Apriori Algorithm Implementation Using Python

Data Mining

CS634

By - Teja Kulkarni

GITHUB : https://github.com/TejaKulkarni/Apriori_Algorithm

1. Introduction

In the period of unconvinced life and digital shopping and everything could be available online still will try to find a time to go supermarkets for shopping or quick pick up. But did we think why certain items are placed together and are there any reason behind their placement such as milk near cookies, conditioners placed near shampoo and so on ...
The shopkeeper knows the customers' sentiment so they look to optimising the sales and makes a profit out of it.
Market Basket Analysis is a common technique used *specially* by large retailers to find hidden patterns on customer behaviours.

1.1 Apriori Algorithm Overview

It is an algorithm for discovering frequent itemsets in transaction databases. This algorithm, introduced by R Agrawal and R Srikant in 1994 has great significance in data mining.
Apriori is one of the algorithms that use in E-commerce website as recommendation features websites to get recommended contents. Applied when we have transactions databases.In final we are looking to get frequent item sets i.e. items which are bought most frequently.

1.2 Association Rule Mining

Association Rule Mining is used when you want to find an association between different objects in a set, find frequent patterns in a transaction database, relational databases or any other information repository. The applications of Association Rule Mining are found in Marketing, Basket Data Analysis (or Market Basket Analysis) in retailing, clustering and classification. The most common approach to find these patterns is Market Basket Analysis, which is a key technique used by large retailers like Amazon, Flipkart, etc to analyse customer buying habits by finding associations between the different items that customers place in their "shopping baskets". The discovery of these associations can help retailers develop marketing strategies by gaining insight into which items are frequently purchased together by customers.

Support

Support tells about the items that are frequently bought together. Support count is the frequency of occurrence of an item-set.

Confidence

If items A & B are bought together, Confidence tells us the number of times that A & B are bought together, given the number of times A is bought. For every purchase of A, Confidence tells us the number of times that B was also bought along with A.

Confidence c = frequency(A & B)/frequency(A)

2. Steps to implement:

Step 1: Scan the whole transaction database to fetch the support value S for each item.

Step 2: If the Support S is more than or equal to the minimum threshold, add the item to frequent itemset (L1), else go to step 1.

Step 3: Join Lk-1 and Lk-1, and generate the set of candidate k-itemsets.

Step 4: For each k-itemset, get the support S and check the minimum support threshold.

Step 5: Repeat the iteration in step 4, if support is not more than or equal to the minimum value.

Step 6: If S is more than the required value, add to the frequent k-itemsets.

Step 7: If there are no itemsets, stop the algorithm.

Step 8: Till there are frequent itemsets, for each frequent itemset L, get all the non-empty subsets.

Step 9: For each frequent subset of L, find the confidence C.

Step 10: If the Confidence C is more than or equal to the minimum required Confidence, add it to the strong rules, else move to the next frequent subset.

3. AlgorithmArchitecture:

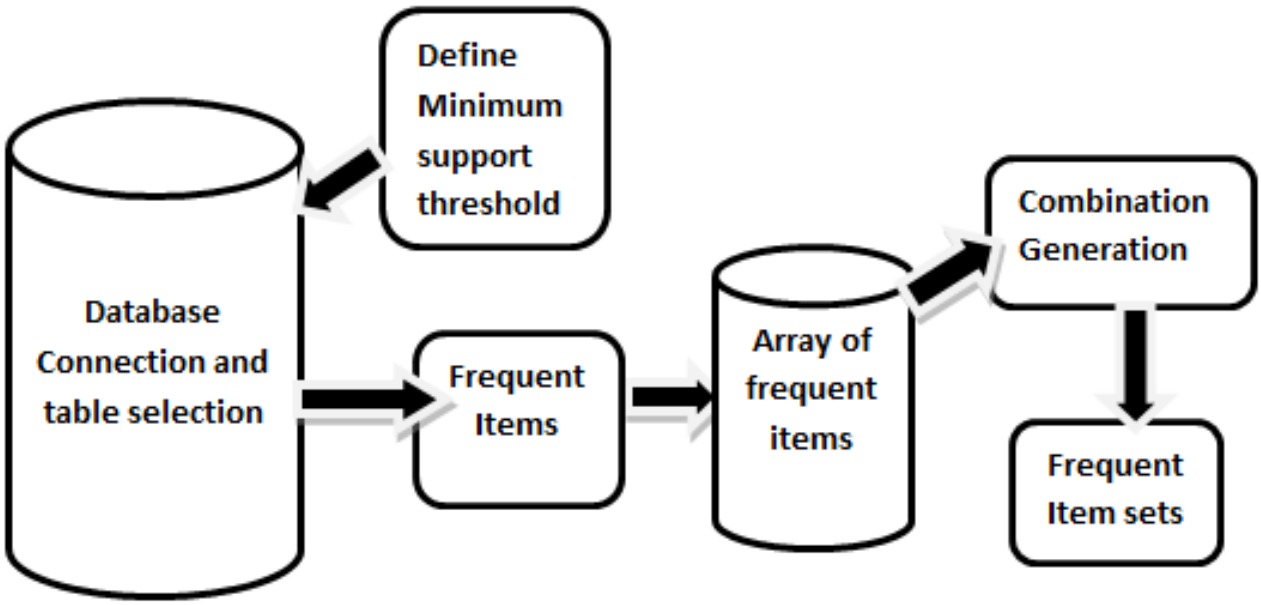

Figure 1 : Block Diagram implementation

5. Project Description:

Write a Apriori algorithm, generate and print out all the association rules and the input transactions for each of the 4 transactional databases given (support and confidence should be user-specified parameters, so the output should show different support and confidence values with respect to different databases).

