Department of Computer Science & Engineering

Final Year B. Tech. (CSE) – I : 2022-23

5CS462 : PE5 - Data Mining Lab

Assignment No. 7

Use / extend the data analysis tool (menu driven GUI) developed in Assignment No. 1 to perform the following task :

1. Implement the Apriori algorithm for generating Association Rules

2. Experiment with different values of support, confidence, and maximum rule length.

3. Tabulate the results containing frequent item sets, total number of rules generated for different support and confidence.

4. Find the interesting rules from above obtained rules using following metrics/measures

   a. Lift
   b. Chi-Square Test x^2
   c. All_confidence measure
   d. Max_confidence measure
   e. Kulczynski measure
   f. Cosine measure

5. Use the following data sets from UCI machine learning repository :

   a. US Votes data set
      https://archive.ics.uci.edu/ml/datasets/congressional+voting+records
   b. Extended BAKERY dataset (1000, 5000, 20,000 and 75,000 transactions per set)
      https://users.csc.calpoly.edu/~dekhtyar/466-Spring2018/labs/lab01.html

Deadline : 09/10/2022

Dr. B. F. Momin
Course Teacher