

# Educational Data Mining For Student Support in Interactive Learning Environment

**Gaurav Jindal<sup>1\*</sup>, Sakshi Garg<sup>2</sup>**

<sup>1</sup>Assistant Professor, RDIAS

<sup>2</sup>BBA Student, RDIAS

DOI: <https://doi.org/10.26438/ijcse/v7i5.15581565> | Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

Accepted: 16/May/2019, Published: 31/May/2019

**Abstract**—Educational Data Mining is an emerging interdisciplinary research area that deals with the development of methods to explore data originating in an educational context. EDM uses computational approaches to analyze educational data in order to study educational questions. Firstly, it introduces EDM and describes and why we Need of EDM then, how EDM support Students in Interactive Learning, Applications of EDM, then goes on to list the most typical/common tasks in the educational environment that have been resolved through data mining techniques, LMS Environment, Software used in EDM (MOODLE), Advantages of DM in Education System and what Challenges we are facing in the Field of EDM and finally some of the most promising future lines of research are discussed.

**Keywords:** EDM, LMS, DM Techniques, Educational Systems.

## I. INTRODUCTION

Educational Data Mining (EDM) is the application of Data Mining (DM) techniques to educational data, and so, its objective is to analyze these types of data in order to resolve educational research issues. DM can be defined as the process involved in extracting interesting, interpretable, useful and novel information from data. It has been used for many years by businesses, scientists and governments to sift through volumes of data like airline passenger records and the supermarket scanner data that produces market research reports. EDM is concerned with developing methods to explore the unique types of data in educational settings and, using these methods, to better understand students and the settings in which they learn. On one hand, the increase in both instrumental educational software as well as state databases of student information has created large repositories of data reflecting how students learn. On the other hand, the use of Internet in education has created a new context known as e-learning or web-based education in which large amounts of information about teaching-learning interaction are endlessly generated and ubiquitously available. All this information provides a gold mine of educational data. EDM seeks to use these data repositories to better understand learners and learning, and to develop computational approaches that combine data and theory to transform practice to benefit learners.

## II. Overview of Data Mining and its Techniques

Data mining refers to “*Extracting*” or “*Mining*” knowledge from large amounts of data.

1. The process that finds a small set of precious nuggets from a great deal of raw material.
2. It allows users to analyze data from many different dimensions by applying various data mining techniques.
3. Data mining as a synonym for another popularly used term, Knowledge Discovery from Data, or KDD.
4. It is a Knowledge Discovery process which helps in determining relationship between different objects and help to analyze the pattern.