5.1 Minimum Requirements:

1. Hardware Specification:

Hard Disk – 1TB
RAM – 8 GB
OS – MAC OS
Processor – Dual-Core Intel Core i5

2. Software Specification:

Anaconda ,Python, Jupyter NoteBook(6.2.0)
Libraries : Panda , NumPy , CSV

SOURCE CODE :

```
import pandas as pd
import numpy as np
from itertools import combinations

choice = input("Please, Select your Dataset for \n 1 Amazon.\n 2 bestbuy.\n 3 K_mart.\n 4 Nike.\n 5 Generic. \n ")
choice = int(choice)

if choice == 1:
    df =pd.read_csv('/Users/teja/Kulkarni_Teja_MidtermProj/amazon.csv')
    print(df)
elif choice == 2:
    df = pd.read_csv('/Users/teja/Kulkarni_Teja_MidtermProj/bestbuy.csv')
    print(df.head())
elif choice == 3:
    df = pd.read_csv('/Users/teja/Kulkarni_Teja_MidtermProj/kmart.csv')
    print(df)
elif choice == 4:
    df = pd.read_csv('/Users/teja/Kulkarni_Teja_MidtermProj/nike.csv')
    print(df)
elif choice == 5:
    df = pd.read_csv('/Users/teja/Kulkarni_Teja_MidtermProj/generic.csv')
    print(df)
```

```
else:
    print("Wrong Choice")

min_sup = input("Please, input your Min. Support \n")
min_sup = float(min_sup)
min_con = input("Please, input your Min. confidence \n")
min_con = float(min_con)

names = list(df.columns)
tid = df[names[0]]
items = df[names[1]]
uni_items = df[names[1]].unique()
uni_tid = df[names[0]].unique()

def build_transactions(uni_tid, tid, items):
    transactions = []
    for i in uni_tid:
        temp_list = []
        for j in range(0, len(tid)):
            if tid[j] == i:
                temp_list.append(items[j])
        transactions.append(temp_list)
    return(transactions)

transactions = build_transactions(uni_tid, tid, items)
num_trans = len(transactions)
def check_pattern(list1, list2):
    x = 0
    if(all(x in list2 for x in list1)):
        x = 1
    return x

def update_fre_items (a, b):
    f = []

    for i in a:
        for j in i:
            f.append(j)

    temp = []
    for i in b:
        if i  in f:
            temp.append(i)

    return temp

pat_size = 1
fre_pat = []
```

```
#Number of pattrens
fre_pat_count = []
temp_fre_pat = [1]
fre_items = list(uni_items)
while (temp_fre_pat):

    # generate acceptable patterns
    pats = combinations(fre_items, pat_size)
    temp_fre_pat = [] # frequent patterns
    for f in list(pats):
        count = 0
        for t in transactions:
            count = count + check_pattern(f, t)
        if count >= min_sup * num_trans:
            temp_fre_pat.append(f)
            fre_pat_count.append(count)

    fre_pat = fre_pat + temp_fre_pat
    pat_size += 1
    # update frequent items list for creating new patterns
    fre_items = update_fre_items(temp_fre_pat, fre_items)

print('frequent patterns   \n',fre_pat)

print('\nAssociation rules')

for i in fre_pat:

    if len(i) > 1:

        sub_groups = list(combinations(i, len(i) - 1))
        #print(sub_groups)
        for j in sub_groups:
            temp = []
            for k in j:
                temp.append(k)
            z = list(set(i).difference(set(temp)))
            confidence = fre_pat_count[fre_pat.index(i)] / fre_pat_count[fre_pat.index(j)]
            if confidence > min_con:
                print(j,' ---> ', z ,'   confidence = ‘,confidence)
```

SCREENSHOTS :

