Research Article

Descriptive And Predictive Data Mining Techniques To Improve Student Academics And Employability

B.Dhanalaxmi

Abstract

In current trends, data mining is the most important domain in the real world aspects. By using data mining Techniques, we can identify the knowledge of different areas and get the best patterns. Data Mining Techniques are used in various applications i.e., Health care, Customer Relationship Management, Market Prediction, Fraud Detection…, One of the important applications where data mining used is in Education. In the present knowledge-based era, education plays a major role in the progress of a nation's economy and development. Thus, the research on development in education is an important work and is actually required. Institutions are applying data mining technologies on the huge data generated in class room including academic, behavioral, demographic data of students and faculty data as well to find out the useful patterns and fill the gap between the student academics and employability

Introduction

Data Mining is the important process to identify correct data from huge amount of data and selecting appropriate data based upon the user requirement. The best example is our daily real-life i.e when we attend a party huge number of food items available from which we select our choice food item. The most important task in data mining is to extract necessary data from large amounts of data. Below figure indicates the various steps involved in date mining to generate knowledge

Data Mining Process

The main aim of Data mining is to find correct pattern from huge data. Like in Google Search we search for a particular topic from huge number of hyper links which is displayed, among which we select the best hyper link. Data mining goals are classified into two categories: descriptive and predictive.

---

[1]Assistant Professor, St.Martins Engineering College Komapally Hyderabad Contact Number 9705081326 Mail id bdhanalaxmicse@smec.ac.in

Descriptive Mining goal is to divide the large amount of data into small so that it is easy to identify the correct pattern. The data mining methods used for this are Association Rule and Clustering

Association Rule

Association rule is simple a procedure to find frequent patterns, correlations and association from data sets which is found in different databases available.

Clustering

Clustering is a method to group similar objects under one Domain where Domain is a collection of similar objects. Clustering is defined as a set of datum, each having a set of attributes and similarity measure. Each datum in same cluster is similar to one another where as datum in separate clusters are less similar to one another.

Euclidean distance formula is used find the how much difference is there between one cluster with another cluster

Predictive Mining goal is to identify the future problem based upon on the current data. The data mining methods used for this are Regression and Classification

Regression

Regression is used to predict a range of numeric values on a given dataset. We call this method in simple words also "predictive power." When we use a regression analysis we want to predict the value of a given (continuous) feature based on the values of other features in the data, assuming a linear or nonlinear model of dependency

Classification

The main goal of Classification in data mining is to assigns items in a collection or categories or classes. Classification which exactly predict the target class for each case in the dataset. For example in Educational Department where each student records contains a set of attributes, i.e., Roll No, Student Name, Father Name and class from this data we identify the class

Need For Study/Research

At present Educational institutions and universities are facing problems in terms of student employability. It became a big task for the educational institutions to resolve these problems. To resolve these problems some data mining techniques are identified to resolve the gap between the student academics and faculty Employability. To remove the gap which includes the process of analyzing the students' details using different attributes such as students' name, roll number, previous semester marks, attendance, assignment, seminar performance, lab work and gender are used to evaluate the students' performance (Pass / Reappear).

The various prediction algorithms like Classification, decision tree algorithm, C4.5, Feature with Graph structure, Bayesian, RIPPER, and SVM. The above algorithms compares the best performances of students based on the criteria .Among these techniques, classification which is used for educational data mining. The classification process is based on C5.0 algorithm with good

classification accuracy. The study helps to the learners as well as to the teachers for the academic performance evaluation. It is a caution system for the students' to improve their study performance

EDUCATIONAL DATA MINING (EDM) FLOW CHART
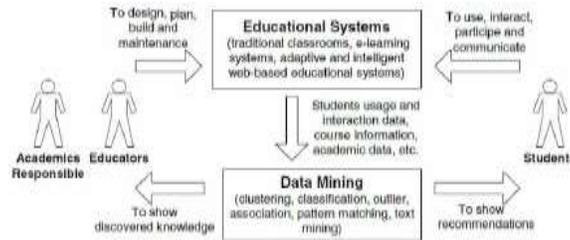
Research Methodology

This paper concentrates on various aspects to increase student performance

- Predicting students' future learning behavior with the use of student modeling which includes detailed information such as their knowledge, behaviors and motivation to learn.
- By discovering or improving student domain models through the various methods and applications of EDM. Discovery of new or existing include illustrating the educational content to engage learners and regulate instructional sequences to support the student's learning style.
- Studying the effects of educational support that can be achieved through learning systems.
- By using Advance scientific knowledge about learning and learners by building and incorporating student models that uses software also.

With the help of systematic review on Educational data mining (EDM) helps in growth of academic performance and increase

According to the literature study, EDM research pertains mainly on three

- According to the way data is collected
- Traditional face to face or the offline education system based on data generated in the classroom.
- E-learning in which the learning is provided through online content based on online activity logs.

TOOLS FOR ANALYSIS

Conventional educational refers to class room teaching which is mostly used by many institutions throughout the world. As, students takes admission in a better Institution with high academic performance and hope for better employability after completion.

In Descriptive data mining the two methods used are:-

- An association technique that uses Apriori algorithms to characterize student performance to discover co-relations among set of items. The Apriori algorithm is applied to database

containing academic records of various students and try to extract association rules in order to profile students based on various parameters like exam scores, term work grades, attendance and practical.

- The Clustering algorithm (k-means) used, for grouping students which assigns a set of observations into subsets

Educational data mining is very useful in discovering valuable information which can be used to characterize student performance based on their academic record.

In Predictive data mining the two methods used are:-

- Regression method is used to evaluate student academic performance for getting good employability. This method help the management identify students of poor academic performance .The process of prediction involves application of various data mining algorithms, to predict the dependent variables based on independent factors.
- Classification method which predicts performance of student's .A decision tree algorithm has been used in the research for classifying students according to their attributes (grades). The different decision tree algorithms are J48, NBtree, Reptree and Simple cart.

The conventional KDD process has been used as a methodology. The WEKA (Waikato Environment for Knowledge Analysis) tool was used for analysis and prediction

Implications of the study

This paper mainly focus on student career and employability by improving standards in teaching .Various Data Mining Techniques in Education helps every institution to improve and discover new teaching process so that the student academic standard will increase which indirectly helps the student to get good job after their successful completion of their course

References

1. E. Osmanbegovic, M. Suljic, Data mining approach for predicting student performance, Economic Review
2. Tan P. N., Steinbcah M and Kumar V. 2006 Introduction to Data mining, Introduction to Data mining(MA: Pearson Education)
3. Bhardwaj B. K. and Pal S. 2011 Data Mining: A prediction for performance improvement using classification (IJCSIS) International Journal of Computer Science and Information.
4. Suman, P.Mittal, Comparison and Analysis of Various Clustering Methods in Data mining On Education Data set using the Weka tool, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 3(2), 2014, 240–244.
5. K. R. Kavyashree, LakshmiDurga – "A Review on Mining Students Data for Performance Prediction" - International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 4, April 2016 [2] TriptiDwivedi, Diwak