A PUBLICATION OF MOTHERHOOD UNIVERSITY, ROORKEE

(Recognized by the UGC with the right to award degrees u/s 22(1) of the UGC act 1956 andestablished under Uttarakhand Government Act No. 05 of 2015)

Motherhood International Journal of Multidisciplinary Research & Development

A Peer Reviewed Refereed International Research Journal

Volume I, Issue IV, May 2017, pp. 01-09

ONLINE ISSN-2456-2831

A Study on Data Mining in Education System

[1]Kamaljeet Kaur, [2]Atul Garg
[1]Research Scholar, [2]Associate Professor,
MM University, Mullana, Ambala

Abstract

Mining of data in education system is growing nowadays in research field. Various tools and techniques are developed to discover hidden patterns existing in the data from an educational view point. Application of the mining techniques in the education data helps us to deal with problems which are difficult to solve without them. The steps which are normally involved in the educational mining are defining the problem, preparing and gathering the data, building and evaluation of model and development of knowledge. Therefore, this paper studies the role of data mining in an education sector. It focuses on the importance of data mining process in educational systems and the various techniques used for increasing the effectiveness of education system in educational institutes.

Keywords: *Data Mining Process, Educational Mining, Education system,Mining Techniques*

Introduction

In present scenario, data mining is being used in government sector, businesses, research, technology etc to filter enormous data [2]. The colleges/Institutes have huge data of the students and this data can be stored in hard/soft form. It is very difficult to manage and search the student's huge data in hard form. So, the data is stored in soft form. It is helpful for the administrators to manage resources efficiently and academicians to inspect their student's learning behavior, performance and to guess their future interests. Techniques of data mining such as classification, association, decision tree and clustering can be useful to educational

systems according to the requirements. Data Mining plays a vital role in extracting patterns from the knowledge/ information. Basic concept of using data mining is that it can proficiently handle large data sets, gain sensitive knowledge and mine actionable trends [13]. Data Mining uses techniques & tools to inspect large data sets from artificial intelligence and statistics together with database systems [19]. According to the demography of the students, they are divided into groups for which classification and clustering data mining techniques are used. It results in designing specially made knowledge group that supports effective group studying.

Data Mining process is divided into various subtasks which are as follows [20]:-

1. Resource selection for mining.
2. Information gathering and it's pre-processing.
3. Pattern discovery and relationship among data by Generalization.
4. Interpreting the discovered patterns by Analyzing.

Related research done by renowned researchers is discussed in section II, in section III basics and benefits of Educational Data Mining and in section IV methodologies used in Educational Data Mining are discusses. In last section conclusion of the study is discussed.

Literature review

A lot of research work has been carried out in data mining field to handle large datasets and many techniques, tools and methods have been developed to solve these real world problems. Data Mining is being used in various applications like decision support system, bioinformatics, data warehousing etc [13].

The prime purpose of mining is to first predict the unknown variables or likely to be future variables from the given database and then illustrates the human interpretation of describing the data. Techniques mostly used in mining methods are pattern recognition, machine learning, clustering, regression, classification etc [4].

By the researchers in [16] educational mining is an efficient and influential method. Inference of mining of data in education is a continuous process through which the knowledge is extracted, which is further used in directing and improving the learning, teaching and decision making process.

Analyzing the educational data helps in studying the learning behavior of the students. Student's failure or the performance of the students in the various courses can be easily evaluated with the

help of the features like student's interests or their skills, internal assessments, presentation and communication skills [9].

The knowledge discovered is being used not only by the providers but also by the students. Thus, this application of data mining in education systems is designed to support the particular needs of students, professors, administration and also resulting in improving the performance of the system [15].Educational data mining is recognized field in its own way. Researchers have contributed a lot in analyzing a data from the educational software. Educational data mining is widely used for the up gradation of the education system [11].In educational management system Educational Data mining (EDM) is playing a significant role. EDM play a vital role in course management system also. It deals with its theoretical and practical features. In today world, for learning resources are available but it is very huge [17].

Educational data mining

Techniques of data mining can be used to find out valuable information that are used in decisive assessment to help academicians set up a instructive basis for results/decisions when modifying or designing a teaching approach.

