better ways to teach those identified students (Kumar & Vijayalakshmi, 2011). The same technique is also used in (Superby et al., 2006; Yadav & Pal, 2012; Kovacic, 2010) on various groups of students. The outcome of the decision trees show a prediction of the number of students who are likely to pass or fail, or be at low, medium or high risk of being dropped out and so on. There are clear evidences of the usefulness of such prediction in helping the relevant authority, i.e. the university to devise strategies to assist the identified students (Yadav & Pal, 2012). Other examples on the use of decision tree in education are listed in Table 1.

**Table 1:** Applications of Decision Tree in Education

Year	Author	Description
2010	Jiabin et al.	Teaching evaluation model in adult higher education.
2012	Delibasic et al.	Comparative study on decision tree
2012	Qian et al.	Used to discover the relationship between different etiological factor of the weather
2012	Jayasimman& Raj	Decision tree induction algorithm to predict the requirement of the <b>future learner's need</b>
2011	Lu, Yang &Chen	Perceived understanding on decision tree algorithm
2009	Chen &Yen-Chu Hung	Employs a modified ID3 inductive learning to generate a decision tree for tennis or baseball coaches.
2009	Wook et al.	Clustering and Decision Tree for predicting academic performance
2008	Huang et al.	Dynamic assessment strategy for authentic learning
2007	Bresfelean	<b>To predict the students' choice in continuing their education</b>

### Bayesian Network

The essence of Bayesian classification is to utilize the characteristics of priori information to solve classification problems (Anderson &Matessa, 2008) and this classification is used in complex educational assessment (Mislevy et al., 2008). A lot of work has been devoted to the Bayesian Network structure learning for classification purposes. The related work shows that some structural learning approaches do take into account the existence of the class variable, and probably the most extended approach is to use the classification accuracy of the net as the met-ric value (Acid, Campos &Castellano, 2005; Sierra et al, 2001). Roos et al. (2005) proposed learning parameter for BNs devoted to classification tasks, maximizing the conditional (super-vised) likelihood instead of the joint (unsupervised) one. Friedman et al.(2003)employ struc-tural learning method that takes into account the existence of the class variable and obtains a tree-shaped structure, known as a Tree Augmented Network (TAN), in which the class variable is the root node. Super Parent method learns TAN structures not by means of probability, but guided by the accuracy known as greedy (Keogh &Pazzani, 2002). Another method, known as ELR algorithm maximizes the conditional likelihood of the class node to augment the discrimi-native capabilities of the acquired Bayesian Network (Greiner & Zhou, 2003). Grossman and Domingos (2004) use the BNC algorithm to learn the structure of a BN by maximizing the conditional likelihood. This method is similar to a greedy algorithm which combines user knowledgeand statistical data.

The naive Bayesian classifier is based on Bayes' theorem. Suppose that there are m clas-ses, C1, C2... Cm. The classifier predicts an unseen example X as belonging to the class having the highest posterior probability<sub>i</sub> conditioned<sub>(i)</sub> on X. In other words, X is assigned to a class if and only if.

Bayesian classification is used in education as demonstrated by Osmanbegovic&Suljic (2012); Bharadwaj and Pal (2011); Pai et al. (2010) and Pandey et al.(2011). Semi-Naive Bayesian model classifier (Di & He, 2010) is used to construct the characteristics of academic perfor-mance evaluation question answering system. Some of the applications of Bayesian Classifier in Education are listed in Table 2. Examples of applications of k-Nearest Neighbor

**Table 2:** Applications of Bayesian Classifier in Education

Year	Author	Description
2010	Colace & De Santo	Developing ontology in e-learning contest
2007	Yang et al.	Uncertainty factors in Student model
2010	Tan &Tay	Modeling the education domain to measure cognitive performance
2008	Zhi, Huaguo&Wenting	BN model to assess Learnerspsilas knowledge level in Data structure course.

2006	Colace et al.	New approach to design edutainment for medical education
2006	Colace & De Santo	Characterizing students, expertise
2009	Cheng, Rodriguez, & Basu	Multilevel tutoring model with multimedia to reduce lacking in adaptive tutoring
2009	Thai-Nghe, Busche & Schmidt-Thieme	Predict student performance at university level using Decision Tree and Bayesian Network
2011	Xiang	An intelligence question & answering database system for network education

### K-Nearest Neighbor

The k-nearest neighbor classifier is intuitive, easy to implement and effective in practice. It can construct a different approximation to the target function for each new example to be classified, which is advantageous when the target function is very complex, but can be described by a collection of less complex local approximations (Mitchell, 2007). The K-nearest neighbor algorithm requires three user-specified parameters: number of clusters K, cluster initialization, and distance metric (Tibshirani et al., 2001). The basic K-nearest neighbor algorithm has been extended in many different ways. Some of these extensions deal with additional heuristics involving the minimum cluster size and merging and splitting clusters such as Fuzzy c-means (Bezdek 2011; Dunn 2010 & Eschrich et al., 2003).

The k-nearest neighbor classifier classifies an unknown example to the most common class among its k nearest neighbors in the training data. It assumes all the examples correspond to points in an n-dimensional numeric attributes. Each sample represents a point in an n-dimensional space (An, 2005). The training samples are stored in an n-dimensional pattern space. When given an unknown sample, a k-nearest neighbor classifier searches the pattern space for the k training samples that are closest to the unknown sample. "Closeness" is defined

in terms of Euclidean distance (Thair, 2009), where the Euclidean distance, where the Euclidean distance between two points,  $X=(X_1, X_2, \dots, X_n)$  and  $Y=(Y_1, Y_2, \dots, Y_n)$  is

$$\sqrt{\sum}$$

The unknown sample is assigned to the most common class among its k nearest neighbors in k-nearest neighbor classifier. When  $k=1$ , the unknown sample is assigned the class of the training sample that is closest to it in pattern space. The limitation of such algorithm comes from its crisp nature in assigning cluster membership to data points. Since real world data seldom arranged in clear cut groups, there are apparent extraneous data points clearly not belonging to any of the clusters. The k-means algorithm is not capable of dealing with such a problem (Lei & Wenhui, 2003).

