PSG College of Technology

Department of Applied Mathematics and Computational Sciences

MSc SS VIII Semester – Data Mining

Lab Assignment 1 – 13-01-2021

Your task for this project is to identify and perform an association rule mining task. This involves

1. Selecting an appropriate data set
2. Preparing and preprocessing the data
3. Finding rules, including appropriate parameter setting
4. Determining which of the resulting rules are interesting
5. Figuring out how the interesting rules could be useful

While you are on your own to select an appropriate data set from The UCI Machine Learning Repository or Kaggle. This contains many data sets, not all of which are appropriate for association rules. You are also welcome to identify data from other sources, especially those that you find personally of interest.

**Project Report**

The project report should contain the following:

1. Objectives: What is the domain and what are the potential benefits to be derived from association rule mining. This is high level - not find patterns, but what would improve because of the use of the patterns.
2. Data set description: What is in the data, and what preprocessing was done to make it amenable for association rule mining. Describe your reasoning behind the choices.
3. Rule mining process: Parameter settings, choice of algorithm, and the time required. Experiment with different minimum values for support and confidence so that you get a relatively small but interesting set of rules.
4. Resulting rules: Summary (number of rules, general description), and a selection of those you would show to a client.
5. Recommendations: What should the client do because of the rules discovered.