IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 47-51

Web-Based Data Mining Tools: Performing Feedback Analysis And Association Rule Mining

Shital Gadekar, Prerna Naiskar, Prof. Vijaya Kamble, Monika Dhoble, Nancy Wahane

Department of Computer Science and Engineering Guru Nanak Institute of Engineering and Technology,Dahegaon Nagpur,Maharashtra,India
Department of Computer Science and Engineering Guru Nanak Institute of Engineering and Technology,Dahegaon Nagpur,Maharashtra,India

Abstract: *This paper aims to explain the web-enabled tools for educational data mining. The proposed web-based tool developed using Asp.Net framework and php can be helpful for universities or institutions providing the students with elective courses as well improving academic activities based on feedback collected from students. In Asp.Net tool, association rule mining using Apriori algorithm is used whereas in php based Feedback Analytical Tool, feedback related to faculty and institutional infrastructure is collected from students and based on that Feedback it shows performance of faculty and institution. Using that data, it helps management to improve in-house training skills and gains knowledge about educational trends which is to be followed by faculty to improve the effectiveness of the course and teaching skills.*

Keywords: *Academic, Apriori, Asp.Net, Feedback, Knowledge, Mining, Php*

I. Introduction

With competitive environment prevailing among the educational institutions, the main objective of higher education institutes is to disseminate quality education to its students and to improve the quality of managerial decisions. Quality of education can be improved by gaining knowledge from educational data which facilitate academic planners in higher education institutes to enhance their decision making process, to improve students' academic performance and better understand students' behavior, to assist instructors, to improve teaching and many other benefits and for achieving this data mining plays an important role. Data Mining is the process of applying intelligent methods to extract data patterns. It is a powerful analytical tool that enables educational institutions to better allocate resources and staff, manage student feedback [1].Applying data mining techniques to educational data for knowledge discovery is significant to educational organizations as well as students. Educational data mining is an emerging trend, concerned with developing techniques for exploring, and analyzing the huge data that come from the educational context and using those methods to better understand student interests [2][3].EDM consists of Web-based data mining software tools which helps in managing knowledge to gain competitive advantage, information dissemination and knowledge acquisition, analytic capabilities and enhances our understanding of learning process to focus on identifying, extracting and evaluating variables related to the learning process of students[4].The key advantages of these web-based tools is that they are helpful in future analysis, decision-making and feedback based on dynamic data[5].These web-based data mining tools provide management of the institution with information to improve educational trends. The Web-based applications and data-mining environments are becoming important media for collecting, sharing and distributing information [6].There are learning content management system platforms which facilitate information sharing and communication between participants in a course where massive online courses are provided to students where they can opt courses accordingly, give online tests, prepare assignments, engage in discussions, chats etc [7].

II. Classification Of Existing System

Learning Management Systems (LMS) are being installed more and more by universities, community colleges, schools, businesses, and even individual instructors in order to add web technology to their courses and to supplement traditional face-to-face courses [8]. LMS systems accumulate a vast amount of information which is very valuable for analyzing the students' behaviour and could create a gold mine of educational data [9]. They can record whatever student activities it involves, such as reading, writing, taking tests, performing various tasks, and even communicating with peers. However, due to the vast quantities of data these systems can generate daily, it is very difficult to analyze this data manually. A very promising approach towards this analysis objective is the use of data mining techniques.

Data mining or knowledge discovery in databases (KDD) is the automatic extraction of implicit and interesting patterns from large data collections [10]. Association rules mining is one of the most well studied data mining tasks. It discovers relationships among attributes in databases, producing if-then statements concerning attribute-values [4]. An association rule $X \Rightarrow Y$ expresses that in those transactions in the database where X occurs; there is a high probability of having Y as well. X and Y are called respectively the antecedent and consequent of the rule. The strength of such a rule is measured by its support and confidence. The confidence of the rule is the percentage of transactions with X in the database that contain the consequent Y also. The support of the rule is the percentage of transactions in the database that contain both the antecedent and the consequent.

