Analysis of Student Motivation Behavior on e-Learning Based on Association rule mining

Kunyanuth Kularbphettong, Phanu Waraporn, Cholticha Tongsiri

Abstract—This research aims to create a model for analysis of student motivation behavior on e-Learning based on association rule mining techniques in case of the Information Technology for Communication and Learning Course at Suan Sunandha Rajabhat University. The model was created under association rules, one of the data mining techniques with minimum confidence. The results showed that the student motivation behavior model by using association rule technique can indicate the important variables that influence the student motivation behavior on e-Learning.

Keywords—Motivation Behavior, e-Learning, Moodle log, association rule mining

I. INTRODUCTION

Nowadays, the web-based educational system, with no longer barrier by space and time, has been increasingly used as a significant tool to support students and teachers. The benefits of the system are to facilitate information sharing and collaboration and to communicate between student and teacher in a course. Student can take a web-based class to enhance their knowledge and understanding and teacher can easily monitor student’s performance as well. Moodle is a well known Learning Management System (LMS) that educator able to create the effective online courses. However, it accumulates a huge of information daily which in turn is very valuable for analyzing student’s pattern behavior [1]. Student’s logs in a Moodle can show students’ interactions like reading, writing, taking exam, and doing various tasks [2]. Therefore, it is very difficult to analyze this data manually and although there are some tools that help to report useful information, they do not offer specific features teacher need to track and evaluate all the students’ activities in class [3]. Data Mining Techniques is the promising methodology to extract valuable information in this objective. Data Mining can analyze relevant information results and produce different perspectives to understand more about the students’ activities so as to customize the course for student learning.

The paper presents a prediction model of student motivation behavior by using Moodle in the Information Technology for Communication and Learning Course at Suan Sunandha Rajabhat University. The remainder of this paper is organized as follows. Section 2 reviews about the related literatures and the related methodologies used in this work. Section 3 presents the implementation based on the purposed data mining techniques.

In section 4 the result and discussion is presented. Finally, we conclude the paper with future research issues in section 5.

II. RELATES WORKS AND THE METHODOLOGIES

In this section, we illustrate the literature search and the specified methodologies used in this project.

A. Relates Works

A literature search shows that most of the related researches have deployed data mining techniques to analyze student’s learning behaviors by following this: According to C. Romero at el [4], the research was shown the usefulness of the data mining techniques in course management system and the rules can help to classify students and to detect the sources of any incongruous values received from student activities. Data mining techniques like association rule mining were applied in [5],[6] to extract the patterns and to evaluate the activities of on line course and classification and association rule mining algorithms are discussed and demonstrated in [7]. Also there are many researches that have been investigated in the on-line learning environment. For example, West et al investigated impact of learning style on e-learning by using Statistics [8] and Kerdpasop et al used Rule induction rough set to Classify student knowledge level[9].

B. The Methodologies

Data Mining is the data analyzing process from different perspectives also summarizing the useful information results. The data mining process uses many principles as machine learning, statistics and visualization techniques to discover and present knowledge in an easily comprehensible form. There is another definition as “the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data” [10], [11].

Association rule mining is one of the most popular data mining approaches. It is used to discover interesting relationships between variables in databases. According to Agrawal et al [12], an association rule explains a close correlation between items in a database in the form of $x \Rightarrow y$ where $x$ and $y$ are sets of Item set $I$ ($x \subset I$ and $x \cap y = \phi$). $I = I_1, I_2, \ldots, I_m$ is an Item set of m distinct attributes. The rule is indicated $x \Rightarrow y$ whereby $x$ is called antecedent and $y$ is called consequent. There are two importance thresholds for measurement association rule mining, minimum support and minimum confidence. The support of a rule $x \Rightarrow y$ is the probability of the Item set $x, y$ that means the relevance of the rule and the confidence of a rule $x \Rightarrow y$ is the conditional probability of $y$ given $x$ that indicate the accuracy of the rule.

\[
\text{Confidence (} x \Rightarrow y \text{)} = \frac{\text{support}(x, y)}{\text{support}(x)} \tag{1}
\]

and
Support \(( x \Rightarrow y) = \frac{\text{support}({x, y})}{\text{Total number of transaction in } D} \) 

Set of transactions: \( D = \{d_1, d_2, \ldots, d_n\} \) each \( d_i \subseteq I \)

Hence, confidence is a significant measure of the association rules to indicate how strong the mined rules. If the confidence of the association rule \( x \Rightarrow y \) is 80%, it means that 80% of the transactions that contain \( x \) also contain \( y \), based on users to indicate the specified minimum confidence [13].

Apriori Algorithm is an influential algorithm for association rule mining, proposed by proposed by Agrawal [14] as shown in Fig.1. The Apriori Algorithm is used level-wise search for frequent item sets, the sets of items that have minimum support.

Let \( C_k \) is Candidate itemsets of size \( k \) and \( L_k \) is itemsets of size \( k \).

