

# Use of Data Mining Tools and Techniques in Improving Productivity of Education Industry

Dr. Omkar V. Gadre

Institute of Science, Poona's Institute of Business Management & Research, Wakad Pune, India

**Abstract**— In today's scenario collecting a productive data & scrutinizing it is a tedious job. Data mining is one tool which can help to improve data quality in less time consuming manner. Data mining is known as Knowledge Discovery in Database (KDD). Education industry has grown rapidly over a last few years. Usage of technology has improved teaching learning methodology. A growing number of institutes in both public and private sector are offering huge diversity to all its aspirants students whether graduate, under graduate, post graduate, diploma or vocational courses. It is a concern for today's education system and this gap has to be identified and properly addressed to the learning community. Utilization of DM techniques in education sector is a developing and new growing research area. It is also known as Educational Data Mining. The Educational Data Mining is involved with developing the methods that helps to search specific types of data sets that come from education surroundings. Hence it has become important to understand the requirement of students and their academic progression. Educational Data Mining helps in a big way to answer the issues of predictions and profiling of not only students but other stake holders of education sectors. This paper can highlight some tools and techniques to discuss predictions of students progress and its impact on teaching learning methods.

**Key words:** Data Mining, Educational Data Mining, Knowledge Discovery of Database

## I. INTRODUCTION

Today's business environment is more competitive than ever. The difference between survival and defeat often rests on a thin edge of higher efficiency than the competition. This advantage is often the result of better information technology providing the basis for improved business decisions. The problem of how to make such business decisions is therefore crucial. One answer is through the better analysis of data. Knowledge Discovery in Databases (KDD) is a methodology to assess the value of the data and to leverage that value as an asset to provide large returns on the analytic investment. Private cooperation in building up new organizations, energized by the legislature, constrained advanced education to change its extension and goals to support over the long haul. The administrative bodies have surrounded rules for different foundations, personnel and different assets much of the time. This has been terribly damaged prompting sub-par training, finishing into un-employability of the understudies. All things considered, the objective of advanced education is to upgrade the nature of instruction in the college.

The problem that often confronts researchers new to the field is that there are a variety of KDD techniques available—which one to choose? All these tools give you answers. Some are more difficult to use than others, and they differ in other, superficial ways, but most importantly, the underlying algorithms used differ and the nature of these

algorithms is directly related to the quality of the results obtained and ease of use.

The utilization of such methods in education sector is to boost effectiveness of education sector. In the event that Knowledge Discovery procedures, for example, grouping, bunching, forecast, and affiliation can be connected to advanced education forms, it can offer assistance enhance understudies' execution, their life cycle administration, choice of the course and the major, their standard for dependability and allow/subsidize administration of an establishment.

It is acclaimed globally to improve educational quotient amongst students and teachers of all streams and faculties. Number of ERP softwares introduced and developed is increasing in educational sector.

Utility of Data Mining and Data Warehousing for educational institute along with designing and building the data warehouse for an educational institute by using open source tool freely available. Applications of Data warehouse are often considered as in large organizations and big business setups. But this technology can equally be useful in academics if rightly explored.

## II. OBJECTIVE OF THE PROJECT

- To identify students & teachers behavior
- Identifying trends of in educational sector
- Construction of profiles of all stakeholders pertaining to education industry
- Helping Management in developing business.

## III. IMPORTANCE & PROBLEM STATEMENT OF THIS PROJECT

Making in time decision making is need of an hour, education sector can't be far from that. Pune is considered as educational hub it will be easier to analyze and implement policies for betterment of this sector.

Data mining or knowledge discovery in databases (KDD) is the automatic extraction of implicit and interesting patterns from large data collections. Next to statistics and data visualization, there are many data mining techniques for analyzing the data. Some of the most useful data mining tasks and methods are clustering, classification and association rule mining. These methods uncover new, interesting and useful knowledge based on users' usage data. In the last few years, researchers have begun to apply data mining methods to help instructors and administrators to improve educational systems.

Association rule mining is one of the most well studied data mining tasks. It discovers relationships among attributes in databases, producing if-then statements concerning attribute-values. Association rule mining has been applied to web-based education systems from two points of view: 1) help professors to obtain detailed feedback of the e-learning process: e.g., finding out how the students learn on the web, to evaluate the students based on their navigation

patterns, to classify the students into groups, to restructure the contents of the web site to personalize the courses; and 2) help students in their interaction with the e-learning system: e.g., adaptation of the course according to the apprentice's progress, e.g., by recommending to them personalized learning paths based on the previous experiences of other similar students. Therefore, following research questions needs to be explored:

- 1) How can data mining help find students data?
- 2) How can data mining help to group student's behavior, and solve their problems?
- 3) How can some associative rules be found between students that could cluster
- 4) their types of problem?
- 5) How can data mining be used to identify those students who are at risk
- 6) especially in very large classes?
- 7) How can data mining be used in student's fees details and subjects?
- 8) How can data mining be used to grade student's academic progress

#### IV. RESEARCH APPROACH & METHODOLOGY

##### A. Orientation

The research starts with the orientation on the area of education data mining, what are data mining, education data mining about and which education system issues are in dire need of investigation by consulting websites of current university education system service offerings, reading news articles and discussing data mining application for education system and online education system with professionals.