### Knowledge Discovery Process

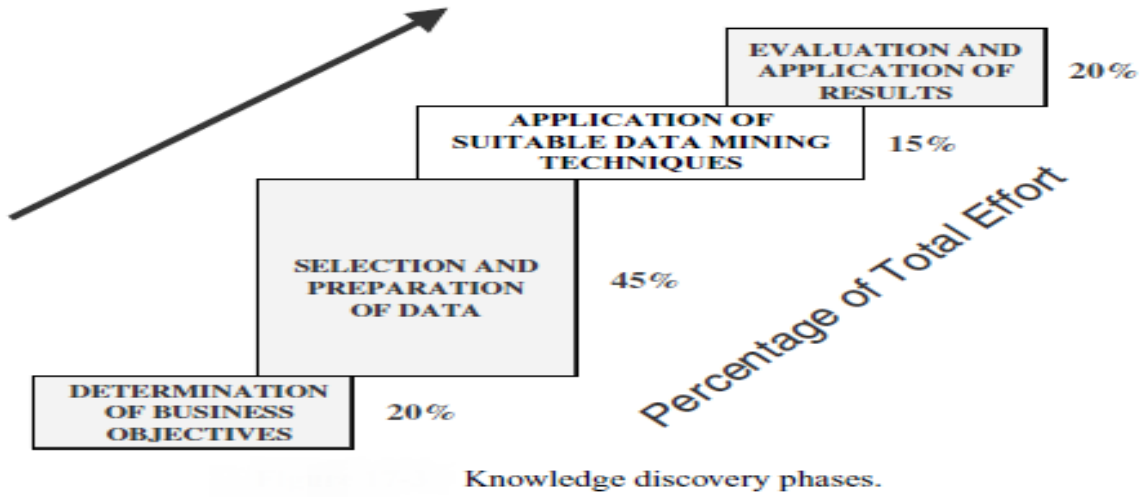


Figure 1: Phases of Knowledge Discovery Process & Effort Distribution

### III. Educational Data Mining

The EDM process converts raw data coming from educational systems into useful information. The data can be collected from historical and operational data reside in the databases of educational institutes. Educational data mining can be applied to discover patterns in non trusted datasets to automate the decision making process of learners, students and administrators. EDM consists of these phases:

- Data Selection
- Pre-processing
- Transformation
- Data mining
- Interpretation/Evaluation

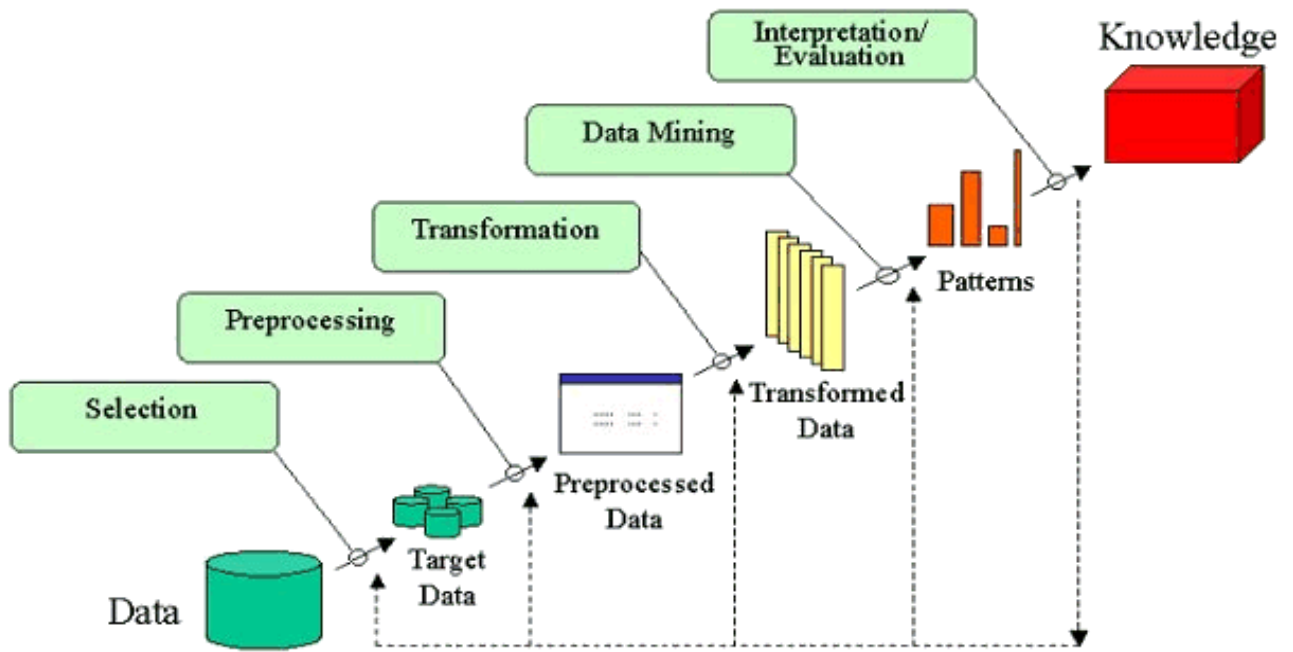


Figure2: Phases of Educational Data Mining

Researchers have used various data mining methodologies in different ways to understand or identify the learning models and learning patterns of the students. Various learning management systems can be used for providing the study programs for the students who can be connected in open distance mode or in a much more hybrid manner. But applying the same teaching principle on all the students in the same way will not cope with the ultimate goal of any learning management systems available, which is to help the students to learn rather helping them to pass. But understanding the nature of the learning in each of the students requires a huge effort such that the researching of educational information need to be followed by a stepwise process to reach the final discoveries. Like in any data mining research, some steps can be given as the basic steps that need to be followed in the knowledge discovery process which is shown in above figure. The same principles can also be applied in the learning analytics processes where the data collecting step will be applied to different educational data sources to aggregate necessary data for the analysis and data pre-processing steps will be applied to arrange the data into a proper arrangement which can be applied on the mining algorithms to discover the hidden information and patterns.