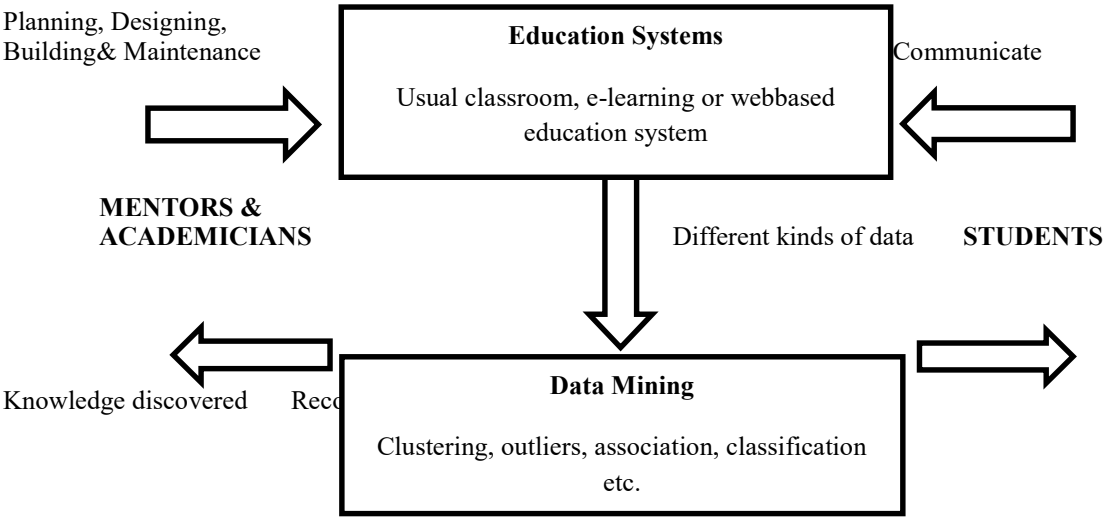

Figure-1 DataMining in Educational Systems

As in Figure-1, planning, designing, building and maintenance of the educational systems is being looked after by the mentors and the academicians. Data is turned into a mined knowledge which further results in decision making.

It is also used in preparation of usual educational process. It is helpful in the preparation and the development of the future courses which help in the curriculum designing, admission procedure etc. Various tools like Decision trees are used to find out students preferences in taking the course or to find out the number of students completing the course.

Educational Data Mining Usage

In businesses, Data Mining was used to discover patterns and relationships. It was also used in defense, medical science, banking, insurance sector, IT sector, food industry etc but now from last few years it is also being used in an education sector. There are certain benefits of mining the data in education sector

- Helps in recognizing student's requirements, courses preference by students and their specialization areas
- Helps in finding the student's performance in their related fields and their areas of interest for the future.
- Predicts the future learning behavior of students by analyzing their areas of interests and also helps in predicting the knowledge gained by the students and results of the students
- Accessibility to students profile becomes easy and records are maintained in a proper and proficient way.
- Beneficial for the management in pre-planning the future business for the profit.

Data mining methodologies

There are various data mining techniques like association, prediction, clustering etc which are being used in information technology. These techniques are used by researchers for analyzing the real world problems.

1. Prediction

Models were built to predict secondary education placements-test using sensitive analysis on predicting models. (SVM, Regression, Decision Tree Algorithm etc) [18].They has recognized various predictors of various tests, student scholarship availability, student's previous years result etc. It is also found that Decision Tree Analysis is the best predictor which is followed by SVM and neural Network.

2. Classification

Data is classified in the fixed number of groups [19] and used for categorical variables [10] usingit for categorical variables [10] is called as classification. It can be classified into two types:

Supervised Classification and Unsupervised Classification. In supervised learning cases or the objects are priorily known whereas in unsupervised learning the objects are not known in advance. Algorithms used for classification model are as follows [6].

- Statistical Analysis
- Support Vector Machines
- Genetic Algorithms
- Decision Tree/Classification
- K-nearest neighbour classifier etc

Association

According to [6] Association rule is defined as the procedure which is used to find frequent patterns and associations. It summarizes the data and also finds out the relationship among data items. Algorithms like GRI and Priori are used in association rules. Researchers show that association rules are used to find the characteristics of the students and to determine the relationship between them [16]. Association rules can be used in the education sector for opening new institutes or stating new courses depending upon certain rules.

Types of Association rules [6]:

- Multidimensional association rule
- Quantitative association rule
- Multilevel association rule

Clustering

Customer analysis is an unsupervised learning technique [20]. Clustering is grouping similar objects. Authors in[1] defines clustering as a process of grouping a set of physical or abstract object into a class of similar objects. Cluster does not classify, estimate or predict the value of target variables but segment the entire data into homogeneous subgroups. Heterogeneous population is classified into number of homogenous subgroups or clusters are referred as clustering. Furthermore, clustering task is an unsupervised classification.

Clustering in education sector can be dependent upon course seleFor example; students can be targeted after segmenting heterogeneous students into similar groups. Furthermore, clustering task in education sector can be based on enrolments, transfer, readmission,course selections, specialization, gender, behavior and age of students. Clustering describes characteristics of the students in the groups [16].