##### B. Data collection

In this phase data are collected, mainly data related to basic concept of data mining, online education system in data mining, essential characteristics, architecture, secure educational key management model, benefit, advantages and disadvantages, client and server architecture, how to configure online education system, challenges and education system goals etc.

##### C. Data Analysis and Experiment

Once the data are collected analysis will be undertaken. During this phase the performance of education data mining system is analyzed which includes various techniques and also secure education key management model in university education system.

#### V. HYPOTHESIS OF STUDY

Considering the objective there is need to implement Knowledge Discovery Data in educational institutes. The following are objectives:

- H<sub>0</sub>: Use of KDD tools is more effective tool to improve e-governance in educational sector.
- H<sub>1</sub>: Use of only statistical methods useful to improve e-governance in educational sector.
- H<sub>2</sub>: No Need to implement the KDD tools for effectiveness.

- H<sub>3</sub>: Traditional Methods are very useful for betterment and improvement of E-Governance in educational institutes.

#### VI. APPLICATION OF DATA MINING IN EDUCATION SECTOR

Data mining can be applied in the following functional areas of an education sector.

- 1) Admission Forecasting: As more and more institutes are established from private sectors, the student expectation from these institutes is also increasing. They are taking admission in any new course only after screening various factors that are considered important for their overall growth. Classification and regression (CART) are the commonly used Data Mining techniques for doing such predictions. On the basis of pre-enrolment data they recommended students who are classified in the category of "High Risk" should be mentored in order to avoid any drop outs.
- 2) Profile of Students: KDD can also be used as an effective tool in profiling students based on both hard as well as soft skills. The hard factors include academic background, grades and achievements while soft factor includes communication, behavior, attitude, hobbies etc. Different Data Mining Techniques and algorithms have been used for this task.
- 3) Measuring Students Performance: It is most effective tool to check the academic progress of the students. Various statistical as well as data mining tools are widely available for this activity.
- 4) Teacher's overall performance: To evaluate teachers performance students' feedback is most common tool. identified factors that affect instructors' teaching performance in university by using stepwise regression and decision tree of data mining techniques.
- 5) Curriculum Development

#### VII. CONCLUSION

Data mining is used in industries for a variety of applications such as predicting and classifying customers and clustering customer's characteristics and prepares marketing strategies to segmented customers for the achievement of profitability. Similarly, universities can also apply data mining for predicting enrolment of students into various courses. Data mining can be applied for classifying and clustering students characteristics based on demographic, psychographic and behavioral variables. Data mining can also be applied by using if-then rule. In addition, it can describe the profile of successful and unsuccessful students based of GPA achieved during the semesters. It can also be used for dropout student, students' academic performance, teachers' performance, and students' complaints

#### REFERENCES

##### A. Bibliography

- [1] [www.kdnuggets.com](http://www.kdnuggets.com)
- [2] [http://www.data-mine.com/white\\_papers/nuggets\\_and\\_datamining](http://www.data-mine.com/white_papers/nuggets_and_datamining)
- [3] [www.mpi-inf.mpg.de/departments/d5/.../ss00/.../kdd99-p398-dorre.pdf](http://www.mpi-inf.mpg.de/departments/d5/.../ss00/.../kdd99-p398-dorre.pdf)

[4] [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1347303&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs\\_all.jsp%3Farnumber%3D1347303](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1347303&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D1347303)

*B. Webliography*

- [1] F. Marcotorchino, "Block seriation problems: A unified approach", *Applied Stochastic Models and Data Analysis*, 3,73-91
- [2] N. Wacholder and Y. Ravin, Disambiguation of Proper Names in Text, *Proc. of the 5th Conference on Applied Natural Language Processing*, April 1997,
- [3] Han, J., Kamber, M. (2001), *Data Mining: Concepts and Techniques*, Morgan-Kaufmann Academic Press, San Francisco
- [4] K. Cios, W. Pedrycz, R. Swiniarski, L. Kurgan, *Knowledge Discovery in Databases (KDD): A Knowledge Discovery Approach*, Springer, ISBN: 978-0-387-33333-5, 2007.
- [5] Daniel T. Larose, *Discovering Knowledge in Data: An Introduction to Knowledge Discovery in Databases (KDD)*, ISBN: 0471666572, John Wiley, 2004
- [6] Graham Williams, *Data Mining Desktop Survival Guide*