Association rule mining has been applied to e-learning systems for traditionally association analysis (finding correlations between items in a dataset), including, e.g., the following tasks: building recommender agents for on-line learning activities or shortcuts [5], automatically guiding the learner's activities and intelligently generate and recommend learning materials, identifying attributes characterizing patterns of performance disparity between various groups of students, discovering interesting relationships from student's usage information in order to provide feedback to course author [8], finding out the relationships between each pattern of learner's behaviour, finding students' mistakes that are often occurring together, guiding the search for best fitting transfer model of student learning, optimizing the content of an e-learning portal by determining the content of most interest to the user, extracting useful patterns to help educators and web masters evaluating and interpreting on-line course activities, and personalizing e-learning based on aggregate usage profiles and a domain ontology.

Association rule mining also has been applied to the learning of sequential patterns mining, which is a restrictive form of association rule mining in the sense that not only the occurrences themselves, but also the order between the occurrences of the items is taken into account. The extraction of sequential patterns has been used in e-learning for evaluating the learners' activities and can be used in adapting and customizing resource delivery; discovering and comparison with expected behavioural patterns specified by the instructor that describes an ideal learning path; giving an indication of how to best organize the educational web space and be able to make suggestions to learners who share similar characteristics; generating personalized activities to different groups of learners; supporting the evaluation and validation of learning site designs; identifying interaction sequences indicative of problems and patterns that are markers of success.

Finally, association rule mining has been used in the e-learning for classification. From a syntactic point of view, the main difference to general association rules is that classification rules have a single condition in the consequent which is the class identifier name. They have been applied in learning material organization, student learning assessments, course adaptation to the students' behaviour and evaluation of educational web sites.

The general KDD process has the next steps: collecting data, pre-processing, applying the actual data mining tasks and post-processing. We particularize these steps for association rule mining in the LMS domain.

Collecting data. Most of the current LMSs do not store logs as text files. Instead, they normally use a relational database that stores all the systems information: personal information of the users (profile), academic results, the user's interaction data, etc. Databases are more powerful, flexible and bug-prone than the typically textual log files for gathering detailed access and high-level usage information from all the services available in the LMS. The LMSs keep detailed logs of all activities that students perform. Not only every click that students make for navigational purposes (low level information) is stored, but also test scores, elapsed time, etc. (high level information).

Data pre-processing. Most of the traditional data pre-processing tasks, such as data cleaning, user identification, session identification, transaction identification, data transformation and enrichment, data integration and data reduction are not necessary in LMS. Data pre-processing of LMS data is simpler due to the fact that most LMS store the data for analysis purposes, in contrast to the typically observational datasets in data mining, that were generated to support the operational setting and not for analysis in the first place. LMSs also employ a database and user authentication (password protection) which allows identifying the users in the logs. Some typical tasks of the data preparation phase are: data discretization (numerical values are transformed to categorical values), derivation of new attributes and selection of attributes (new attributes are created from the existed ones and only a subset of relevant attributes are chosen), creating summarization tables (these tables integrate all the desired information to be mined at an appropriate level, e.g. student), transforming the data format (to format required by the used data mining algorithms or frameworks).

Applying the mining algorithms. In this step it is necessary: 1) to choose the specific association rule mining algorithm and implementation; 2) to configure the parameters of the algorithm, such as support and confidence threshold and others; 3) to identify which table or data file will be used for the mining; 4) and to specify some other restrictions, such as the maximum number of items and what specific attributes can be present in the antecedent or consequent of the discovered rules.

Data post-processing. The obtained results or rules are interpreted, evaluated and used by the teacher for further actions. The final objective is to putting the results into use. Teachers use the discovered information (in form of if-then rules) for making decisions about the students and the LMS activities of the course in order to improve the students' learning. So, data mining algorithms have to express the output in a comprehensible format by e.g., using standardized e-learning metadata.