Few of the clustering techniques [8]

- Partitioning methods
- Hierarchical Method
- Density Based Method
- Partitioning Method

Conclusion

Data mining is a decision making process which is used to discover useful pattern of information.Techniques of mining like association, clustering, outlier detection, classification are used to extract knowledge from education system and manage that retrieved information in an efficient way. The algorithms of data mining are used in the education field for the improvement in the student's, teacher's and as well as in the institution's performance. Therefore, Data mining techniques are used for predicting performance of the students, income analysis by the educational institutes, marketing purposes, planning of the different courses and result analysis. Thus, it has a large array of applications in the higher educational systems.

Referenecs

[1] AmandeepKaur Mann, NavneetKaur, "Survey Paper On Clustering Techniques", International Journal of Science, Engineering and Technology Research (IJSETR) Vol 2, Issue 4, April 2013.

[2] Baker, R.S.J.D, "Data Mining for Education", in B. McGraw, P. Peterson, and E. Baker (Eds.), International Encyclopaedia of Education, Third Edition, Vol. 7, 2010, 112-118.

[3] B. L. Shivakumar, V. Murugananthan, "A Novel Application Framework for Educational Data Mining towards Automated Learning System", International Conference on Intelligent Computing Applications, IEEE, 2014.

[4] Cristóbal Romero, S. Ventura., "Data mining in course management systems: MoMoodle case study and tutorial". Computers & Education, 2009, 51(1), 368-384.

[5] Dr. MohdMaqsood Ali, "Role Of Data Mining In Education Sector", International Journal of Computer Science and Mobile Computing,Vol.2 ,Issue. 4, April- 2013, 374-383.

[6] Gorunescu, Florin, " Data Mining Techniques and Models", Data Mining; Concepts, Models and Techniques, Springer-Verlag Berlin Heidelberg, 2011,185-317.

[7] Johann Ari Larusson Brandon White, "Educational Data Mining and Learning Analytics", Learning Analytics, 2014, 906-929.

[8] Johina, VikasKamra, "A Review: Data Mining Technique Used In Education Sector", International Journal of Computer Science and Information Technologies, Vol. 6 (3), 2015, 2928-2930.

[9] K.ShanmugaPriya, A. Kumar, "Improving the Student's Performance Using Educational Data Mining", International Journal ofAdvanced Networking & Applications, 4(4), 2013, 1680-1685.

[10] Nisbet, Robert and Elder, John and Miner, Gary, "Handbooks of Statistical analysis and Data Mining Applications", Academic Press Publications, 2009.

[11] Oliver Scheuer, B.M. Maclaren, "Educational Data Mining", Encyclopedia of the Sciences of Learning, Springer, 1075-1079, 2012.

[12] PratiyushGuleria and Manu Sood, "Data Mining In Education: A Review On the Knowledge Discovery Perspective", International Journal of Data Mining & Knowledge Management Process, Vol.4, No.5, September 2014.

[13] PritamGundecha, Huan. Liu, "Mining Social Media: A Brief Introduction", Tutorials in Operation Research, 2012, pp1-17.

[14] Ryan S.J.D. Baker, K. Y., "The State of Educational Data Mining in 2009: A Review and Future Visions", Journal of E educationalData Mining, 1(1), 2009, pp3-17.

[15] S. Ventura & C. Romera., "Educational Data Mining: a Survey from 1995 to 2005", Expert Systems with Applications, Elsevier, 2007, 135-146.

[16] S. Ventura & C. Romera, Espejo, P.G. &Hervás C., "Data mining algorithms to classify students", International Conference on Educational Data Mining (EDM), Montreal, 2008, 8-17.

[17] Saddam Khan, Sunny Gupta, Yogesh Kumar Sharma, "A Study on Data Mining Techniques and Genetic Algorithm in EducationSector", International Journal of Computer Science and Mobile Computing, Vol.4 Issue.3, March- 2015, 681-683.

[18] Sen, Baha.,Ucar, Emine., Delen, Dursun. "Role of Data Mining in Education Sector", International Journal of Computer Science and Mobile Computing, Vol. 2, Issue. 4, April 2013, 374 – 383.

[19] Soman, K.P., Diwakar, Shyam and Ajay, V. "Insights into Data Mining: Theory and Practice", PHI, New Delhi, 2006.

[20] Tsai, C.F., Tsai, C.T., Hung, C.S., Hwang, P.S., "Data Mining Techniques for identifying students at risk of failing a computer proficiency test requires for graduation." Australian Journal of Educational Technology, 2011, 27(3): 481-498.