

# Educational Data Mining for Improving Educational Quality

Pooja Gulati

Computer Science Department  
Research Scholar, IIS University  
Jaipur, India

Dr. Archana Sharma

Computer Science and Engineering Department  
Gurukul Institute of Engineering and Technology  
Kota, India

**Abstract**— This paper discusses about educational data mining and how it can be used to improve the functional activities of business of education through students, teachers and the way classes are arranged. Educational data mining help students, teachers and management to organize the way of teaching classes and their schedule to improve students performance and also help the student in taking decision for their academic carrier. Aim of the paper is to highlight how the education quality is improved with the help of educational data mining.

**Keywords**- Data Mining, Knowledge Discovery Database.

## I. INTRODUCTION

Data without the pertinent knowledge is useless. Knowledge can be seen as the patterns or characteristics of the data. Raw data is sometimes meaningless because what we want is the knowledge hidden in the data and not the data as such. That is why a new technology has emerged in the mid 1990's to deal with the discovery of knowledge from data. It is called knowledge discovery in databases (KDD) or simply data mining (DM). Uncovering hidden information is the fundamental goal of data mining. KDD(knowledge discovery in the databases) is particularly used for finding data patterns from the databases that give us new knowledge that in turn used in decision making process of the organization. KDD uses numbers of steps (Figure.1) to find these patterns i.e.

- 1) Take organization database
- 2) Select Target data
- 3) Arrange the data (free from errors and noise,)
- 4) Transform the data
- 5) Apply data mining techniques
- 6) Result will be some patterns

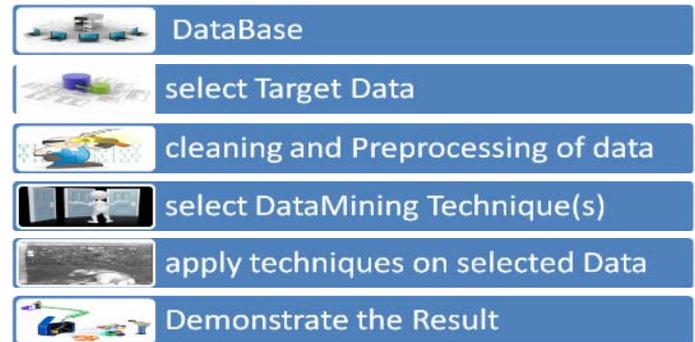


Figure 1 Steps Of KDD

The terms knowledge discovery and data mining are distinct. KDD refers to the overall process of discovering useful knowledge from data. It involves the evaluation and possibly interpretation of the patterns to make the decision of what qualifies as knowledge. It also includes the choice of encoding schemes, preprocessing, sampling, and projections of the data prior to the data mining step. Data mining refers to the application of algorithms for extracting patterns from data without the additional steps of the KDD process. Data mining is used when database is very large and from that large data base we extract some useful information that help the organization to handle task, in taking decisions based on the data etc.

## II. EDUCATIONAL DATA MINING

The field of Data Mining is concerned with finding new patterns in large amounts of data. Widely used in Business, it has scarce applications to Education. Data related to the field of education industry and mining of all such educational data is called educational data mining. It consists of data of student, teacher, faculty, courses, teaching hours, schools, colleges, universities etc. Educational data mining is used for predicting student performance and other learning techniques to remove the drawbacks in educational sector and improve the performance of students. Data mining describe the criteria of finding various patterns in data that are not explicitly part of

the data. These patterns are interesting which can be used to tell us something new. Educational data mining has emerged as an independent research area in recent years. Data related to the field of education industry and mining of all such education data is called educational data mining. It consists of data of student, teacher, faculty, courses, teaching hours, schools, colleges, universities etc. Educational data mining (also referred to as "EDM") is defined as the area of scientific inquiry centered around the development of methods for making discoveries within the unique kinds of data that come from educational settings, and using those methods to better understand students and the settings which they learn in[8]. Educational data mining methods often differ from methods from the broader data mining literature, in explicitly exploiting the multiple levels of meaningful hierarchy in educational data. Methods from the psychometrics literature are often integrated with methods from the machine learning and data mining literatures to achieve this goal. For example, in mining data about how students choose to use educational software, it may be worthwhile to simultaneously consider data at the keystroke level, answer level, session level, student level, classroom level, and school level. Issues of time, sequence, and context also play important roles in the study of educational data.

