Application of Association Rules for Finding Correlations Among Students Preliminary Knowledge

Sylvia Encheva[1] and Sharil Tumin[2]

[1] Stord/Haugesund University College, Bjørnsonsg. 45, 5528 Haugesund, Norway
sbe@hsh.no
[2] University of Bergen, IT-Dept., P.O. Box 7800, 5020 Bergen, Norway
edpst@it.uib.no

Abstract. This paper aims at finding an efficient way for discovering which specific knowledge each student does not possess in order to successfully start a new course or to proceed with another section in a current subject. Most existing tutoring systems respond to students' mistakes by providing links to a collection of teaching materials. Such an approach does the individual needs of each student. Our idea is to apply a holistic approach that involves looking at the whole system of each student knowledge within an subject rather than just concentrating on single mistakes, lack of knowledge or misconception.

1 Introduction

Association rules are widely used for detecting relationships between variables. Association rules show attribute value conditions that occur frequently together in a given dataset. They deal with statements of the form 'the presence of attributes α and β often also involves attribute γ'. This approach has an application in different fields such as market basket analysis [5], medical research [7], web clickstream analysis [14] and census data [10].

The process of determining the set of association rules that hold in a context can be broken down into two steps - finding all frequent subsets of attributes, and generating confident rules from the frequent itemsets. Traditional approaches to find frequent itemsets rely on a minimum support threshold in order to reduce the amount of candidates they have to work with [6].

Most existing tutoring systems provide links to a collection of teaching materials. Such an approach does not relate each student's needs to the suggested help. Our idea is to apply a holistic approach that involves looking at the whole system of each student knowledge within an subject rather than just concentrating on single mistakes or lack of knowledge.

To make everything more clearly visible a graphic display of objects and attributes in a lattice structure is provided. Reducing the computational complexity while searching for association rules in the case at least one attribute

Y. Luo (Ed.): CDVE 2006, LNCS 4101, pp. 303–310, 2006.

included in the 'if' part of the statement is known, is presented. A dataset partitioning technique is used to efficiently discover association rules with some given information in the antecedent and without the constraint of support threshold.

The rest of the paper is organized as follows. Related work and definitions and statements from formal concept analysis and rule mining may be found in Section 2. The main results of the paper are placed in Section 3 and Section 4. The paper ends with a conclusion in Section 5.

2 Related Work

A Web-based tutoring tool with mining facilities to improve learning and teaching is described in [12]. An intelligent system for assisting a user in solving a problem was developed in [9].

A system using assessing Bayesian nets for assizing students' knowledge is proposed in [11].

Formal concept analysis [15] started as an attempt of promoting better communication between lattice theorists and users of lattice theory. Since 1980's formal concept analysis has been growing as a research field with a broad spectrum of applications. Various applications of formal concept analysis are presented in [8].

The complexity of mining frequent itemsets is exponential and algorithms for finding such sets have been developed by many authors such as [4] and [6].

Mining association rules is addressed in [1] and [2]. Algorithms for fast discovery of association rules have been presented in [3], [13] [16], and [17].

3 Association Rules Applied to a Calculus Test

Consider students enrolled in a subject where preliminary knowledge is a prerequisite for understanding new content. Students are first suggested to take a test (Test P, Fig. 1) detecting lack of necessary preliminary knowledge or skills. Suppose the test outcome implies insufficient knowledge or skills. The system then provides personalized help to each student based on his/her individual needs. If the test outcome does not imply lack of preliminary knowledge or skills the student is then directed to Chapter 1. In order to proceed with Chapter 2 the student has to take Test 1 (Fig. 1). The procedure is similar to the one describing Test P. However, tests differ in content.

Below we propose an approach for deriving association rules that involves matrices and vectors. The idea is best described with the following example.

For this scenario we consider freshmen enrolled in bachelor program. Students are divided in sixteen groups $G_1, G_2, ..., G_{16}$, according to gender and results from a test providing information about their preliminary knowledge in calculus. The goal is to find the association rules that relate attributes in Table 1 to students' results from the test.