



DATA MINING TECHNOLOGY AN EMERGING TREND IN HIGHER EDUCATION SECTOR

* Mrs. Suwarna Mulay
**Dr. Shubhangi Potdar

*BYK College of Commerce, Nashik, SPPU, India.

**MCA Department, DVVPPF's, IBMRD, Ahmednagar, SPPU, India.

Abstract:

Data Mining is an emerging technology used in higher education sector. The huge data of student from various online sources can be analysed to find meaningful information. New trends and patterns can be easily collected by analyzing the student data using data mining technology. The various applications of data mining technology in the field of higher education sector and different data mining techniques are discussed in this paper. This new technology can be used to improve the quality of higher education institutions and to enhance teaching-learning process.

Keywords: Data Mining, Data Mining Technology, Higher Education Institution

Copyright © 2023 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0) which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

Introduction:

Data mining in higher education is also referred to as Educational data mining (EDM). Higher education institutions have to face biggest challenge to compete with foreign universities. These institutions can be able to survive in the competition by making use of enhanced technologies in the field of education. Data mining technology is an important technology which can help to improve educational quality. The emerging trend of EDM helps higher education institutions to improve their quality and better decision making. EDM provides various facilities for collecting, analyzing and visualizing student data. Academic analytics can be possible after collection of huge academic data from different sources of Internet. Data mining techniques helps to analyse rapid amount of student data to extract meaningful information from large datasets. It is the

need of higher education institutions to embed data mining tools for the essential academic progress. As students are the building blocks of higher education, to impart quality education is very important aspect of the educational system. Data mining techniques can be useful to maintain educational quality by classifying the student and sending time to time intimations to them regarding their progress. Data mining tools are also useful to improve the student CGPA, grade prediction, subject selection, course selection, to enhance teaching-learning process, alumni registration, institution advertising, strategic decision making etc.

Data Mining Definition and Techniques:

Data Mining is the process of extracting information to identify patterns, trends and useful data that would allow the business to take data driven decisions from huge sets of data. Data mining process is also called as Knowledge Discovery of Data (KDD).

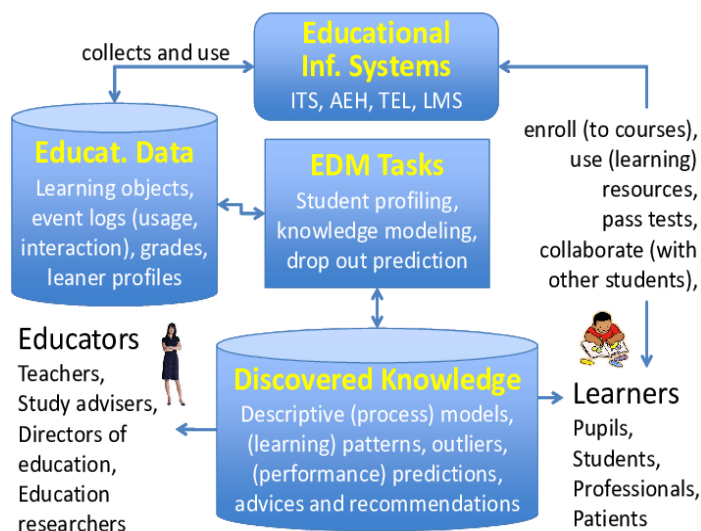


Fig: Use of Data Mining in Higher Education

Data mining was previously used in various business applications such as cross marketing, market-basket analysis, identifying customer purchase behaviour etc. But because of the huge data gathered in the field of education, data mining techniques were started implementing in the field of education to identify different patterns and trends in higher education. This is referred to as Educational Data Mining (EDM). Data mining has various tasks that can be applicable in the field of higher education. These tasks are as follows:

- 1) **Clustering:** Clustering refers to identify similar classes of objects. Clustering techniques are useful to identify sparse and dense regions of objects. It is used to identify correlation between the attributes of data. Clustering technique is cheaper than the classification technique of data mining and can be mostly used in data preprocessing, attribute selection, subset selection etc. e.g. in higher education system we can make clusters of similar students according to their grasping power like low, medium and average students.
- 2) **Classification:** Classification is an important data mining technique that can be used to classify population in different classes according to the similar features. Classification technique uses

previous knowledge to build a learning model and then uses this model as a binary variable for the new data. Bayesian classification, Neural Network, Decision Tree, Support Vector Machine are some of the classification algorithms that can be applicable in educational sector.

- 3) **Association Rule:** Association and correlation are used to find frequent item set among the huge data sets. It is used to find relationship between different variables in the dataset. This technique is also used to find strength of relationship between different variables. This is the most widely used data mining technique in Educational Data Mining. Association rule mining develops the relationship in the form of *If then Else*. E.g. {If the student SGPA < 3 → more chances of dropout from the course} etc.
- 4) **Prediction:** Prediction is also an important data mining technique used for data analysis. It studies the past data accordingly predict the future events. This technique uses algorithm that derives the model or predict according to the training dataset. Here, the model uses continuous variables or ordered value. Regression algorithms are generally used for prediction. E.g. predicting the value of house depending on various factors such as area, number of rooms etc.