### III. LITERATURE REVIEW

Al-Radaideh et.al.[3], discussed in his paper that Student performance in university courses is of great concern to the higher education managements where several factors may affect the performance . This paper is an attempt to use the data mining processes, particularly classification, to help in enhancing the quality of the higher educational system by evaluating student data to study the main attributes that may affect the student performance in courses. Ayesha et.al. [7], conducted a study on student learning behavior. The students evaluation factors like class quizzes mid and final exam assignment are studied. It is recommended that all these correlated information should be conveyed to the class teacher before the conduction of final exam. This study will help the teachers to reduce the drop out ratio to a significant level and improve the performance of students. Bresfelean [4], made an analysis on the basis of which the evaluation of students capabilities of scoring marks can easily be predicted . Cortez et.al.[5], did research on education system of "Portugal", the results showed that a good predictive accuracy can be achieved, provided that the first and/or second school period grades are available. As a direct outcome of this research, more efficient student prediction tools can be developed, improving the quality of education and enhancing school resource management. Ramasubramanian et.al.[6], predict aspects of higher education students. In this paper they analyze that One of the biggest challenges that higher education faces today is predicting the behavior of students. Institutions would like to know, something about the performances of the students group wise. He proposed a problem to investigate the performances of the students when the large data base of Students information system (SIS) is given. Generally students'

problems will be classified into different patterns based on the level of students like normal, average and below average. In this paper we attempt to analyze SIS database using rough set theory to predict the future of students. Ramaswami et.al.[15], states "Educational data mining" (EDM) is a new growing research area and the essence of data mining concepts are used in the educational field for the purpose of extracting useful information on the behaviors of students in the learning process. Ramaswami et.al.[11], developed a predictive data mining model for students' performance so as to identify the slow learners and study the influence of the dominant factors on their academic performance. Ranjan et.al.[10], has applied data mining techniques on higher education data. The paper aims to purpose the use of data mining techniques to improve the efficiency of higher educational institutions. They have given the guidance to better decision making procedures and will improve the quality of instructions. According to Sun [12], the student learning result evaluation system is an essential tool and approach for monitoring and controlling the learning quality. This paper conducts a research on student learning result based on data mining. With this model in practice, student learning can become more energetic, more interesting, more challenging, and more suited to the times and this research paper will help to understand student learning evaluation system to generate theories.

### IV. ADVANTAGES OF EDUCATION DATA MINING

Education data mining can help both students and management for improving the quality of education. Student data mining is the mining of student data or data related to the students for ex. courses assignments, marks, student background etc. It allows the decision making process to use what-if scenario when we analyze student data. It helps the student in better progress and to improve better educational process. The student data mining process allows to have a better perspective on the student progress throughout the educational processes, and at the same time to analyze the information related to the specifics of the programs, courses, and course assignments[1]. This innovative approach allows the decision making process to use the what-if scenario when analyzing the student data, and other education related information in order to improve educational processes. The data related to the students' progress is retrieved from the students' records, imported into the data mining system, analyzed, and exported back. The educational data mining allows identifying and locating details about educational processes that need improvements, or those that perform very well and could be used as good examples. Educational data mining can assist in the design of the educational content. It can help in improvements in student academic performance. Data mining algorithms can help in discovering pedagogically relevant knowledge contained in databases obtained from Web-based educational systems. These findings can be used both to help teachers with managing their class, understand their students' learning and reflect on their teaching and to support learner reflection and provide proactive feedback to learners[2].