1. AMAZON

```
Last login: Fri Nov 12 20:55:40 on ttys000

The default interactive shell is now zsh.
To update your account to use zsh, please run `chsh -s /bin/zsh`.
For more details, please visit https://support.apple.com/kb/HT208050.
(base) tejas-MacBook-Air:~ teja$ cd Kulkarni_Teja_MidtermProj/
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ ls
Midterm_Project_Data_Mining.py  amazon.csv                      bestbuy.csv                     generic.csv                     kmart.csv                       nike.csv
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 1
                         transaction                         Unnamed: 1  ...                                 Unnamed: 3                         Unnamed: 4
0                    A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
1                    A Beginner's Guide          Java: The Complete Reference  ...                                  NaN                                NaN
2                    A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch  Head First Java 2nd Edition
3   Android Programming: The Big Nerd Ranch         Head First Java 2nd Edition  ...                                  NaN                                NaN
4   Android Programming: The Big Nerd Ranch      Beginning Programming with Java  ...                                  NaN                                NaN
5                    A Beginner's Guide  Android Programming: The Big Nerd Ranch  ...                                  NaN                                NaN
6                    A Beginner's Guide         Head First Java 2nd Edition   ...                                  NaN                                NaN
7            Java: The Complete Reference                   Java For Dummies  ...                                  NaN                                NaN
8                      Java For Dummies  Android Programming: The Big Nerd Ranch  ...          Beginning Programming with Java                                NaN
9          Beginning Programming with Java                  Java 8 Pocket Guide  ...                                  NaN                                NaN
10                   A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
11                   A Beginner's Guide          Java: The Complete Reference  ...  HTML and CSS: Design and Build Websites                                NaN
12                   A Beginner's Guide          Java: The Complete Reference  ...                  Java 8 Pocket Guide  HTML and CSS: Design and Build Websites
13                     Java For Dummies  Android Programming: The Big Nerd Ranch  ...                                  NaN                                NaN
14                     Java For Dummies  Android Programming: The Big Nerd Ranch  ...                                  NaN                                NaN
15                   A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
16                   A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
17           Head First Java 2nd Edition       Beginning Programming with Java  ...                                  NaN                                NaN
18  Android Programming: The Big Nerd Ranch         Head First Java 2nd Edition  ...                                  NaN                                NaN
19                   A Beginner's Guide          Java: The Complete Reference  ...                                  NaN                                NaN

[20 rows x 5 columns]
Please, input your Min. Support
```

```
Midterm_Project_Data_Mining.py  amazon.csv                      bestbuy.csv                     generic.csv                     kmart.csv                       nike.csv
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 1
                         transaction                         Unnamed: 1  ...                                 Unnamed: 3                         Unnamed: 4
0                    A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
1                    A Beginner's Guide          Java: The Complete Reference  ...                                  NaN                                NaN
2                    A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch  Head First Java 2nd Edition
3   Android Programming: The Big Nerd Ranch         Head First Java 2nd Edition  ...                                  NaN                                NaN
4   Android Programming: The Big Nerd Ranch      Beginning Programming with Java  ...                                  NaN                                NaN
5                    A Beginner's Guide  Android Programming: The Big Nerd Ranch  ...                                  NaN                                NaN
6                    A Beginner's Guide         Head First Java 2nd Edition   ...                                  NaN                                NaN
7            Java: The Complete Reference                   Java For Dummies  ...                                  NaN                                NaN
8                      Java For Dummies  Android Programming: The Big Nerd Ranch  ...          Beginning Programming with Java                                NaN
9          Beginning Programming with Java                  Java 8 Pocket Guide  ...                                  NaN                                NaN
10                   A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
11                   A Beginner's Guide          Java: The Complete Reference  ...  HTML and CSS: Design and Build Websites                                NaN
12                   A Beginner's Guide          Java: The Complete Reference  ...                  Java 8 Pocket Guide  HTML and CSS: Design and Build Websites
13                     Java For Dummies  Android Programming: The Big Nerd Ranch  ...                                  NaN                                NaN
14                     Java For Dummies  Android Programming: The Big Nerd Ranch  ...                                  NaN                                NaN
15                   A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
16                   A Beginner's Guide          Java: The Complete Reference  ...  Android Programming: The Big Nerd Ranch                                NaN
17           Head First Java 2nd Edition       Beginning Programming with Java  ...                                  NaN                                NaN
18  Android Programming: The Big Nerd Ranch         Head First Java 2nd Edition  ...                                  NaN                                NaN
19                   A Beginner's Guide          Java: The Complete Reference  ...                                  NaN                                NaN

[20 rows x 5 columns]
Please, input your Min. Support
0.1
Please, input your Min. confidence
0.2
frequent patterns
 [(' Java: The Complete Reference',), (' Head First Java 2nd Edition ',), (' Beginning Programming with Java',), (' Android Programming: The Big Nerd Ranch',), (' Java For Dummies',), (' Java 8 Pocket Gui
de',), (' Head First Java 2nd Edition',), (' Java: The Complete Reference', ' Head First Java 2nd Edition '), (' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch'), (' Head First J
ava 2nd Edition ', ' Beginning Programming with Java'), (' Head First Java 2nd Edition ', ' Android Programming: The Big Nerd Ranch'), (' Head First Java 2nd Edition ', ' Head First Java 2nd Edition'), (' 
 Beginning Programming with Java', ' Head First Java 2nd Edition'), (' Java: The Complete Reference', ' Head First Java 2nd Edition ', ' Android Programming: The Big Nerd Ranch'), (' Head First Java 2nd E
dition ', ' Beginning Programming with Java', ' Head First Java 2nd Edition')]

Association rules
(' Java: The Complete Reference',)  --->  [' Head First Java 2nd Edition ']    confidence =  1.0
(' Head First Java 2nd Edition ',)  --->  [' Java: The Complete Reference']    confidence =  0.5
(' Java: The Complete Reference',)  --->  [' Android Programming: The Big Nerd Ranch']    confidence =  1.0
(' Android Programming: The Big Nerd Ranch',)  --->  [' Java: The Complete Reference']    confidence =  0.5
(' Head First Java 2nd Edition ',)  --->  [' Beginning Programming with Java']    confidence =  0.5
(' Beginning Programming with Java',)  --->  [' Head First Java 2nd Edition ']    confidence =  0.5
(' Head First Java 2nd Edition ',)  --->  [' Android Programming: The Big Nerd Ranch']    confidence =  0.5
(' Android Programming: The Big Nerd Ranch',)  --->  [' Head First Java 2nd Edition ']    confidence =  0.5
(' Head First Java 2nd Edition ',)  --->  [' Head First Java 2nd Edition']    confidence =  0.5
(' Head First Java 2nd Edition',)  --->  [' Head First Java 2nd Edition ']    confidence =  1.0
(' Beginning Programming with Java',)  --->  [' Head First Java 2nd Edition']    confidence =  0.5
(' Head First Java 2nd Edition',)  --->  [' Beginning Programming with Java']    confidence =  1.0
(' Java: The Complete Reference', ' Head First Java 2nd Edition ')  --->  [' Android Programming: The Big Nerd Ranch']    confidence =  1.0
(' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch')  --->  [' Head First Java 2nd Edition ']    confidence =  1.0
(' Head First Java 2nd Edition ', ' Android Programming: The Big Nerd Ranch')  --->  [' Java: The Complete Reference']    confidence =  1.0
(' Head First Java 2nd Edition ', ' Beginning Programming with Java')  --->  [' Head First Java 2nd Edition']    confidence =  1.0
(' Head First Java 2nd Edition ', ' Head First Java 2nd Edition')  --->  [' Beginning Programming with Java']    confidence =  1.0
(' Beginning Programming with Java', ' Head First Java 2nd Edition')  --->  [' Head First Java 2nd Edition ']    confidence =  1.0
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$
```