Data Mining Techniques: Data mining technology includes various techniques that can be either supervised or unsupervised. Supervised data mining technique uses labeled data and have clear objective, while unsupervised data mining methods uses unlabeled data and have more exploratory purpose. The most widely used data mining techniques which are used in higher education sector are described below:

- 1) **Regression:** Regression is a powerful tool for summering the nature of the relationship between variables and for making predictions of likely values of the dependent variables. Regression analysis can be applied in data mining to predict



students' CGPA, subject selection, to identify factors that affect student CGPA etc. Regression analysis can be categorized as simple regression and multiple regression.

- 2) **Decision Tree:** Classification and prediction of large data can be possible with the help of decision tree technique. Decision tree is used for student profiling. It consists of nodes and branches, nodes are connected by branches, time flows from left to right, each branch represents a decision or a possible event. Decision tree make classification easy and understandable and also result-oriented. Higher education institutions can use this technique to classify students' performance, behaviour, and expectations.
- 3) **Outlier Detection:** A database may contain data objects that do not comply with the general behavior or model of the data. These data objects are Outliers. The investigation of OUTLIER data is known as OUTLIER MINING. An outlier may be detected using statistical tests which assume a distribution or probability model for the data, or using distance measures where objects having a small fraction of "close" neighbors in space are considered outliers. Rather than utilizing factual or distance measures, deviation-based techniques distinguish exceptions/outlier by inspecting differences in the principle attributes of items in a group.
- 4) **Neural Network:** This technique is used for analyzing large and complex data. It is used in educational institutions to study the course selected by the student, course satisfaction by the student and specialization selection etc. The input data is represented by an entity known as neuron. These neurons are connected with the related neurons inside the same cluster.

Applications of Data Mining in Higher Education Institutions:

1) Students Performance Prediction:

Data mining techniques are used to predict the students' performance at the early stage of completion of their degree. There are many factors which affect the performance of the student like age, gender, yearly income of the family, education of parents, whether parents are working or not working etc. By considering all these factors, performance will be predicted at early stage of their final examination, so that the student can get ample time for the improvement.

2) Teachers Teaching Performance Prediction:

Teachers teaching performance can be generated by collecting the teachers' feedback from the students. Various factors like subject knowledge of the teacher, class control, communication skill, use of audio/visual aids etc. affect the performance of the teacher. By considering all these factors teachers teaching performance can be calculated. This performance can be evaluated using appropriate data mining techniques and necessary actions can be taken accordingly to improve the performance.

3) Syllabus Framing:

Data mining techniques like decision tree, decision forest can help to correlate course category and the enrolled students which is very useful to design the course curriculum. Also, depending on the previous results new changes can be made in the course syllabus which helps the students to easily understand the subject.

4) Student Enrollment Management:

This technique is used by the higher education institutions to describe their enrollment strategies. Enrollment management is an organizational concept and a systematic set of activities designed to enable educational institutions to exert more influence over their student enrollments. Such practices often include marketing, admission policies, retention programs, and financial aid awarding. Strategies and tactics are informed by

collection, analysis, and use of data to project successful outcomes. Activities that produce measurable improvements in yields are continued and/or expanded.

5) Data Visualization:

Data Visualization is one of the important application of data mining in education sector. Students as well as teachers data can be visualized with the help of various data mining tools like WEKA, Rapid Miner and Orange etc. Students data can be analyzed easily which can be further used for appropriate decision making. Data Visualization can help the institutions to maintain or improve the quality of education.

Conclusion:

Data mining technology is an emerging technology in the higher education sector. This technology is used to improve the quality of education by improving the results. Clustering, classification, prediction and association task of data mining helps higher education institutions for administrative task such as strategic decision making, admission management etc. So, this paper concludes that data mining technology is the future trend which has various application in higher education sector.

References:

- Naeimeh DELAVARI, Somnuk PHON-AMNUAISUK, Mohammad Reza BEIKZADEH, Multimedia University, Malaysia, “Data Mining Application in Higher Learning Institutions”
- Mrs. Bharati M. Ramageri, “Data Mining Techniques & Applications”, Indian Journal of Computer Science and Engineering, Vol. 1 No. 4 301-305
- Abdulmohsen Algarni, “Data Mining in Education”, International Journal of Advanced Computer Science and Applications, Vol 7, No. 6, 2016.
- Dr. Mohd Maqsood Ali, “Role of Data Mining in Education Sector.”, IJCSMC, Vol-2, Issue-4, April 2013, pg-374-383.
- Roberto Llorente and Maria Morant, “Data Mining in Higher Education”.
- <https://www.researchgate.net/figure/Educational-data-mining-in-a-nutshell-fig1-254008370>
- <https://www.javatpoint.com/data-mining>
- https://www.researchgate.net/publication/304808426_Data_Mining_in_Education
- Data Mining in Education
Abdulmohsen Algarni
Data Mining in Education
Abdulmohsen Al

Cite This Article:

Mrs. Mulay S. & Dr. Potdar S. (2023). DATA MINING TECHNOLOGY AN EMERGING TREND IN HIGHER EDUCATION SECTOR. In Aarhat Multidisciplinary International Education Research Journal: Vol. XII (Number VI, pp. 152–155). AMIERJ. <https://doi.org/10.5281/zenodo.10517990>