

# 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.

Kunyanuth Kularbphettong is with Computer Science Program, Suan Sunandha Rajabhat University, Bangkok, Thailand (phone: 662-150-1169; e-mail: kobkulriss@gmail.com).

Phanu Waraporn Computer Science Program, Suan Sunandha Rajabhat University, Bangkok, Thailand (phone: 662-150-1169; e-mail: phanu.waraporn@gmail.com).

Cholticha Tongsiri is Faculty of Information and Communication Technology, Silpakorn University, Bangkok, Thailand (phone: 662-233-4995; e-mail: c\_tongsiri@hotmail.com).

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 Kerdprasop 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  $I$  ( $x$  and  $y \subset I$ ) and  $x \cap y = \emptyset$ .  $I = I_1, I_2, \dots, I_m$  is a Item set of  $m$  distinct attributes. The rule is indicated  $x$  implies  $y$  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.

$$\text{Confidence}(x \Rightarrow y) = \frac{\text{support}(\{x, y\})}{\text{support}(x)} \quad (1)$$

and

Support (x => y) = support({x, y}) / Total number of transaction in D (2)

Set of transactions: D = {d1, d2, ..., dn} each di ⊆ I

Hence, confidence is a significant measure of the association rules to indicate how to strength of the mined rules. If the confidence of the association rule x => 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 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 Ck is Candidate itemsets of size k and Lk is itemsets of size k.

- 1) L1 = {large 1-itemsets};
2) for (k = 2; Lk-1 ≠ ∅; k++) do begin
3) Ck = apriori-gen(Lk-1); // New candidates
4) forall transactions t ∈ D do begin
5) Ct = subset(Ck, t); // Candidates contained in t
6) forall candidates c ∈ Ct do
7) c.count++;
8) end
9) Lk = {c ∈ Ck | c.count ≥ minsup}
10) end
11) Answer = ∪k Lk;

Fig. 1 Apriori Algorithm [14]

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).

Table with 6 columns: รหัสวิชา, เวลา, หมายเลขข้อสอบ, ชื่อข้อสอบ, คำตอบที่, ผลการตรวจ. It contains log data for various assignments and quizzes.

Fig. 2 (a) Students' log Moodle

Table with 10 columns: ชื่อ, เริ่มต้น, หมดเวลา, วิชาที่, คะแนน/10, #1, #2, #3, #4, #5, #. It shows student login logs for Moodle.

Fig. 2(b) Students' log Moodle

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.

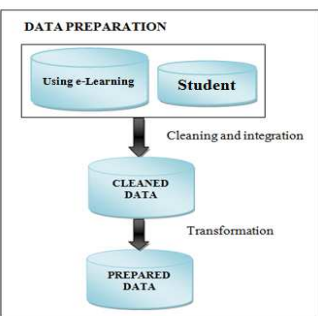


Fig. 3 the data preparation process

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 I
IMPORTANT ATTRIBUTES

Table with 2 columns: Name, Description. Lists various Moodle attributes like PreTest\_number, PostTest\_number, Result-PreTest\_number, etc.

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 algorithm 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.

Best rules found:

1. Midium=Very Low 409 ==> Count-Test-P=Very Low 397 conf:(0.98)
2. Count-P-test=Height 469 ==> Count-Test-P=Very Low 448 conf:(0.98)
3. Count-P-test=Height Midium=Low 341 ==> Count-Test-P=Very Low 324 c
4. Count-P-test=Very Height Count-Test-P=Very Low Grade=High 347 ==> Cou
5. Count-P-test=Very Height Grade=B+ 406 ==> Count-assign=Very Height 38
6. Count-Test-P=Very Low Grade=B+ 374 ==> Count-assign=Very Height 353
7. Grade=High 433 ==> Count-assign=Very Height 408 conf:(0.97)
8. Count-assign=Very Height Grade=High 408 ==> Count-P-test=Very Height
9. Grade=High 433 ==> Count-P-test=Very Height 406 conf:(0.97)
10. Final=Medium Grade=High 341 ==> Count-assign=Very Height 319 conf:
11. Count-assign=Height 441 ==> Count-Test-P=Very Low 412 conf:(0.97)
12. Count-assign=Very Height Count-Test-P=Very Low Grade=High 353 ==> Cou
13. Count-assign=Very Height In-Class=Very Height Project=Very Height 371
14. Final=Medium Grade=High 341 ==> Count-P-test=Very Height 317 conf:

Fig. 4 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 interested rules teacher can use for enhancing their classes. Hence, for this research, the selected rules are explained in the Fig.5, based on decision of teacher.

Discovered rules	Conf.
Count-P-test=Very High Final=Medium Count-assign=Very High ==> Grade= Very High	0.98
Final=Medium Count-assign=Very High ==> Grade= Very High	0.98
Count-assign=Very High In-Class=Very High Count-P-test=Very High ==> Grade= Very High	0.98
Count-P-test=High Project=High Count-Test-P=Very Low ==> Grade= Very High	0.98
Count-assign=Very High Final=Medium Count-P-test=Very High ==> Grade= Very High	0.98
In-Class=Very High Count-P-test=Very High ==> Grade= Very High	0.97
Count-P-test=Very High In-Class=Very High Count-assign=Very High ==> Grade= Very High	0.97
Count-P-test=Very High Count-assign=Very High ==> Grade= Very High	0.97
In-Class=Very High Count-assign=Very High ==> Grade= Very High	0.97
Count-Test-P=Very Low Medium=Low Count-assign=Very High ==> Grade= High	0.97
Count-P-test=Very High Medium=Low Count-assign=Very High ==> Grade= High	0.97
Mid=Low Count-assign=Very High ==> Grade= High	0.96
Mid=Very Low Count-Test-P=Very Low ==> Grade= High	0.96
Final=Medium Count-assign=Very High Count-P-test=Very High ==> Grade= Very High	0.96
Count-P-test=Very High In-Class=Very High Count-assign=Very High ==> Grade= High	0.95

