A FRAMEWORK FOR DESIGNING EDUCATIONAL DATA MINING SYSTEM

ADITYA KUMAR
MBA BANKING TECHNOLOGY
CENTRAL UNIVERSITY OF PONDICHERRY

ABSTRACT
The aim of this paper is to propose a framework for designing an educational data mining system in order to improve the overall teaching-learning methodology.

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

Moodle is an open source virtual learning environment to help educators create online courses with a focus on interaction and collaborative construction of content.

All the teachers and students are provided with the login ids to access their account in Moodle, which is installed in the department intranet. As the students access the account, log file of the particular student is generated giving details of the student.

On the generated data, we can apply various data mining techniques, out of which we have focused on Clustering and Classification. Clustering is a descriptive data mining technique which partitions data into meaningful subclasses, called clusters. Classification is a predictive data mining technique where data objects are classified into predefined classes.

Weka is a freely available specialized data mining tool. It has a comprehensive collection of data preprocessing and mining techniques.

We have used Hierarchical Agglomerative clustering and C4.5 classification algorithms to analyze educational data in Weka. The obtained results are used for predicting the academic performance of students, so that the teaching process can be modified accordingly.

KEYWORDS: Educational data mining, Course Management System, Data preprocessing, Clustering, Classification

REFERENCES


