

Midterm Project Report

Implementation & Code Usage

The main concept of the Apriori Algorithm is to create associations. In order to create associations, I had to figure out what items were most frequent when given the list of transactions. Once the items that are most frequent were found, depending on the user's support parameter, the support would have to be calculated for each item. After calculating the support value for each item, we can eliminate the items that do not meet the user-defined support parameter. Next, I can then generate set combinations and calculate the support for those item-sets by cross-referencing them to the list of transactions and examine if these item-sets existing in the transactions list.

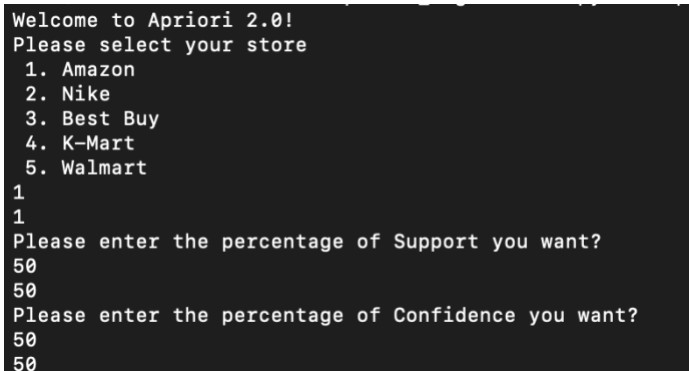
We would repeat this process and increase the number of items in the set by one until I no longer meet the user-defined support parameter. Once I have the item-sets that meet the support then I would have to generate the permutations and calculate the confidence with each item set.

When having the support and confidence of each item-set we can then filter the item-sets that meet the both user-defined parameters of support and confidence and acquire our final associations.

As for code usage, it was in my best interest to use the notions of sets, counters, data frames, and lambdas, loops, and well-defined functions to carry out specific tasks to achieve the Apriori Algorithm.

Screenshots

Below are screen shots to show that the program runs.



```
Welcome to Apriori 2.0!
Please select your store
1. Amazon
2. Nike
3. Best Buy
4. K-Mart
5. Walmart
1
1
Please enter the percentage of Support you want?
50
50
Please enter the percentage of Confidence you want?
50
50
```

Figure 1: Follow the prompts when asked

Michael Woo
28 March 2021
Professor Yasser Abdualлах
CS 634 104 Data Mining

The final output should be the following:

```
Store Name: Amazon

Final Associations that meet the user standards....
Support: 50%    Confidence: 50%

      Association  Support  Confidence
0  (Java: The Complete Reference, Java For Dummies)    50.0    100.0
1  (Java For Dummies, Java: The Complete Reference)    50.0    77.0
```

Figure 2: These are the following Associations given the user parameters

Other

The source code (.py file) and data sets (.csv files) will be attached to the zip file.

Link to Git Repository

<https://github.com/MichaelWoo-git/Apriori Algorithm>