2.1 Web enabled Association Rule Mining tool

This tool proposed has used an SQL query mechanism for querying the discovered knowledge in the form of association rules. ARM techniques are applied to databases of transactions where each transaction consists of a set of items [11]. The suggested web-based tool, developed using ASP.NET framework, can be helpful for universities or institutions providing the students with elective courses. It can be utilized to:

• Generate the combinations of elective courses most opted on the basis of feedback of students.

• Generate the combinations of elective courses best recommended on the basis of feedback from industry experts.

User can develop web forms using ASP.Net and follow the different steps:

i) Data Selection:

The data is input from the students using a web-form.

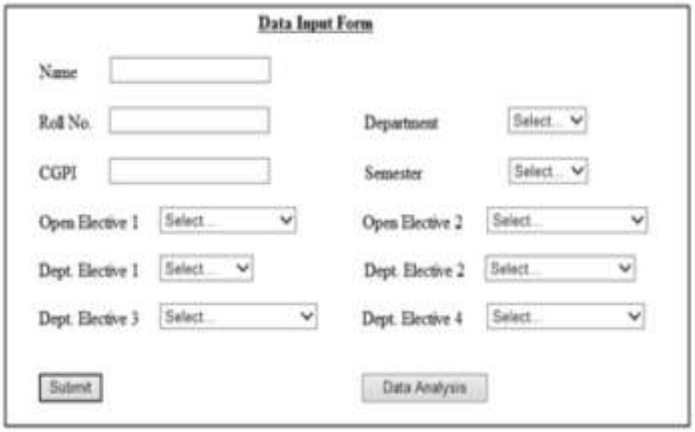

Fig 1: Input Form for students to select Open Elective and Department Elective Courses.

ii) *Database Structure*:

The structure of the database being used to store transactions.

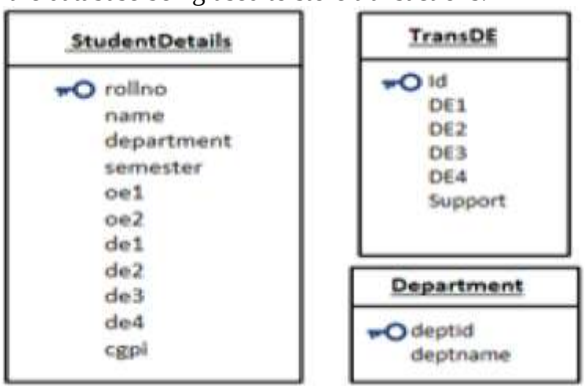

Fig 2. Design of database structure

iii) *Purpose of Tables:*

Different database tables are used in this tool.

Table 1: Description of tables

Table Name	Description
Student Details	To store information regarding the students.
DeptElectives	It contains the list of all departmental electives being offered to the students.
OpenElectives	It contains the list of all open electives being offered to the students.
Departments	It contains list of all the departments.
TransOE	It contains all the transactions of open electives which are to be analyzed.
TransDE	It contains all the transactions of departmental electives which are to be analyzed.

iv) *Sample dataset:*

Sample Dataset containing transactions for implementing Mining.

Open Elective1	Open Elective 2	Deptt Elective 1	Deptt Elective 2	Deptt Elective 3	Deptt Elective 4
Digital Electronics	Communication	Artificial Intelligence	Mobile DB	Adv.Microprocessor	Embedded Systems
IBE	Industrial Env.	Web Tech.	Distributed Systems	Adv. Microprocessor	Embedded Systems
IBE	Industrial Env.	Web Tech.	Distributed Systems	Parallel Algo.	Adv.Comp.Networls
IBE	Nuclear Phy.	Artificial Intelligence	Mobile DB	Parallel Algo.	Adv.Comp.Network
Digital Electronics	Computer Graphics	Artificial Intelligence	Distributed Systems	Adv.Microprocessor	Embedded Systems
Robotics_1	Neural Networks	Artificial Intelligence	Distributed Systems	Adv.Microprocessor	Embedded Systems
Optimization	Nuclear Phy.	Web Tech.	Mobile DB	Parallel Algo.	Adv.Comp.Network
RDBMS	Robotics_2	Web Tech	Distributed Systems	Parallel Algo	Adv.Comp.Network

