INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

AN EMPIRICAL STUDY OF APPLICATIONS OF DATA MINING IN EDUCATIONAL FIELD

MS. LEENA BHAIYA[1], PROF AJAY B GADICHA[2]

1. M.E. Scholar, Department of Computer Science & Engg, P. R. Patil COE&T, Amravati.
2. Assistant Professor, Department of Computer Science & Engg, P. R. Patil COE&T, Amravati.

Abstract: Educational institutions are playing an important role for growth and development of our nation. Few years ago, the information flow in education field was relatively simple and the application of technology was limited. Today, huge amount of data and information are available and the process of transformation of information is becoming much more difficult. Today, one of the biggest challenges that educational institutions are facing the explosive growth of educational data and to use this data to improve the quality of managerial decisions. Data mining techniques are analytical tools that can be used to extract meaningful knowledge from large data sets. Educational data mining is concerned with development of new methods for discovering knowledge from educational database. In this paper, we have discussed the applications of data mining in educational field to extract useful information from the large amount of data sets and use this information for decision making processes.

Keywords: Data Mining, Applications of Data Mining, Educational Data Mining

PAPER-QR CODE

Corresponding Author: MS. LEENA BHAIYA

Access Online On:

www.ijpret.com

How to Cite This Article:

Leena Bhaiya, IJPRET, 2016; Volume 4 (9): 904-909

904

INTRODUCTION

In this modern world, there is availability of a large amount of data which can be effectively used to produce important information. This information can be used in the field of Medical science, Education, Business, Agriculture and so on. As huge amount of data is being collected and stored in the databases. Data Mining or data or knowledge discovery has become the area of growing significance because it helps in analyzing data from different perspectives and summarizing it into useful information. Data mining is the process of discovering interesting knowledge from large amount of data stored in database, data warehouse or other information repositories. Data mining is necessary process where various intelligent techniques are applied for extraction of useful patterns.

Data mining comprises five major elements:

- Extract, convert, and load transaction data into data warehouse system.
- Storage and then management of this data in a multidimensional database system.
- Provide access of this data to information technology professionals and business analysts.
- Analysis of data using application software.
- Present the data in a useful form, such as a table or graph. [1]

Various algorithms and techniques of data mining like Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method etc., are used for knowledge discovery from databases. Each technique has its own importance according to its role. There are various applications of data mining in various fields like education, scientific and engineering, healthcare, business and many more. Mining in educational environment is called Educational Data Mining. Educational data mining is the process to discover knowledge from educational database. EDM converts raw data coming from educational systems into useful information that could potentially have a greater impact on educational research and practice. It helps in improving the performance of students and quality of education in an educational environment. Various techniques of data mining are applied on these applications to identify the best technique among them with the help of parameters like accuracy, processing time, etc.

LITERATURE REVIEW

S. Padmapriya et al[2015] has discussed many technologies available for prediction of students' performance and review that clustering is the efficient method among the existing methods

also concludes that the data mining application produces powerful and accurate results in prediction and clustering.[2]

Monika Goyal et al[2012] have studied various applications of data mining like analysis and visualization of data, grouping of students, student performance analysis and planning and scheduling etc. and these can be implemented using various algorithms of classification and clustering techniques. They discussed that DM techniques helps educational institutions in analysing student's performance, designing course curriculum to group students and to motivate students.[3]

Dina Abdulaziz Alhammad [2013] presented the implementation of student performance prediction using various algorithms (C4.5, k-means clustering, k star and naïve bayes). After implementation he has concluded that C4.5 performs better than all other algorithms. C4.5 has the highest accuracy than all other algorithms.[4]

Mrs. M.S. Mythili et al [2014] presented implementation of student performance analysis using various classification algorithms. After implementation they found that random forest performance is best than that of other algorithms used in the implementation. Random forest algorithm has highest accuracy in student performance analysis.[5]

Brijesh Kumar Baradwaj et al [2011] describes the main objective of higher education institutions is to provide quality education to its students. One way to achieve highest level of quality in higher education system is by discovering knowledge for prediction regarding enrolment of students in a particular course, detection of abnormal values in the result sheets of the students, prediction about students' performance and so on, the classification task is used to evaluate student's performance and as there are many approaches that are used for data classification, the decision tree method is used. [6]

Oyelade, O. J et al [2010] presented a method of using K-means clustering algorithm for the prediction of Students' Academic Performance. This paper is aims to present a systematic review on different clustering techniques applied for educational data mining to predict academic performance of students and its implications. [7]

Abdullah H. Wahbeh, Qasem A et al has concluded that no tool is better than the other if they are used for classification task because classification task itself is affected by the type of dataset and the way the classifier was implemented within the toolkit.[8]