From the perspective of management, it helps in optimizing the organization by the maintenance of education infrastructure, increase in the courses and also increasing the areas of interest. One of the significant facts in higher learning institution is the explosive growth educational data. These data are increasing rapidly without any benefit to the management. The main objective of any higher educational institution is to improve the quality of managerial decisions and to impart quality education. Good prediction of student's success in higher learning institution is one way to reach the highest level of quality in higher education system. There are many prediction model available with difference approach in student performance was reported by researcher, but there is no certainty if there are any predictors that accurately determine whether a student will be an academic genius, a drop out, or an average performer[14].

## V. CONCLUSION

Educational data can provide us with an improved understanding of students' knowledge and better assessments of their progress. The educational systems currently face number of issues. Data mining provides a set of techniques, which can help the educational system to overcome these issues and enhance the quality of education. It will enable the institution to guide the students and help teachers and management in enhancing the performance of the institute. It can help to improve an education system by enabling better understanding of the students. The extra information can help the teachers to manage their classes better and the management to make better policies.

## REFERENCES

- [1] K. Michalski. and R. Michalski , "Analyzing Student Data In Order To Improve Educational Processes", Educational Data Mining and Reporting, pp. 1088-1094, 2004.
- [2]A. Merceron and K.Yacef, "Educational Data Mining: a Case Study", AIED, pp.467-474, 2005.
- [3] Al-Radaideh, A. Qasem, E. M. Al-Shawakfa, and M.I. Al-Najjar, "Mining Student Data Using Decision Trees", International Arab Conference on Information Technology, 2006.
- [4] Bresfelean and V.Paul, "Analysis and Predictions on Students' Behavior Using Decision Trees in Weka Environment", Proceedings of the ITI 2007 29th Int. Conf. on Information Technology Interfaces, 2007.
- [5] Cortez, Paulo and A. Silva, "Using Data Mining to Predict Secondary School Student Performance", Portugal, pp. 5-12, 2008.
- [6]P.Ramasubramanian, K. Iyakutti, and P.Thangavelu, "Enhanced Data Mining Analysis in Higher Educational System Using Rough Set Theory", African Journal of Mathematics and Computer Science Research Vol. 2(9), pp. 184-188, 2009.
- [7]Ayesha, Shaeela, T. Mustafa, A.R. Sattar, M.Inayat Khan, "Data Mining Model for Higher Education System", European Journal of Scientific Research, Vol.43, pp.24-29, 2010.
- [8] R.S.J.D.Baker, "Data mining for education", International Encyclopedia of Education (3rd edition), Oxford, UK: Elsevier, vol. 7, pp. 112-118, 2010.
- [9]M.H.Falakmasir and J. Habibi, "Using Educational Data Mining Methods to Study the Impact of Virtual Classroom in E-Learning", EDM, pp. 241-248, 2010.
- [10] Ranjan, Jayanthi and R.Ranjan, "Application Of Data Mining Techniques In Higher Education In India", Journal of Knowledge Management Practice, Vol. 11, January 2010.

- [11]M.Ramaswami and R. Bhaskaran, "A CHAID Based Performance Prediction Model in Educational Data Mining".IJCSI International Journal of Computer Science Issues, Vol. 7, January 2010.
- [12] Sun, Hongjie , "Research on Student Learning Result System based on Data Mining" .IJCSNS International Journal of Computer Science and Network Security, VOL.10, April 2010.
- [13] A.H.Cherif, "Computational Technologies For Accreditation In Higher Education", International Journal of Computers, Issue 3, Volume 5, 2011.
- [14]S.Sembiring, M. Zarlis, D.Hartama, S.Ramlia and E.Wan "Prediction of Student Academic Performance by an Application of Data Mining Techniques", International Conference on Management and Artificial Intelligence IPEDR, ) IACSIT Press, Bali, Indonesia, Vol.6, 2011.
- [15] M.Ramaswami and R. Bhaskaran, "A Study on Feature Selection Techniques in Educational Data Mining", Journal of Computing, Vol. 1, December 2009.