III. Proposed System

One of the most useful data mining techniques for e-learning is classification. Classification is the processing of finding a set of models which describe and distinguish data classes or concepts. The derived model may be represented in various forms, such as classification (IF-THEN) rules, decision trees, mathematical formulae, or neural networks. We implemented feedback system developed and Apriori application in Asp.Net taking sample of student's data.

Using Association Rule Mining, we find association rules with support and confidence for optional and elective courses mostly opted by students. These methods have been applied to web-based educational systems where associations discovered shows which contents students tend to access together, or which combination of subjects they want to opt.

IV. Classification Using Web Based Tools

One of the most useful data mining techniques for e-learning is classification. Classification is the processing of finding a set of models which describe and distinguish data classes or concepts. The derived model may be represented in various forms, such as classification (IF-THEN) rules, decision trees, mathematical formulae, or neural networks [8]. We implemented feedback system developed in Php and Apriori application in Asp.Net taking sample of student's data. Using Association Rule Mining, we find association rules with support and confidence for optional and elective courses mostly opted by students. These methods have been applied to web-based educational systems where associations discovered shows which contents students tend to access together, or which combination of subjects they want to opt.

V. Conclusion

With the growth in Web-based applications which includes massive open online courses, e-learning, there is significant increase in analyzing educational trends through web-based tools and apply that knowledge

to better serve users, students. Web-Based Data mining tools have shown significant results in educational sector, in-depth analysis of data and results generated.

Reference

[1]. Luan Jing (2001). Data Mining Applications in Higher Education. A chapter in the upcoming New Directions for Institutional Research. Josse-Bass. San Francisco.

[2]. Sachin R.B, Vijay M.S (2012). A Survey and Future Vision of Data Mining in Educational Field. pp.96 – 100.ISBN:978-1-4673-0471-9

[3]. Goyal Monika, Vohra Rajan (2012). Applications of Data Mining in Higher Education. IJCSI. Vol. 9, Issue 2, No 1. ISSN (Online): 1694-0814

[4]. Bhardwaj Brijesh Kumar, Pal Saurabh (2011) Mining Educational Data to Analyze Students Performance. International Journal of Advanced Computer Science and Applications. Vol. 2, No. 6.

[5]. Heinrichs H. John, Lim Su-Jeen (2003). Integrating web-based data mining tools with business models for knowledge management. Decision Support Systems 35,103– 112.

[6]. Zaiane O. R., Luo J. (2001). Towards evaluating learners' behaviour in a web-based distance learning environment. In Advanced Learning Technologies, 2001. Proceedings. IEEE International Conference on (pp. 357-360). IEEE.

[7]. Sarma Deva Kumar Pankaj, Roy Rahul (2010) A Data Warehouse for Mining Usage Pattern in Library Transaction Data. Assam University Journal of Science &Technology: Physical Sciences and Technology Vol. 6 Number II, 125-129.

[8]. Rice, W.H.: Moodle E-learning Course Development. A complete guide to successful learning using Moodle. Packt publishing (2006).

[9]. Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., Heiner, C.: An educational data mining tool to browse tutor-student interactions: Time will tell! In: Proc. of the Workshop on Educational Data Mining (2005) 15–22.

[10]. Klosgen, W., & Zytkow, J.: Handbook of data mining and knowledge discovery. Oxford University Press, New York (2002).

[11]. Hamalainen, W., Vinni, M.: Comparison of machine learning methods for intelligent tutoring systems. In: Proc. of Int. Conf. in Intelligent Tutoring Systems (2006) 525–534.