### **EDM Stakeholders**

There are four main users and stakeholders involved with educational data mining.

1. **Learners** - Learners are interested in learning methods to improve the learner's experience and performance.
2. **Educators** - Educators attempt to understand the learning process and the methods they can use to improve their teaching methods.
3. **Researchers** - Researchers focus on the development and the evaluation of data mining techniques for effectiveness.
4. **Administrators** - Administrators are responsible for allocating the resources for implementation in institutions.

## **IV. WHY WE NEED EDUCATIONAL DATA MINING**

Educational data mining provides large amount of knowledge which is diverse and distributed globally. It includes sharing and manipulation of knowledge with distance interactive in real time. It is easy to manage diverse educational materials for teachers by virtual platform. It helps to view and analyze the outline performance of each student in the given activity on the course. Special needs-Students often drop out of education because they require specific attention to a certain need, such as dyslexia or other learning disabilities.

### **EDM support students in Interactive Learning**

1. It deals with the assessment of student's learning performance.
2. It provides course adaptation and learning recommendations based on student's learning behavior.
3. In dealing with the evaluation of learning material and educational web-based courses.
4. It involves feedback to both teachers and students in e-learning courses.
5. Development for detection of a typical student's learning behaviors.

### **Applying DM in education System**

There are increasing research interests in using data mining in education. This new emerging field EDM, concerns with developing methods that discover knowledge from data originating from educational environments. The data can be collected from various educational institutes that reside in their databases. The data can be personal or academic which can be used to understand students' behavior, to assist instructors, to improve teaching, to evaluate and improve e-learning systems, to improve curriculums and many other benefits.

