Telematics and Informatics 33 (2016) 773–792

Contents lists available at ScienceDirect

Telematics and Informatics

journal homepage: www.elsevier.com/locate/tele

Change discovery of learning performance in dynamic educational environments

Tony Cheng-Kui Huang [a,*], Chih-Hong Huang [a], Yung-Ting Chuang [b]

[a] Department of Business Administration, National Chung Cheng University, University Rd., Taiwan, ROC
[b] Department of Information Management, National Chung Cheng University, University Rd., Taiwan, ROC

ARTICLE INFO

Article history:
Received 3 April 2015
Accepted 28 October 2015
Available online 31 October 2015

Keywords:
Educational data mining
Association rule
Change mining
Learning performance
Teaching strategies

ABSTRACT

In recent years, as information technology has become more prevalent, learning management systems have arisen around e-learning and web-based platforms. As a result, huge quantities of data about students' learning process have been recorded and stored. Teachers can apply data-mining techniques to mine students' learning performance. One such technique is association rule mining, which can find correlations between student characteristics and performance. For instance, a rule (Attendance = Middle) ∧ (Gender = Male) → (Semester = Low) indicates that the semester grade of students is at the Low level if their gender is Male and attendance rate is Middle, where Low and Middle are predetermined linguistic terms given by teachers. Teachers can rely on such rules to formulate their teaching strategies. However, these rules may be varied in different semesters because of the change of student characteristics or teaching method of teachers. The above rule is used to describe student behavior during the last semester, yet, within this semester, the rule changes to (Attendance = Low) ∧ (Gender = Female) → (Semester = Low). Without updating this knowledge, teachers might adopt inappropriate teaching strategies for students who are learning in different ways across different semesters. In this study, we propose a new change mining model to discover the change in student learning performance and characteristics on the basis of association rules. We conducted experiments with real-life datasets to evaluate the effectiveness of the proposed model.

1. Introduction

With the increasing prevalence of information technologies in the education field, learning management systems (LMS) have been maturely developed and widely adopted to store a wide range of data, including students' characteristics, learning histories, achievements, testing scores, and grades. Teachers use the LMS data to conduct summative assessments to examine class achievement and develop criteria about what students are expected to achieve at certain stages (Chuang, 2015; Gikandi et al., 2011; Harlen and James, 1997). As a result, they can clearly comprehend students' learning behaviors and adjust teaching strategies in a timely manner.

An LMS or similar system includes education-data repositories, which store the learning histories and records of students. These data repositories are treasure troves of valuable learning experiences, and we can analyze them to enhance student achievement. Many previous studies have analyzed learning-history data by applying various data-mining methods to investigate student learning behavior.

* Corresponding author.
E-mail address: bmahck@ccu.edu.tw (T.C.-K. Huang).

http://dx.doi.org/10.1016/j.tele.2015.10.005
0736-5853/

Baker and Yacef (2009) surveyed the most-cited papers in educational data mining (EDM) and proposed several methods, including relationship mining, prediction, clustering, human judgment, discovery with models and none of the above. Calders and Pechenizkiy (2011) pointed out particular problems in LMS logs, and then mapped basic EDM problems to traditional data-mining problems, especially emerging-pattern mining. Romero and Ventura (2007) explored key areas of concern related to the data mining and education realms, particularly where EDM intersects with analytics. In EDM, relationship mining is a popular method for analyzing learners' data, helping students to understand and optimize their learning journey. Effective EDM techniques include classification, clustering, Bayesian modeling, association rule, and text mining (Romero and Ventura, 2010, 2013).

The field of data mining has expanded dramatically in recent years, and the most popular topic in this field is mining association rules (Agrawal et al., 1993; Liu et al., 1996; Srikant and Agrawal, 1995, 1996), in which the objective is to find useful information and extractive patterns from one's database. The classic application is market basket analysis. For example, consider a hypothetical situation in which 30% of the people who bought flowers and cosmetics on transaction-database records also bought chocolate 60% of the time. In other words, it can be said that flowers and cosmetics are associated with chocolate in these transactions: (Flower) ∧ (Cosmetic) → (Chocolate) (*support* = 0.3 and *confidence* = 0.6).

In a transaction database, this rule can not only encompass information about items but also takes into account the quantities of those items being purchased. In the above example, the rule of quantity is "(Flower = 10) ∧ (Cosmetic = 1) → (Chocolate = 2)", indicating that a customer who buys ten floral items and one cosmetic item is likely to buy two chocolate items in a supermarket. Managers of these supermarkets can apply the association rule method to transaction databases in order to identify which items are frequently bought together and determine transaction quantities.

In an LMS database, student characteristics can be treated as items with quantitative or qualitative data, students' test scores can be treated as quantities, and association rule mining can discover connections between characteristics and quantities. For example, consider an association rule, (Attendance = 60) ∧ (Gender = Male) ∧ (FinalReport = 80) → (Semester = 70), describing a student whose semester score is 70 (quantity), attendance score is 60 (quantity), gender is male (quality), and final report score is 80 (quantity) in this semester. Thus, according to this rule, teachers can understand the correlation between learning performance and learning characteristics. The quantity represents a numerical attribute, meaning that the item has a quantitative value, and the quality represents a categorical attribute, meaning that the item has a qualitative class.

Srikant and Agrawal (1996) and Rastogi and Shim (2002) have proposed a model to discover association rules containing quantitative attributes. In this model, a numerical item with its quantity is called quantitative-item, and a categorical item with its class is qualitative-item. Hence, the quantity (a score) can be transformed into a linguistic term by functions. For instance, consider the rule (Attendance = Low) ∧ (Gender = Male) ∧ (FinalReport = High) → (Semester = Middle), where Low, Middle, and High are predetermined linguistic terms supplied by educators.

Although applying the method of association rule in education is workable, such a method still fails to consider changes in the education environment. Understanding the correlation of changes in students' learning performance and changes in their characteristics over time is of importance for teachers. Consider the example below.

In the last semester, teachers have developed the following rule based on student data:

(Attendance = Low) ∧ (Gender = Male) ∧ (FinalReport = High) → (Semester = Middle).

However, in this semester, they found this rule as well:

(Attendance = Low) ∧ (Gender = Female) ∧ (FinalReport = Low) → (Semester = Middle).

Obviously, the gender of the students being examined has changed from Male to Female, and the value of the final report has changed from High to Low. If teachers do not update these two erroneous premises from this semester, then:

(1) Teachers may still believe that students will acquire the semester grade Middle if their gender is Male and their final report grade is High;
(2) In fact, students who acquire the semester grade Middle are that the gender class is Female not Male and the final report grade is Low not High.

Without updating their knowledge and beliefs, teachers might adopt inappropriate teaching strategies and pedagogies for students who are learning in different time-periods. No studies, to our knowledge, have discussed this changing issue in education. To address this research gap, we propose a novel change mining model, Association Rule Change Mining (*ARCM*), to detect the change in student learning performance and characteristics.

The remainder of this paper is organized as follows. Section 2 reviews related works. Section 3 defines a similarity measurement for association rules. Section 4 presents the *ARCM* model for detecting the change in student learning performance and characteristics with the form of association rules. Section 5 shows the experimental results of the proposed model. Conclusions and discussions are drawn in Section 6.

2. Related works

In general, data-mining techniques are usually applied to static settings; that is, users apply them once to dig out rules or patterns from old databases but do not re-run them for new databases. They do not consider that the discovered rules or patterns may vary over time. In a real world, however, data are generated and appended daily to an old database, creating