2.BESTBUY

```
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 2
  transaction    Unnamed: 1        Unnamed: 2        Unnamed: 3           Unnamed: 4     Unnamed: 5           Unnamed: 6 Unnamed: 7 Unnamed: 8 Unnamed: 9
0    Desk Top       Printer       Flash Drive   Microsoft Office            Speakers     Anti-Virus                  NaN        NaN        NaN        NaN
1     Lab Top   Flash Drive   Microsoft Office      Lab Top Case           Anti-Virus            NaN                  NaN        NaN        NaN        NaN
2     Lab Top       Printer       Flash Drive   Microsoft Office           Anti-Virus   Lab Top Case   External Hard-Drive        NaN        NaN        NaN
3     Lab Top       Printer       Flash Drive         Anti-Virus   External Hard-Drive   Lab Top Case                  NaN        NaN        NaN        NaN
4     Lab Top   Flash Drive      Lab Top Case         Anti-Virus                  NaN            NaN                  NaN        NaN        NaN        NaN
Please, input your Min. Support
```

```
Please, input your Min. Support
0.15
Please, input your Min. confidence
0.1
frequent patterns
 [(' Printer',), (' Flash Drive',), (' External Hard-Drive',), (' Lab Top',), (' Desk Top',), (' Speakers',), (' Lab Top Case',), (' Printer', ' Flash Drive'), (' Printer', ' External Hard-Drive'), (' Pri
nter', ' Desk Top'), (' Flash Drive', ' External Hard-Drive'), (' Flash Drive', ' Desk Top'), (' External Hard-Drive', ' Desk Top'), (' Lab Top', ' Desk Top'), (' Lab Top', ' Speakers'), (' Lab Top', ' La
b Top Case'), (' Desk Top', ' Speakers'), (' Desk Top', ' Lab Top Case'), (' Speakers', ' Lab Top Case'), (' Printer', ' Flash Drive', ' External Hard-Drive'), (' Printer', ' Flash Drive', ' Desk Top'), (
' Printer', ' External Hard-Drive', ' Desk Top'), (' Flash Drive', ' External Hard-Drive', ' Desk Top'), (' Lab Top', ' Desk Top', ' Speakers'), (' Lab Top', ' Desk Top', ' Lab Top Case'), (' Lab Top', '
Speakers', ' Lab Top Case'), (' Desk Top', ' Speakers', ' Lab Top Case'), (' Printer', ' Flash Drive', ' External Hard-Drive', ' Desk Top'), (' Lab Top', ' Desk Top', ' Speakers', ' Lab Top Case')]

Association rules
(' Printer',)  --->  [' Flash Drive']    confidence =  0.5
(' Flash Drive',)  --->  [' Printer']    confidence =  0.3333333333333333
(' Printer',)  --->  [' External Hard-Drive']    confidence =  0.5
(' External Hard-Drive',)  --->  [' Printer']    confidence =  1.0
(' Printer',)  --->  [' Desk Top']    confidence =  0.5
(' Desk Top',)  --->  [' Printer']    confidence =  0.5
(' Flash Drive',)  --->  [' External Hard-Drive']    confidence =  0.3333333333333333
(' External Hard-Drive',)  --->  [' Flash Drive']    confidence =  1.0
(' Flash Drive',)  --->  [' Desk Top']    confidence =  0.3333333333333333
(' Desk Top',)  --->  [' Flash Drive']    confidence =  0.5
(' External Hard-Drive',)  --->  [' Desk Top']    confidence =  1.0
(' Desk Top',)  --->  [' External Hard-Drive']    confidence =  0.5
(' Lab Top',)  --->  [' Desk Top']    confidence =  1.0
(' Desk Top',)  --->  [' Lab Top']    confidence =  0.5
(' Lab Top',)  --->  [' Speakers']    confidence =  1.0
(' Speakers',)  --->  [' Lab Top']    confidence =  1.0
(' Lab Top',)  --->  [' Lab Top Case']    confidence =  1.0
(' Lab Top Case',)  --->  [' Lab Top']    confidence =  1.0
(' Desk Top',)  --->  [' Speakers']    confidence =  0.5
(' Speakers',)  --->  [' Desk Top']    confidence =  1.0
(' Desk Top',)  --->  [' Lab Top Case']    confidence =  0.5
(' Lab Top Case',)  --->  [' Desk Top']    confidence =  1.0
(' Speakers',)  --->  [' Lab Top Case']    confidence =  1.0
(' Lab Top Case',)  --->  [' Speakers']    confidence =  1.0
(' Printer', ' Flash Drive')  --->  [' External Hard-Drive']    confidence =  1.0
(' Printer', ' External Hard-Drive')  --->  [' Flash Drive']    confidence =  1.0
(' Flash Drive', ' External Hard-Drive')  --->  [' Printer']    confidence =  1.0
(' Printer', ' Flash Drive')  --->  [' Desk Top']    confidence =  1.0
(' Printer', ' Desk Top')  --->  [' Flash Drive']    confidence =  1.0
(' Flash Drive', ' Desk Top')  --->  [' Printer']    confidence =  1.0
(' Printer', ' External Hard-Drive')  --->  [' Desk Top']    confidence =  1.0
(' Printer', ' Desk Top')  --->  [' External Hard-Drive']    confidence =  1.0
(' External Hard-Drive', ' Desk Top')  --->  [' Printer']    confidence =  1.0
(' Flash Drive', ' External Hard-Drive')  --->  [' Desk Top']    confidence =  1.0
(' Flash Drive', ' Desk Top')  --->  [' External Hard-Drive']    confidence =  1.0
(' External Hard-Drive', ' Desk Top')  --->  [' Flash Drive']    confidence =  1.0
(' Lab Top', ' Desk Top')  --->  [' Speakers']    confidence =  1.0
(' Lab Top', ' Speakers')  --->  [' Desk Top']    confidence =  1.0
(' Desk Top', ' Speakers')  --->  [' Lab Top']    confidence =  1.0
(' Lab Top', ' Desk Top')  --->  [' Lab Top Case']    confidence =  1.0
(' Lab Top', ' Lab Top Case')  --->  [' Desk Top']    confidence =  1.0
(' Desk Top', ' Lab Top Case')  --->  [' Lab Top']    confidence =  1.0
(' Lab Top', ' Speakers')  --->  [' Lab Top Case']    confidence =  1.0
(' Lab Top', ' Lab Top Case')  --->  [' Speakers']    confidence =  1.0
(' Speakers', ' Lab Top Case')  --->  [' Lab Top']    confidence =  1.0
(' Desk Top', ' Speakers')  --->  [' Lab Top Case']    confidence =  1.0
(' Desk Top', ' Lab Top Case')  --->  [' Speakers']    confidence =  1.0
(' Speakers', ' Lab Top Case')  --->  [' Desk Top']    confidence =  1.0
(' Printer', ' Flash Drive', ' External Hard-Drive')  --->  [' Desk Top']    confidence =  1.0
(' Printer', ' Flash Drive', ' Desk Top')  --->  [' External Hard-Drive']    confidence =  1.0
(' Printer', ' External Hard-Drive', ' Desk Top')  --->  [' Flash Drive']    confidence =  1.0
```