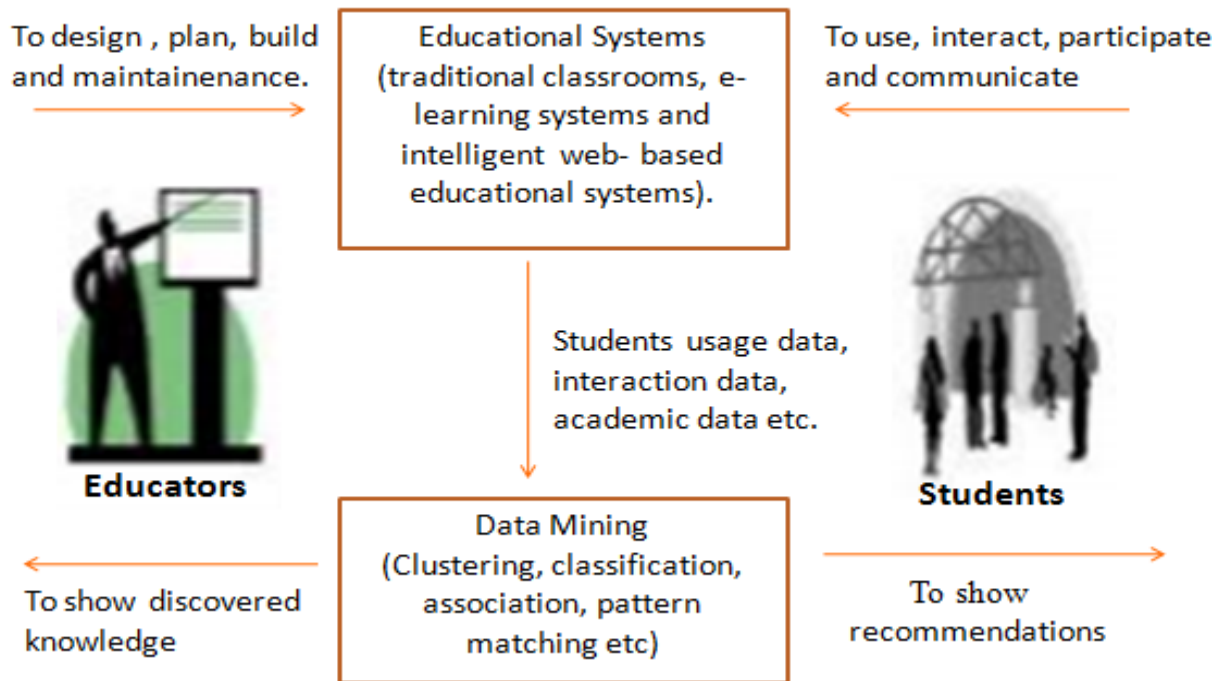


Figure 3 : Illustrating Data Mining Application in Education System

The above figure illustrates how the data from the traditional classrooms and web based educational systems can be used to extract knowledge by applying data mining techniques which further helps the educators and students to make decisions.

#### **DM technique applied in EDM**

In this study, data gathered from college students was analyzed using a data mining technique namely k-means clustering.

#### **Example using Clustering Technique:**

In this we are using clustering technique that can be used to distinguished intelligent student from dull student according to their performance in exam.

1. **Preparation :** In this step data stored in different tables was joined in a single table after joining process errors were removed. Here we are considering various cases.
  - Case 1 – If mid-term grade = Low, Internal Exam Grade = Low, Practical work = poor, Final Term = Average, then final grade = Low.
  - Case 2 – If mid-term grade = Average, Internal Exam Grade = Average, Practical work = Good, Final Term =Average, then final grade = Average.
  - Case 3 – If mid-term grade = High, Internal Exam Grade = Good, Practical work = High, Final Term = High, then final grade = Good.
2. **Implementation of mining model:** In this steps k-means clustering algorithm was applied to the proposed data and get valuable information, k- means is an old and most widely used clustering algorithm by Mac Queen in 1967.
 

**Algorithm Basic K-means Algorithm**

  1. Select K points as the initial centroids.
  2. Repeat.
  3. From K- cluster by assigning all points to the closest centroids.
  4. Recomputed the centroid of each cluster.
  5. Until the centroid don't change.

3. **Results:** We grouped the students regarding their final grades in three ways:

- i. Assign possible labels that are the same as number of possible grades
- ii. Group the students in three classes “High”, “Medium”, “Low”.
- iii. Categorized the students with one of two class labels “Passed” for marks greater than or equal to 40 and “Failed” for marks less than 40.

**Table1: Showing Grades for Students**

Class	Marks	No. of Students	Percentage
Fail	$> 40$	6	5
Pass Class	$40 < \text{percentage} < 50$	14	11.67
Second Class	$50 < \text{percentage} < 55$	23	19.17
Higher Second Class	$55 < \text{percentage} < 60$	17	14.17
First class	$60 < \text{percentage} < 70$	47	39.16
First class with distinction	$70 \leq$	13	10.83