Fig. 5 Results of selected rules

Though numbers of attributes have been defined, only one attribute, "Grade", is significantly attributed to student motivation behavior. Accordingly, it was being discretized in to five levels: VERY LOW, LOW, MEDIUM, HIGH and VERY HIGH. Based on the derived testing result, 30% of the data set used for training, the result is depicted in Figure 6.

Student Motivation Behavior Levels	Testing set	Accuracy	
		Number of Record	Percentage
Very High	224	187	83.5
High	284	254	89.44
Medium	164	125	76.22
Low	112	87	77.68
Very Low	117	78	66.67
Total	901	731	81.13

Fig. 6 Results of testing mined rules

#### V. CONCLUSION AND FUTURE WORK

In this paper, we presents the preliminary result showing a promising progress in this prototypes model for the ongoing improvement of e-Learning course and also this model can be beneficial to similar courses to share and discover students' motivation behavior. However, in term of the future experiments, we are looking forward to research about other data mining techniques to enhance this project and also apply the tool to help teachers in their class.

#### ACKNOWLEDGMENT

The authors gratefully acknowledge the financial subsidy provided by Suan Sunandha Rajabhat University.

#### REFERENCES

- [1] E. García, C. Romero, S. Ventura, and C. Castro, "Using Rules Discovery for the Continuous Improvement of e-Learning Courses," *Lecture Notes in Computer Science*, 2006, Volume 4224/2006, 887-895.
- [2] J. Mostow, J. Beck, H. Cen, A. Cuneo, E. Gouvea, and C. Heiner, "An educational data mining tool to browse tutor-student interactions: Time will tell", In Proceedings of the Workshop on Educational Data Mining, Pittsburgh, USA (pp. 15-22), 2005.
- [3] M. E. Zorrilla, E. Menasalvas, D. Marin, E. Mora, and J. Segovia, "Web usage mining project for improving web-based learning sites", In Web Mining Workshop. Cataluna pp. 1-22, 2005.
- [4] C. Romero, S. Ventura, E. García, "Data mining in course management systems: Moodle case study and tutorial" *Computers & Education*, Volume 51, Issue 1, August 2008, pp. 368-384.
- [5] H.H. Hsu, C.H. Chen, and W.P. Tai, "Towards Error-Free and Personalized Web-Based Courses", In: The 17th International Conference on Advanced Information Networking and Applications, AINA'03. March 27-29, Xian, China, pp. 99-104, 2003.
- [6] A. Kumar, "Rule-Based Adaptive Problem Generation in Programming Tutors and its Evaluation", In: The 12th International Conference on Artificial Intelligence in Education. July 18-22, Amsterdam, pp. 36-44, 2006.
- [7] F. Berzal, J.C. Cubero, N. M. Sánchez, J.M. Serrano, and A. Vila, "Association rule evaluation for classification purposes" *Actas del III Taller Nacional de Minería de Datos y Aprendizaje, TAMIDA2005*, pp.135-144 ISBN: 84-9732-449-8, 2005.
- [8] W. West, B.R.S. Rosser, S. Monani, and L. Gurak, "How Learning Styles Impact ELearning: a Case Comparative Study Of Undergraduate Students Who Excelled, Passed Or Failed An Online Course", In *Scientific/Technical Writing, E-learning*, pp. 534-543, 2006.
- [9] N. Kerdprasop, N. Muenrat, and K. Kerdprasop, "Decision Rule Induction in a Learning Content Management System" *Proceedings of World Academy of Science, Engineering and Technology*, pp.77-81, 2008.
- [10] U. M. Fayyad, G. Pitetsky-Shapiro, P. Smyth, and R. Uthurasamy, "Advances in Knowledge Discovery and Data Mining", AAAI/MIT Press, 1996.
- [11] W. Frawley, G. Piatetsky-Shapiro, and C. Matheus, "Knowledge Discovery in Databases: An Overview". *AI Magazine*, Fall 1992, pp. 213-228.
- [12] R. Agrawal, T. Imielinski, and A.N. Swami, A. N., "Mining association rules between sets of items in large databases", In Proceedings of the

1993 ACM SIGMOD International Conference on Management of Data, pp. 207-216,1993.

- [13] Z. Qiankun, "Association Rule Mining:A Survey, Technical Report",CAIS, Nanyang Technological University, Singapore , 2003
- [14] R. Agrawal, and R. Srikant, "Fast algorithms for mining association rules", In Proc. 20<sup>th</sup> Int. Conf. Very Large Data Bases, VLDB, pp. 487-499, 1994.
- [15] <http://www.cs.waikato.ac.nz/ml/weka/>
- [16] E. García, C. Romero, S. Ventura, C. Castro, and T. Calders, "Chapter 7: Association Rule Mining in Learning Management Systems." In: Handbook of Educational Data Mining, Taylor&Francis Group, 2010.

**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 Tongsir** is interested in Web-Based Educational Learning, and Data mining techniques.