3. K-MART

```
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 3
             transaction             Unnamed: 1             Unnamed: 2           Unnamed: 3     Unnamed: 4           Unnamed: 5 Unnamed: 6
0      Decorative Pillows                 Quilts   Embroidered Bedspread                  NaN            NaN                  NaN        NaN
1   Embroidered Bedspread                  Shams          Kids Bedding   Bedding Collections      Bed Skirts           Bedspreads     Sheets
2      Decorative Pillows                 Quilts   Embroidered Bedspread                Shams   Kids Bedding   Bedding Collections        NaN
3            Kids Bedding     Bedding Collections                 Sheets           Bedspreads     Bed Skirts                  NaN        NaN
4      Decorative Pillows            Kids Bedding     Bedding Collections               Sheets     Bed Skirts           Bedspreads        NaN
5     Bedding Collections             Bedspreads             Bed Skirts               Sheets          Shams          Kids Bedding        NaN
6      Decorative Pillows                 Quilts                    NaN                  NaN            NaN                  NaN        NaN
7      Decorative Pillows                 Quilts   Embroidered Bedspread                  NaN            NaN                  NaN        NaN
8              Bedspreads            Bed Skirts                  Shams          Kids Bedding         Sheets                  NaN        NaN
9                  Quilts   Embroidered Bedspread     Bedding Collections                  NaN            NaN                  NaN        NaN
10    Bedding Collections             Bedspreads             Bed Skirts          Kids Bedding          Shams               Sheets        NaN
11     Decorative Pillows                 Quilts                    NaN                  NaN            NaN                  NaN        NaN
12  Embroidered Bedspread                  Shams                    NaN                  NaN            NaN                  NaN        NaN
13                 Sheets                  Shams             Bed Skirts          Kids Bedding            NaN                  NaN        NaN
14     Decorative Pillows                 Quilts                    NaN                  NaN            NaN                  NaN        NaN
15     Decorative Pillows            Kids Bedding             Bed Skirts                Shams            NaN                  NaN        NaN
16     Decorative Pillows                 Shams             Bed Skirts                  NaN            NaN                  NaN        NaN
17                 Quilts                 Sheets          Kids Bedding                  NaN            NaN                  NaN        NaN
18                  Shams            Bed Skirts          Kids Bedding               Sheets            NaN                  NaN        NaN
19     Decorative Pillows            Bedspreads                 Shams               Sheets     Bed Skirts          Kids Bedding        NaN
Please, input your Min. Support
```

```
 3
             transaction             Unnamed: 1             Unnamed: 2           Unnamed: 3     Unnamed: 4           Unnamed: 5 Unnamed: 6
0      Decorative Pillows                 Quilts   Embroidered Bedspread                  NaN            NaN                  NaN        NaN
1   Embroidered Bedspread                  Shams          Kids Bedding   Bedding Collections      Bed Skirts           Bedspreads     Sheets
2      Decorative Pillows                 Quilts   Embroidered Bedspread                Shams   Kids Bedding   Bedding Collections        NaN
3            Kids Bedding     Bedding Collections                 Sheets           Bedspreads     Bed Skirts                  NaN        NaN
4      Decorative Pillows            Kids Bedding     Bedding Collections               Sheets     Bed Skirts           Bedspreads        NaN
5     Bedding Collections             Bedspreads             Bed Skirts               Sheets          Shams          Kids Bedding        NaN
6      Decorative Pillows                 Quilts                    NaN                  NaN            NaN                  NaN        NaN
7      Decorative Pillows                 Quilts   Embroidered Bedspread                  NaN            NaN                  NaN        NaN
8              Bedspreads            Bed Skirts                  Shams          Kids Bedding         Sheets                  NaN        NaN
9                  Quilts   Embroidered Bedspread     Bedding Collections                  NaN            NaN                  NaN        NaN
10    Bedding Collections             Bedspreads             Bed Skirts          Kids Bedding          Shams               Sheets        NaN
11     Decorative Pillows                 Quilts                    NaN                  NaN            NaN                  NaN        NaN
12  Embroidered Bedspread                  Shams                    NaN                  NaN            NaN                  NaN        NaN
13                 Sheets                  Shams             Bed Skirts          Kids Bedding            NaN                  NaN        NaN
14     Decorative Pillows                 Quilts                    NaN                  NaN            NaN                  NaN        NaN
15     Decorative Pillows            Kids Bedding             Bed Skirts                Shams            NaN                  NaN        NaN
16     Decorative Pillows                 Shams             Bed Skirts                  NaN            NaN                  NaN        NaN
17                 Quilts                 Sheets          Kids Bedding                  NaN            NaN                  NaN        NaN
18                  Shams            Bed Skirts          Kids Bedding               Sheets            NaN                  NaN        NaN
19     Decorative Pillows            Bedspreads                 Shams               Sheets     Bed Skirts          Kids Bedding        NaN
Please, input your Min. Support
0.1
Please, input your Min. confidence
0.1
frequent patterns
 [(' Quilts',), (' Shams',), (' Bedding Collections',), (' Kids Bedding',), (' Bedspreads',), (' Bed Skirts',), (' Embroidered Bedspread',), (' Sheets',), (' Quilts', ' Shams'), (' Quilts', ' Kids Bedding
'), (' Quilts', ' Bedspreads'), (' Shams', ' Kids Bedding'), (' Shams', ' Bedspreads'), (' Kids Bedding', ' Bedspreads'), (' Embroidered Bedspread', ' Sheets'), (' Quilts', ' Shams', ' Kids Bedding'), ('
Quilts', ' Shams', ' Bedspreads'), (' Quilts', ' Kids Bedding', ' Bedspreads'), (' Shams', ' Kids Bedding', ' Bedspreads'), (' Quilts', ' Shams', ' Kids Bedding', ' Bedspreads')]

Association rules
(' Quilts',)  --->  [' Shams']    confidence =  1.0
(' Shams',)  --->  [' Quilts']    confidence =  0.3333333333333333
(' Quilts',)  --->  [' Kids Bedding']    confidence =  1.0
(' Kids Bedding',)  --->  [' Quilts']    confidence =  1.0
(' Quilts',)  --->  [' Bedspreads']    confidence =  1.0
(' Bedspreads',)  --->  [' Quilts']    confidence =  0.5
(' Shams',)  --->  [' Kids Bedding']    confidence =  0.3333333333333333
(' Kids Bedding',)  --->  [' Shams']    confidence =  1.0
(' Shams',)  --->  [' Bedspreads']    confidence =  0.3333333333333333
(' Bedspreads',)  --->  [' Shams']    confidence =  0.5
(' Kids Bedding',)  --->  [' Bedspreads']    confidence =  1.0
(' Bedspreads',)  --->  [' Kids Bedding']    confidence =  0.5
(' Embroidered Bedspread',)  --->  [' Sheets']    confidence =  1.0
(' Sheets',)  --->  [' Embroidered Bedspread']    confidence =  1.0
(' Quilts', ' Shams')  --->  [' Kids Bedding']    confidence =  1.0
(' Quilts', ' Kids Bedding')  --->  [' Shams']    confidence =  1.0
(' Shams', ' Kids Bedding')  --->  [' Quilts']    confidence =  1.0
(' Quilts', ' Shams')  --->  [' Bedspreads']    confidence =  1.0
(' Quilts', ' Bedspreads')  --->  [' Shams']    confidence =  1.0
(' Shams', ' Bedspreads')  --->  [' Quilts']    confidence =  1.0
(' Quilts', ' Kids Bedding')  --->  [' Bedspreads']    confidence =  1.0
(' Quilts', ' Bedspreads')  --->  [' Kids Bedding']    confidence =  1.0
(' Kids Bedding', ' Bedspreads')  --->  [' Quilts']    confidence =  1.0
(' Shams', ' Kids Bedding')  --->  [' Bedspreads']    confidence =  1.0
(' Shams', ' Bedspreads')  --->  [' Kids Bedding']    confidence =  1.0
(' Kids Bedding', ' Bedspreads')  --->  [' Shams']    confidence =  1.0
(' Quilts', ' Shams', ' Kids Bedding')  --->  [' Bedspreads']    confidence =  1.0
(' Quilts', ' Shams', ' Bedspreads')  --->  [' Kids Bedding']    confidence =  1.0
(' Quilts', ' Kids Bedding', ' Bedspreads')  --->  [' Shams']    confidence =  1.0
(' Shams', ' Kids Bedding', ' Bedspreads')  --->  [' Quilts']    confidence =  1.0
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$
```