**Table 2: Student Classification Summary**

Class	No. of Students
High	13
Medium	87
Low	20

**Table 3: Student Categorization Scheme**

Class	Marks	No. of Students	Percentage
Passed	$\geq 40$	114	95
Failed	$< 40$	6	5

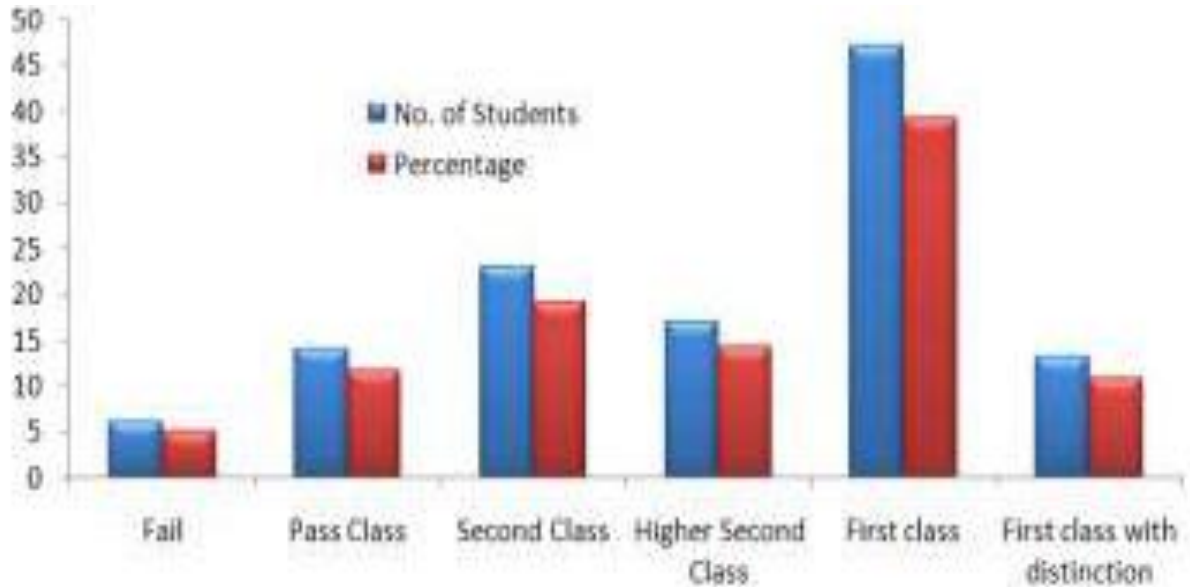


Figure 4: Student Performance Chart

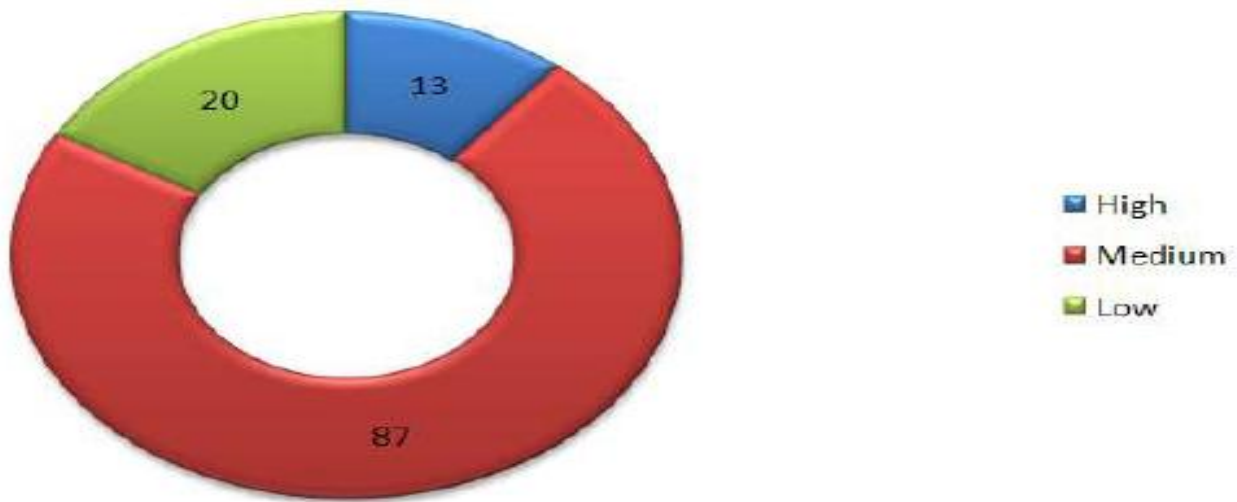


Figure 5: Student Performance Report Analysis

### Applications of EDM

- 1. Providing feedback for supporting instructors:** The objective is to provide feedback to support course authors/teachers/administrators in decision making (about how to improve students' learning, organize instructional resources more efficiently, etc) and enable them to take appropriate proactive and/or remedial action. It is important to point out that this task is different than data analyzing and visualizing tasks, which only provide basic information directly from data (reports, statistics, etc.). Moreover, providing feedback divulges completely new, hidden and interesting information found in data.
- 2. Predicting student performance:** The objective of prediction is to estimate the unknown value of a variable that describes the student. In education the values normally predicted are performance, knowledge, score or mark. This value can be numerical/continuous value (regression task) or categorical/discrete value (classification task). Regression analysis finds the relationship between a dependent variable and one or more independent variables. Prediction of a student's performance is one of the oldest and most popular applications of DM in education, and different techniques and models have been applied.
- 3. Detecting undesirable student behaviors:** The objective of detecting undesirable student behavior is to discover/detect those students who have some type of problem or unusual behavior such as: erroneous actions, low motivation, playing games, misuse, cheating, dropping out, academic failure, etc. Several DM techniques (mainly,

classification and clustering) have been used to reveal these types of students in order to provide them with appropriate help in plenty of time.

4. **Grouping students:** The objective is to create groups of students according to their customized features, personal characteristics, etc that can be used by the instructor/developer to build a personalized learning system, to promote effective group learning, to provide adaptive contents.
5. **Constructing courseware:** Constructing courseware is to help instructors and developers to carry out the construction/development process of courseware and learning contents automatically.
6. **Planning and scheduling:** The objective of planning and scheduling is to enhance the traditional educational process by planning future courses, helping with student course scheduling, planning resource allocation, helping in the admission and counseling processes, developing curriculum.

