

HW9 Q2

Dataset taken: Transactions/purchases of customers in a grocery store.

I have taken a sample size of 200 unique transactions.

Each line of a file is a transaction of a single user.

Python code:

Please run the hw9.py python file.

There are two txt files : 'user_items_id.txt' , which has each user's transaction with transaction ID starting with 0.

And another: 'user_trans.txt', which only has transactions of all the 200 users.

For finding association rules, I used minimum support as 0.5 .

I got following rules :

Supp Conf Rule

0.1 1.0 juice -> fruit/vegetable

0.1 1.0 fruit/vegetable -> juice

0.1 0.3 rolls/buns -> vegetables

0.1 0.2 vegetables -> rolls/buns

0.1 0.3 rolls/buns -> milk

0.1 0.2 milk -> rolls/buns

0.1 0.8 butter -> milk

0.1 0.2 milk -> butter

0.1 0.5 vegetables -> milk

0.1 0.4 milk -> vegetables

0.1 0.2 vegetables -> fruit

0.1 0.3 fruit -> vegetables

0.1 0.2 milk -> yogurt

0.1 0.6 yogurt -> milk

0.1 0.2 milk -> fruit

0.1 0.4 fruit -> milk

From the rules, we can say that those customers who buy juice, always buys fruits/vegetables.

Another high confidence is 0.8 for butter and milk.

And 0.6 for yogurt and milk.

I have also printed association rules for each user.

Application:

From the rules, we can recommend the grocery store which items are being purchased together frequently.

From that we can also imply that, If any two items have high confidence, they both should be placed together for better outcome and profit.