4. NIKE

```
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 4
       transaction       Unnamed: 1       Unnamed: 2       Unnamed: 3     Unnamed: 4       Unnamed: 5   Unnamed: 6       Unnamed: 7       Unnamed: 8       Unnamed: 9
0     Running Shoe            Socks      Sweatshirts     Modern Pants            NaN              NaN          NaN              NaN              NaN              NaN
1     Running Shoe            Socks      Sweatshirts              NaN            NaN              NaN          NaN              NaN              NaN              NaN
2     Running Shoe            Socks      Sweatshirts     Modern Pants            NaN              NaN          NaN              NaN              NaN              NaN
3     Running Shoe       Sweatshirts     Modern Pants              NaN            NaN              NaN          NaN              NaN              NaN              NaN
4     Running Shoe            Socks      Sweatshirts     Modern Pants     Soccer Shoe              NaN          NaN              NaN              NaN              NaN
5     Running Shoe            Socks      Sweatshirts              NaN            NaN              NaN          NaN              NaN              NaN              NaN
6     Running Shoe            Socks      Sweatshirts     Modern Pants      Tech Pants       Rash Guard      Hoodies              NaN              NaN              NaN
7    Swimming Shirt            Socks      Sweatshirts              NaN            NaN              NaN          NaN              NaN              NaN              NaN
8    Swimming Shirt       Rash Guard   Dry Fit V-Nick          Hoodies      Tech Pants              NaN          NaN              NaN              NaN              NaN
9    Swimming Shirt       Rash Guard              Dry              NaN            NaN              NaN          NaN              NaN              NaN              NaN
10  Swimming Shirt       Rash Guard   Dry Fit V-Nick              NaN            NaN              NaN          NaN              NaN              NaN              NaN
11    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants      Soccer Shoe   Rash Guard          Hoodies       Tech Pants   Dry Fit V-Nick
12    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants      Soccer Shoe   Rash Guard        Tech Pants   Dry Fit V-Nick          Hoodies
13    Running Shoe   Swimming Shirt       Rash Guard       Tech Pants        Hoodies   Dry Fit V-Nick          NaN              NaN              NaN              NaN
14    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants   Dry Fit V-Nick   Rash Guard        Tech Pants              NaN              NaN
15  Swimming Shirt      Soccer Shoe          Hoodies   Dry Fit V-Nick     Tech Pants       Rash Guard          NaN              NaN              NaN              NaN
16    Running Shoe            Socks              NaN              NaN            NaN              NaN          NaN              NaN              NaN              NaN
17           Socks       Sweatshirts     Modern Pants      Soccer Shoe        Hoodies       Rash Guard   Tech Pants   Dry Fit V-Nick              NaN              NaN
18    Running Shoe   Swimming Shirt       Rash Guard              NaN            NaN              NaN          NaN              NaN              NaN              NaN
19    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants      Soccer Shoe      Hoodies       Tech Pants       Rash Guard   Dry Fit V-Nick
Please, input your Min. Support
```

```
(' Shams', ' Bedspreads')  --->  [' Kids Bedding']    confidence =  1.0
(' Kids Bedding', ' Bedspreads')  --->  [' Shams']    confidence =  1.0
(' Quilts', ' Shams', ' Kids Bedding')  --->  [' Bedspreads']    confidence =  1.0
(' Quilts', ' Shams', ' Bedspreads')  --->  [' Kids Bedding']    confidence =  1.0
(' Quilts', ' Kids Bedding', ' Bedspreads')  --->  [' Shams']    confidence =  1.0
(' Shams', ' Kids Bedding', ' Bedspreads')  --->  [' Quilts']    confidence =  1.0
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 4
       transaction       Unnamed: 1       Unnamed: 2       Unnamed: 3     Unnamed: 4       Unnamed: 5   Unnamed: 6       Unnamed: 7       Unnamed: 8       Unnamed: 9
0     Running Shoe            Socks      Sweatshirts     Modern Pants            NaN              NaN          NaN              NaN              NaN              NaN
1     Running Shoe            Socks      Sweatshirts              NaN            NaN              NaN          NaN              NaN              NaN              NaN