Ayesha, Mustafa et al [2010] describe the use of k-means clustering algorithm to predict student's learning activities. The information generated after the implementation of data mining technique may be helpful for instructor as well as for students.[9]

S. Yadav et al [2009] obtained the university students data like attendance, class test, seminar and assignment marks from the students' previous database, to predict the performance at the end of the semester with the help of three decision trees. It was observed that C4.5 is the best algorithm.[10]

PROPOSED METHDOLOGY

Prediction of student's future learning behavior, discovering or improving the learning models, studying the effects of different strategies of teaching that learning software can provide etc. are the goals of educational data mining. Educational data mining mainly helps in studying the performance of the students based on past records. Predicting the performance of students is a challenging task because it involves extraction of large amounts of information from the universities containing student records and analyzes the progress of each student's academic progress. Clustering in this field can be mainly used to group all the students based on their similarity measures. By using clustering methods like k-means clustering, it helps in creation of groups with students having similar learning style which can be improved and can also be made faster.

K-Means Clustering Algorithm:

K-mean clustering algorithm, clusters are fully dependent on the selection of the initial cluster centroids. K data elements are selected as initial centers and then the distances of all data elements are calculated by Euclidean distance formula. Data elements having less distance to centroids are moved to the appropriate cluster. The process is continued until no more changes occur in clusters. The following are the steps of k-means clustering algorithm :

Step 1: Accept the number of clusters to group data into and the dataset to cluster as input values

Step 2: Initialize the first K clusters

- Take first k instances or

- Take Random sampling of k elements

Step 3: Calculate the arithmetic means of each cluster formed in the dataset.

Step 4: K-means assigns each record in the dataset to only one of the initial clusters

- Each record is assigned to the nearest cluster using a measure of distance (e.g Euclidean distance).

Step 5: K-means re-assigns each record in the dataset to the most similar cluster and re-calculates the arithmetic mean of all the clusters in the dataset.

CONCLUSION:

Since the application of data mining brings a lot of advantages in the educational field. There are various algorithms of classification and clustering techniques that are used in educational data mining. In this paper we have discussed clustering technique of data mining which can support education system via generating strategic information in the areas like predict the performance of student, optimization of resources, prediction of retainment of faculties in the university, to find the gap between the number of candidates applied for the post, number of applicants responded, number of applicants appeared, selected and finally joined and can also help to improve the accuracy and processing time which is the problem that other algorithms are x facing.

REFERENCES

1. Hardeep Kaur. A Review of Applications of Data Mining in the Field of Education. International Journal of Advanced Research in Computer and Communication Engineering.

2. S. Padmapriya, Dr. L. Jayasimman, Dr. Nisha Jebaseeli, B. Senthil Kumar, Mining Educational Data Using Data Mining Techniques And Algorithms –A Review, International Research Journal of Engineering and Technology.

3. Monika Goyal, Rajan Vohra. Applications of Data Mining in Higher Education, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.

4. Dina Abdulaziz Alhammadi, "Data Mining in Education- An Experimental Study", International Journal of Computer Applications Volume 62, No.15, January 2013.

5. Mrs. M.S. Mythili, Dr. A.R.Mohamed Shanavas, "An Analysis of students' performance using classification algorithms", IOSR Journal of Computer Engineering, Volume 16, Issue 1, January 2014.

6. Brijesh Kumar Baradwaj, Saurabh Pal, Mining Educational Data to Analyze Students" Performance, International Journal of Advanced Computer Science and Applications.

7. Oyelade, O. J, "Application of k-Means Clustering algorithm for prediction of Students' Academic Performance", (IJCSIS) International Journal of Computer Science and Information Security, Vol.7.

8. Abdullah H. Wahbeh, Qasem A. Al-Radaideh, Mohammed N. Al-Kabi, and Emad M. Al-Shawakfa. A Comparison Study between Data Mining Tools over some Classification Methods International Journal of Advanced Computer Science and Applications.

9. Shaeela Ayesha, Tasleem Mustafa, Ahsan Raza Sattar, M. Inayat Khan, "Data mining model for higher education system", Europen Journal of Scientific Research, Vol.43, No.1, pp.24-29, 2010.

10. S. K. Yadav, B.K. Bharadwaj and S. Pal, "Data Mining Applications: A comparative study for predecting students' performance", International Journal of Innovative Technology and Creative Engineering (IJITCE), Vol 1, No. 12, ISSN: 2045-8711, 2011.

11. Neha D. B. M. Vidyavathi," A Survey on Applications of Data Mining using Clustering Techniques, " International Journal of Computer Applications (0975 – 8887) Volume 126 – No.2, September 2015.