

Project 1: Association Analysis

UE 141 on Data Mining Spring 2013

Each team needs to present project in class and submit a written report before class on March 20. The maximum number of members in a team is 2.

The goal of this project is to evaluate your understanding of association rule mining in real-world applications, and get you familiar with the association analysis function in Weka.

You are asked to complete the following two tasks.

- We talked about the application of association rule mining in supermarket business. In fact, association analysis can be applied to various domains. Please choose one application to discuss how association analysis can help. What are the data sets? What is the expected output? How can people make decisions based on the rules mined from the data sets?
- Find a data set and load it into Weka. Discretize the data if the attributes are numeric. Run the Apriori algorithm on the data and get the association rules. Change the support and confidence thresholds. What changes you observe in the rules?

Your report should include: 1) A short description of the application and answers to the three questions in part 1, and 2) A short description of the data set you use for Weka experiments. Report a few interesting rules and discuss how the extracted rules change with respect to the changes in confidence and support.

Your presentation should be a short overview of your report. You can ignore the details but show the most interesting points.

Note that plagiarism/copying is not allowed and may result in an F in the grades of all the team members. Academic integrity policy can be found at <http://www.cse.buffalo.edu/shared/policies/academic.php>