2     Running Shoe            Socks      Sweatshirts     Modern Pants            NaN              NaN          NaN              NaN              NaN              NaN
3     Running Shoe       Sweatshirts     Modern Pants              NaN            NaN              NaN          NaN              NaN              NaN              NaN
4     Running Shoe            Socks      Sweatshirts     Modern Pants     Soccer Shoe              NaN          NaN              NaN              NaN              NaN
5     Running Shoe            Socks      Sweatshirts              NaN            NaN              NaN          NaN              NaN              NaN              NaN
6     Running Shoe            Socks      Sweatshirts     Modern Pants      Tech Pants       Rash Guard      Hoodies              NaN              NaN              NaN
7    Swimming Shirt            Socks      Sweatshirts              NaN            NaN              NaN          NaN              NaN              NaN              NaN
8    Swimming Shirt       Rash Guard   Dry Fit V-Nick          Hoodies      Tech Pants              NaN          NaN              NaN              NaN              NaN
9    Swimming Shirt       Rash Guard              Dry              NaN            NaN              NaN          NaN              NaN              NaN              NaN
10  Swimming Shirt       Rash Guard   Dry Fit V-Nick              NaN            NaN              NaN          NaN              NaN              NaN              NaN
11    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants      Soccer Shoe   Rash Guard          Hoodies       Tech Pants   Dry Fit V-Nick
12    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants      Soccer Shoe   Rash Guard        Tech Pants   Dry Fit V-Nick          Hoodies
13    Running Shoe   Swimming Shirt       Rash Guard       Tech Pants        Hoodies   Dry Fit V-Nick          NaN              NaN              NaN              NaN
14    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants   Dry Fit V-Nick   Rash Guard        Tech Pants              NaN              NaN
15  Swimming Shirt      Soccer Shoe          Hoodies   Dry Fit V-Nick     Tech Pants       Rash Guard          NaN              NaN              NaN              NaN
16    Running Shoe            Socks              NaN              NaN            NaN              NaN          NaN              NaN              NaN              NaN
17           Socks       Sweatshirts     Modern Pants      Soccer Shoe        Hoodies       Rash Guard   Tech Pants   Dry Fit V-Nick              NaN              NaN
18    Running Shoe   Swimming Shirt       Rash Guard              NaN            NaN              NaN          NaN              NaN              NaN              NaN
19    Running Shoe   Swimming Shirt            Socks      Sweatshirts   Modern Pants      Soccer Shoe      Hoodies       Tech Pants       Rash Guard   Dry Fit V-Nick
Please, input your Min. Support
0.3
Please, input your Min. confidence
0.4
frequent patterns
 [(' Socks',), (' Sweatshirts',), (' Rash Guard',), (' Swimming Shirt',), (' Soccer Shoe',), (' Socks', ' Sweatshirts'), (' Socks', ' Rash Guard'), (' Socks', ' Swimming Shirt'), (' Socks', ' Soccer Shoe'
), (' Sweatshirts', ' Swimming Shirt'), (' Rash Guard', ' Soccer Shoe'), (' Socks', ' Sweatshirts', ' Swimming Shirt'), (' Socks', ' Rash Guard', ' Soccer Shoe')]

Association rules
(' Socks',)  --->  [' Sweatshirts']    confidence =  0.5
(' Sweatshirts',)  --->  [' Socks']    confidence =  0.5
(' Socks',)  --->  [' Rash Guard']    confidence =  0.5
(' Rash Guard',)  --->  [' Socks']    confidence =  1.0
(' Socks',)  --->  [' Swimming Shirt']    confidence =  0.5
(' Swimming Shirt',)  --->  [' Socks']    confidence =  1.0
(' Socks',)  --->  [' Soccer Shoe']    confidence =  0.5
(' Soccer Shoe',)  --->  [' Socks']    confidence =  1.0
(' Sweatshirts',)  --->  [' Swimming Shirt']    confidence =  0.5
(' Swimming Shirt',)  --->  [' Sweatshirts']    confidence =  1.0
(' Rash Guard',)  --->  [' Soccer Shoe']    confidence =  1.0
(' Soccer Shoe',)  --->  [' Rash Guard']    confidence =  1.0
(' Socks', ' Sweatshirts')  --->  [' Swimming Shirt']    confidence =  1.0
(' Socks', ' Swimming Shirt')  --->  [' Sweatshirts']    confidence =  1.0
(' Sweatshirts', ' Swimming Shirt')  --->  [' Socks']    confidence =  1.0
(' Socks', ' Rash Guard')  --->  [' Soccer Shoe']    confidence =  1.0
(' Socks', ' Soccer Shoe')  --->  [' Rash Guard']    confidence =  1.0
(' Rash Guard', ' Soccer Shoe')  --->  [' Socks']    confidence =  1.0
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$
```