Examples of applications of K-Nearest Neighbor method are presented in Table 3.

**Table 3:** Applications of K-Nearest Neighbor in Education

Year	Author	Description
2010	Li & Tang	Optimize the multi-object in K-NN query algorithm
2009	Zou & Fan	The performance of the moving objects in mobile computing
2004	Huang et al.	Intelligent learning diagnosis system for web based thematic learning model

2004	Tsiriga&Virvou	Student model in a web based Intelligent tutoring system
2001	Pal & Ghosh	Few algorithms have been tested on both real and synthetic data sets and compared with the k-nearest neighbour rule (k-NN), m-multivariate rank nearest neighbour rule (m-MRNN), and k-nearest neighbour Dempster Shafer theory rule (k-NNDST).
1999	Kim et al.	Efficient query processing in multi-dimensional indexing structures in multimedia data applications
1999	Yamauchi et al.	Incremental learning methods with retrieval of inferred pattern in neural network based system.

### CONCLUSIONS

There are many useful techniques for mining student data for improving academic performance. So we see that Classification is more powerful than logistic regression and many other statistical experiments. In this paper, several applications on the classification methods have been presented. The literature indicates that among the three methods of classification with respect to academic performance, the most widely used classification techniques in the educational domain is decision tree.

### REFERENCES

- [1]. Al-Radaideh, Q., Al-Shawakfa, E. and Al-Najjar, M. (2006) 'Mining Student Data Using Decision Trees', The 2006 International Arab Conference on Information Technology (ACIT'2006) – Conference Proceedings.
- [2]. Ayesha, S., Mustafa, T., Sattar, A. and Khan, I. (2010) 'Data Mining Model for Higher Education System', European Journal of Scientific Research, vol. 43, no. 1, pp. 24-29.
- [3]. Baradwaj, B. and Pal, S. (2011) 'Mining Educational Data to Analyze Student's Performance', International Journal of Advanced Computer Science and Applications, vol. 2, no. 6, pp. 63-69.
- [4]. Chandra, E. and Nandhini, K. (2010) 'Knowledge Mining from Student Data', European Journal of Scientific Research, vol. 47, no. 1, pp. 156-163.
- [5]. El-Halees, A. (2008) 'Mining Students Data to Analyze Learning Behavior: A Case Study', The 2008 International Arab Conference on Information Technology (ACIT2008) – Conference Proceedings, University of Sfax, Tunisia, Dec 15- 18.
- [6]. Romero, C. and Ventura, S. (2007) 'Educational data Mining: A Survey from 1995 to 2005', Expert Systems with Applications (33), pp. 135-146.
- [7]. Shannaq, B., Rafael, Y. and Alexandro, V. (2010) 'Student Relationship in Higher Education Using Data Mining Techniques', Global Journal of Computer Science and Technology, vol. 10, no. 11, pp. 54-59.
- [8]. S. T. Hijazi, and R. S. M. M. Naqvi, "Factors affecting student performance: A Case of Private Colleges", Bangladesh e-Journal of Sociology, Vol. 3, No. 1, 2006.
- [9]. Acid, S., Campos, L. M. D. & Castellano, J. G. (2005). Learning Bayesian Network classifiers: searching in a space of partially directed acyclic graphs, Machine Learning 59, 213–235.
- [10]. Bharadwaj, B. K. & Pal, S. (2011). Data mining: a prediction for performance improvement using classification. International Journal of Computer Science and Information Security (IJCSIS), 9(4), 136-140.
- [11]. Bezdek, J.C. (2011). Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, 2(3), 196-202.
- [12]. Bresfelean, V.P.; , "Analysis and Predictions on Students' Behavior Using Decision Trees in Weka Environment," Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on , pp.51-56.
- [13]. Cheng, I.; Rodriguez, S.; Basu, A.; , "Multimedia and games incorporating student modeling for education," Technology for Education, 2009. T4E '09. International Workshop on , pp.91-94.
- [14]. Chenn-Jung Huang; Ming-Chou Liu; San-Shine Chu; Chin-Lun Cheng; , "Application of machine learning techniques to Web-based intelligent learning diagnosis system," Hybrid Intelligent Systems, 2004. HIS '04. Fourth International Conference on , pp. 242- 247.
- [15]. Dunn, J.C.(2010). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. Journal of Cybernetics. 3, 32–57.

- [16]. Colace, F.; De Santo, M.;, "Ontology for E-Learning: A Bayesian Approach," Education, IEEE Transactions on , vol.53, no.2, pp.223-233.
- [17]. Eschrich, S., Ke, Jingwei, Hall, L.O., &Goldgof, D.B. (2003).Fast accurate fuzzy clustering through data reduction.IEEE Transaction on Fuzzy Systems.11 (2), 262–270.
- [18]. Dunn, J.C.(2010). A fuzzy relative of the ISODATA process and its use in detectingcompact well-separated clusters.Journal of Cybernetics. 3, 32–57.
- [19]. Sierra, B., Serrano, N., Larrañaga, P., Plasencia, E. J., Inza, I., Jiménez, J. J., Revuelta, P. &Mora, M. L. (2001).Using Bayesian Networks in the construction of a bi-level multi- classifier, Arti ficial Intelligence in Medicine 22, 233–248.