1. \( L_1 = \{\text{large 1-item sets}\} \)
2. for \((k = 2; \ L_{k-1} \neq \emptyset ; k++\) \) do begin
3. \( C_k = \text{apriori-gen}(L_{k-1}); \) // New candidates
4. for all transactions \( t \in D \) do begin
5. \( \forall c \in C_k \) do begin
6. \( c \cdot \text{count}++; \) end
7. end
8. \( L_k = \{c \in C_k | c \cdot \text{count} \geq \text{minsup}\} \)
9. end
10. \( \text{Answer} = \bigcup L_k; \)

![Fig. 1 Apriori Algorithm [14]](image)

### III. EXPERIMENTAL SETUP

In our research, we collected the student’s data from the Information Technology for Communication and Learning Course at Suan Sunandha Rajabhat University, in the two semesters of 2011. The number of students was 3,002. The data is composed of personal records, course (face-to-face) records and students’ log file from e-Learning system. Moodle has been used for this course. Moodle is an well known open source software for learning management system that offers teacher create and manage online classes effectively. From this Course, in the class (face-to-face) room, student must be required to attend in course room, to do exercises, to take post exams in class, and to work in group for doing project. Also, in e-Learning class, student needs to take pre and post quizzes online, to review and use materials on e-Learning system and to participate in exercises as shown in Fig.2 (a) and (b).

Also, the results of students’ grade are collected in the last. From Fig.3, the data is preprocessed, and transformed to be appropriated format in order to apply data mining techniques to discover association rules.

The table I presents the important information of this research and then all continuous attributes have been transformed to nominal attributes so as to conveniently discover the rules and to easily understand. There are various methodologies to transform numerical attributes to discrete attributes like equal width, equal frequency, clustering principles and etc. In our case, we used the equal width method to partition the value of continuous attributes into five nominal values: VERY LOW, LOW, MEDIUM, HIGH and VERY HIGH.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreTest_number</td>
<td>Identification number of pretest</td>
</tr>
<tr>
<td>PostTest_number</td>
<td>Identification number of posttest</td>
</tr>
<tr>
<td>Result PreTest_number</td>
<td>Mark obtained from pretest</td>
</tr>
<tr>
<td>Result PostTest_number</td>
<td>Mark obtained from posttest</td>
</tr>
<tr>
<td>Time-of-PreTest_number</td>
<td>Time spent on pretest</td>
</tr>
<tr>
<td>Time-of-PostTest_number</td>
<td>Time spent on posttest</td>
</tr>
<tr>
<td>SumPreTest</td>
<td>Total mark obtained from all pretest</td>
</tr>
<tr>
<td>SumPostTest</td>
<td>Total mark obtained from all posttest</td>
</tr>
<tr>
<td>Time-of-viewing–material</td>
<td>Total time spent on viewing material</td>
</tr>
<tr>
<td>Number-of-Attendance</td>
<td>Number of Attendance in class</td>
</tr>
<tr>
<td>Assignment_Number</td>
<td>Identification number of assignment</td>
</tr>
<tr>
<td>Result_Assignment_Number</td>
<td>Mark obtained from assignment</td>
</tr>
<tr>
<td>InClassTest_number</td>
<td>Identification number of in class test</td>
</tr>
<tr>
<td>Result_InClassTest_number</td>
<td>Mark obtained from in class test</td>
</tr>
<tr>
<td>Project_score</td>
<td>Mark obtained from project</td>
</tr>
<tr>
<td>Midterm-score</td>
<td>Mark obtained from Midterm</td>
</tr>
<tr>
<td>Final-score</td>
<td>Mark obtained from Final</td>
</tr>
<tr>
<td>Grade</td>
<td>Final mark obtain from this class</td>
</tr>
</tbody>
</table>

![Fig. 2(a) Students’ log Moodle](image)

![Fig. 2(b) Students’ log Moodle](image)

![Fig. 3 the data preparation process](image)
After preparation data, we apply the association rule mining algorithm to discover valuable patterns. Data was analyzed by WEKA. WEKA, the Waikato Environment for Knowledge Analysis, is a collection of machine learning algorithms to analyze data set for data mining tasks [15]. Association rule mining is very useful for educational objectives because it presents significant relationships between the activities of students on this class and their final scores [16]. Apriori algorithm has been used for this research and evaluated with a minimum support of 0.2 and a minimum confidence of 0.9. Fig 4 shows examples from the results of Apriori algorithm.

**IV. RESULTS OF EXPERIMENTAL**

Due to a huge number of discovered rules, there are many unnecessary rules teachers are not interested in. However, there are also many interesting rules teachers can use for enhancing their classes. Hence, for this research, the selected rules are explained in the Fig.5, based on decision of teacher.

![Fig. 5 Results of selected rules](image)

**REFERENCES**


Kunyanuth Kularbphettong received the B.S. degree in Computer Business, M.S. degree in Computer Science, and Ph.D degree in Information Technology. Her current research interests are in Multi-agent System, Web Services, Semantic Web Services, Ontology and Data mining techniques.

Phanu Waraporn received M.S. degree in Computer Information Systems and currently a doctoral student in Computer Engineering. His current research interests are in the area of Ontology based medical decision making.

Cholticha Tongsiri is interested in Web-Based Educational Learning, and Data mining techniques.