Association Rule Analysis (Market Basket Analysis) Assignment

Dr. George Azzopardi

university of groningen

Groceries data set

- The file groceries.txt contains 9835 rows with comma separated items

- Each row can be considered as a receipt with a transaction of multiple grocery items

Tasks

1. Show the histogram of all items in the 9835 transactions

2. Use the given Matlab implementation (associationRules.m) and improve the selection of association rules by making it more efficient. Use slide 43 of Lecture 1 as a reference.

3. Compare the time needed to generate all association rules based on the given implementation to that of the improved implementation (as per point 2 above).

4. How many rules satisfy the following input parameter values? **_support = 0.001, confidence = 0.8_**

5. Sort the association rules by confidence in descending order and list the top 30 rules.