

## AN OVERVIEW OF METHODS AND ALGORITHMS IN EDUCATIONAL DATA MINING

*Dr.K.Devasenapathy<sup>1</sup>*

### ABSTRACT

The objective of the paper is to deal with the area of Education Data Mining (EDM), which enables us to monitor the performance of the students. The educational institution's information is operated by the Educational data mining to enhance the performance of students as well as faculty. This paper deals with various algorithms and methods in data mining to be implemented in the field of EDM. This paper provides a general view of EDM.

**Keyword:** Education, Data Mining, EDM

### I. INTRODUCTION

Today data mining is used in various developing fields. Baker and Yacef [1] have proposed that Data Mining (DM) is the automatic extraction of large volumes of data to determine the patterns and trends from large data sets. The Special Interest Group on Knowledge Discovery and Data Mining-SIGKDD propose to tap information from the process of data mining from a large database and modify it into a logical construction for extended use [2].

The various tools in data mining can be used to identify the upcoming activities and ideas to be followed in the area of marketing to take some valuable ideas. Also, the tools can examine different cases and make the user deal with the situation. The website “www.theartling.com” defines, data mining as a tool to predict the appropriate information from a large data repository and get accurate information beyond human knowledge. [3].

DM retrieves useful information from the database. It can be used in various fields and is applied in many fields such as retail sales, bio-informatics and counter-terrorism. Nowadays, EDM is commonly used in educational institutions.

The learning from DM falls under two categories, namely: Specifically predictive data modelling and descriptive data modelling. The predictive modelling or directed data mining expands the data by using a target field of the data set. The descriptive data modelling or undirected data mining expresses the data without using a target field of the data set. There is concord that data mining is not a single-step process and knowledge-finding. It is the result or successive processes.

### II. DATA MINING IN EDUCATION

The United States' Department of Educational technology [w.ed.gov/edblogs](http://w.ed.gov/edblogs), defines the purpose of education to be to obtain talent and be aware of the surroundings to face problems in life. Also its goal is to develop a love of learning that stays throughout one's life. The goal of educational institutions is to impart adequate knowledge and skills to uplift students' career within a particular period [4].

There are a lot of techniques for data mining in data analytics. One can't analyze the huge volume of data in student database without the help of automated technique of analysis. The data mining techniques in EDM are used to guide the students to learn or understand the educational phenomena.

“[www.educationaldatamining.org](http://www.educationaldatamining.org), indicates that EDM is a developing field, related to emerging methods to show the

---

<sup>1</sup>Asst. Professor, Department of CS, CA & IT,  
Karpagam Academy of Higher Education  
Coimbatore, Tamilnadu,India

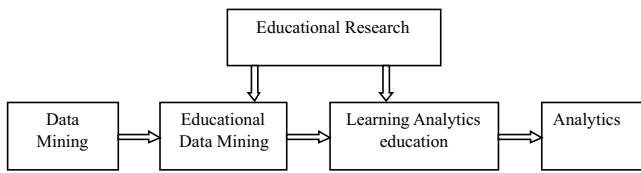


Figure 1: EDM and Learning Analytics

distinct and large volume of data from the educational background”[5].

Figure 1 describes EDM and learning analytics. Educational Data Mining may give an institution the information necessary to take action before a student dropout or to allocate resources with an accurate estimate of how many students will take a particular course. This study addresses the capabilities of data mining and its applications in higher education. The aim of EDM is to know the processes or methodologies followed to analyze educational software or teacher support. The objective of EDM is to predict students' success.

**Data mining algorithms in education** There are many algorithms used in data mining. Some important algorithms used in education field are given below:

**Supervised Learning**

Supervised Learning algorithm consists of a dependent variable. This learning is used to predict from a given set of predictors. One can generate a function that maps source to the required target. It is iterative process of model that helps to attain the essential accuracy level.

Table 1: Machine Learning Algorithms Used in EDM Supervised Learning Unsupervised Learning

Supervised Learning	Unsupervised Learning
<ul style="list-style-type: none"> <li>• Decision Trees</li> <li>• K-Nearest Neighbour</li> <li>• Multi-Layer Perceptron</li> <li>• Naive Bayes</li> <li>• Random Forest</li> <li>• Regression Algorithm</li> <li>• Support Vector Machine</li> </ul>	<ul style="list-style-type: none"> <li>• Apriori Algorithm</li> <li>• K-means Clustering</li> </ul>

**Unsupervised Learning**

Unsupervised learning does not contain any target value. This learning involves clustering and segmenting of data into different groups.

**Functionalities of EDM**

Baker and Yacef have suggested that EDM is an emerging discipline which contains the unique data types that originate from educational settings and the methods in EDM are useful for a better understanding of students and their learning environment.

EDM extracts useful data from an unknown educational database which delivers information on the performance and assessment of students' learning method. An analysis of data can be done from the educational settings in which students are studying. Research in EDM using data mining is increasing steadily in the field of education. The process of EDM involves collection and storage of data for the purpose of analysis and it may help in the motivation of students' learning and evaluation.

**III. EDM LEVELS**

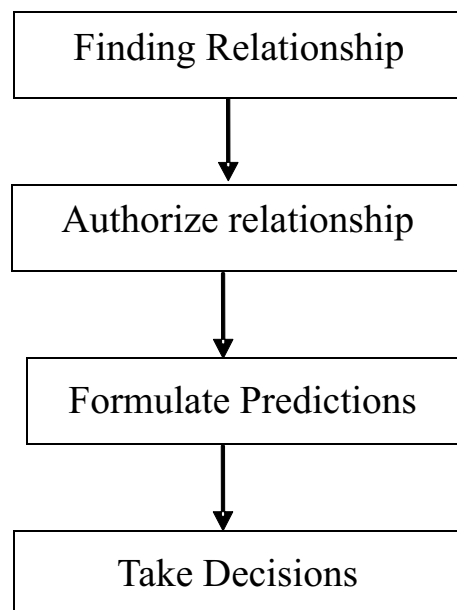


Figure 2: Levels in EDM

**EDM goals are:**

- Identifying students' learning behaviour by consolidating student details
- Evolving area of student interest to learn
- Investigating the consequences of various academic support delivered by learning software
- Assessing the technical knowledge to create models to integrate the learner and learner domain

4. US Department of Education office of Educational Technology Available from : <http://www.ed.gov/edblogs/files/2012/03/edm-la-brief.pdf>> . [10 April, 2012].

5. International Educational Data Mining Society Available from: <http://www.educationaldatamining.org>.

**IV. EDM METHODS**

Generally, EDM methods are found to be same, but some differences exist depending on the educational data set. Baker and Yacef have classified EDM areas as Classification and Prediction, Clustering, Association mining, Characterization of data from human judgment and discover with model.

**IV. CONCLUSION**

This paper discussed the various algorithms in EDM. Different methods are to be discussed in the process of EDM. Various functionalities in data mining and different levels are clearly seen. This paper is generic approach to Educational Data mining.

**VI. REFERENCES**

1. Baker, RSJd & Yacef, K 2009, 'The State of Educational Data Mining in 2009: A Review and Future Visions', Journal of Educational Data Mining. vol. 1, no. 1, pp. 1-16.
2. Data Mining Curriculum ACM SIGKDD 2006-04-30. Retrieved 2014-01-27
3. An Introduction to Data Mining, 2004 Available from:< [http:// www.theartling.com/text/dmwhite/dmwhite.htm](http://www.theartling.com/text/dmwhite/dmwhite.htm)>. [2 April, 2004].