Apriori Algorithm for Association Rule Mining with Python.

By Lionel Ngobesing Alangeh (2020)

Bilgisayar Mühendisliği, Fen Bilimleri Enstitüsü. GAZİ ÜNİVERSİTESİ

ANKARA, TURKEY

Folder Contents:

- myapriori.py: Contains code for the Apriori Algorithm for association rule mining.
- dataset.csv: Big dataset containing data for marketing transactions in EXCEL format
- dataset.txt: Big dataset containing data for marketing transactions in TEXT format if needed
- testdataset.csv: Small dataset containing testing data
- output.txt: File containing output of running program if needed
- README.docx

Programming tools:

- Python 3
- PyCharm IDE

Program Function

- Accept minimum support and minimum confidence from a user as inputs
- Program generates k-itemsets through joining and pruning and displays k-itemset at each step
- Finally program displays various association rules which satisfy minimum support and minimum confidence conditions.

Sample Output

```
C:\Users\lionel\PycharmProjects\PracticeExercises\venv\Scripts\python.exe C:/Users/lionel/PycharmProjects/PracticeExercises/venv/Apriori/myapriori.py
Enter Minimum support value in %: 5
Enter Minimum confidence value in %: 10
********************* The TEST DATA SET **************************
[['citrus fruit', 'semi-finished bread', 'margarine', 'ready soups'], ['tropical fruit', 'yogurt', 'coffee'], ['whole milk'], ['pip fruit', 'yogurt', 'cream ch
****************************************************************
********************* FREQUENT 1-ITEMSET *************************
[['citrus fruit'], ['margarine'], ['tropical fruit'], ['yogurt'], ['coffee'], ['whole milk'], ['pip fruit'], ['other vegetables'], ['butter'], ['rolls/buns'],
****************************************************************
********************* FREQUENT 2-ITEMSET ************************
[['whole milk', 'yogurt'], ['other vegetables', 'whole milk'], ['rolls/buns', 'whole milk']]
****************************************************************
********************* FREQUENT 3-ITEMSET ************************
[]
****************************************************************
******************** GENERATED ASSOCIATION RULES ********************
******* RULE ******                    ******* {SUPPORT, CONFIDENCE} ****
----------------------------------------------------------------
Rule# 1 : ['whole milk'] --> ['yogurt']              {sup = 551/9835 ,conf = 551/2513}
Rule# 2 : ['yogurt'] --> ['whole milk']              {sup = 551/9835 ,conf = 551/1372}
Rule# 3 : ['other vegetables'] --> ['whole milk']        {sup = 736/9835 ,conf = 736/1903}
Rule# 4 : ['whole milk'] --> ['other vegetables']        {sup = 736/9835 ,conf = 736/2513}
Rule# 5 : ['rolls/buns'] --> ['whole milk']        {sup = 557/9835 ,conf = 557/1809}
Rule# 6 : ['whole milk'] --> ['rolls/buns']        {sup = 557/9835 ,conf = 557/2513}
----------------------------------------------------------------

Process finished with exit code 0
```