### LMS (Learning Management System)

A learning management system (LMS) is a software application or Web-based technology used to plan, implement, and assess a specific learning process. LMS provides an instructor, a way to create and deliver content, monitor student participation, and assess student performance. LMS provides students with the ability to use interactive features such as video conferencing and discussion forums. Examples: MOODLE, Edmode, SkillSoft, Sakai, Dokeos etc.



Figure 6: Overview of MOODLE

### MOODLE

MOODLE is the acronym for “*Modular Object-Oriented Dynamic Learning Environment*”, It is a free online learning management system. It enables educators to create their own private website. It is used all over the world by universities, schools, companies and independent teachers. Users can create dynamic courses using different features that extend learning anytime and anywhere.



Figure 7: Various Roles of MOODLE

### MOODLE ACTIVITIES

Some of the MOODLE activities are as follows:

- CHAT
- Database
- Glossary
- Quiz
- Forum
- Choice
- Assignment

### IV. CONCLUSION

EDM has been introduced as an up and coming research area related to several well-established areas of research including e-learning, web mining, data mining, etc. The purpose of this research was to determine if DM technique will help us group students. Evaluating performance in the e-learning systems becomes a massive challenge because of the different factors which affect the learning models. Applying data mining methodologies on the educational data has brought a new research discipline where the existing methodologies have been used to model the learning behaviors of the learners. Therefore models for evaluating the student performance with acceptable accuracy levels and quality predications still need to be researched more and existing learning analytics should be implemented in such a manner in which they can be used by the knowledge producers with more user friendliness effective manner. Recommendation for future research is the development of a Module for MOODLE dynamic and a real time Data collection, and also a recommendation system for student. The goal is to help student learn better, and prevent behavior which leads to poor student.

### REFERENCES

- [1]. An Empirical Study of the Applications of Data Mining Techniques in Higher Education Dr. Varun Kumar Department of Computer Science and Engineering ITM University
- [2]. Educational Data Mining: A Review of the State-of-the-Art By Cristóbal Romero, Member, IEEE, Sebastián Ventura, Senior Member, IEEE.
- [3]. Educational Data Mining for Grouping Students in E-learning System Divna Krpan, Slavomir Stankov Faculty of Science, Teslina 12, Split, 21000 Croatia.
- [4]. Importance of Data Mining in Higher Education System” 1Bhise R.B., 2Thorat S.S., 3Supekar A.K.
- [5]. Use of Data Mining Methodologies in Evaluating Educational Data Thilina Ranbaduge Department of Information Technology, University of Moratuwa, Srilanka.