5.GENERIC

```
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 5
   transaction Unnamed: 1 Unnamed: 2 Unnamed: 3 Unnamed: 4
0            A          B          C          D        NaN
1            B          C        NaN        NaN        NaN
2            E          F          I        NaN        NaN
3            A          G          J        NaN        NaN
4            D          H          I        NaN        NaN
5            I          J        NaN        NaN        NaN
6            B          D          F          H          J
7            C          D          E        NaN        NaN
8            A          E          J        NaN        NaN
9            F          G          H          I        NaN
10           B          C          D          E          F
11           D          G          H          I          J
12           B          F        NaN        NaN        NaN
13           C          D        NaN        NaN        NaN
14           F          G          H          I        NaN
15           H          I          J        NaN        NaN
16           A          B          C        NaN        NaN
17           A          B          C          D          E
18           G          H          I          J        NaN
19           E          F          G          H          I
Please, input your Min. Support
```

```
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$ python Midterm_Project_Data_Mining.py
Please, Select your Dataset for
 1 Amazon.
 2 bestbuy.
 3 K_mart.
 4 Nike.
 5 Generic.
 5
   transaction Unnamed: 1 Unnamed: 2 Unnamed: 3 Unnamed: 4
0            A          B          C          D        NaN
1            B          C        NaN        NaN        NaN
2            E          F          I        NaN        NaN
3            A          G          J        NaN        NaN
4            D          H          I        NaN        NaN
5            I          J        NaN        NaN        NaN
6            B          D          F          H          J
7            C          D          E        NaN        NaN
8            A          E          J        NaN        NaN
9            F          G          H          I        NaN
10           B          C          D          E          F
11           D          G          H          I          J
12           B          F        NaN        NaN        NaN
13           C          D        NaN        NaN        NaN
14           F          G          H          I        NaN
15           H          I          J        NaN        NaN
16           A          B          C        NaN        NaN
17           A          B          C          D          E
18           G          H          I          J        NaN
19           E          F          G          H          I
Please, input your Min. Support
0.1
Please, input your Min. confidence
0.15
frequent patterns
 [('B',), ('C',), ('F',), ('G',), ('H',), ('J',), ('D',), ('E',), ('I',), ('B', 'G'), ('B', 'E'), ('C', 'F'), ('C', 'D'), ('F', 'D'), ('G', 'H'), ('G', 'E'), ('B', 'G', 'E'), ('C', 'F', 'D')]

Association rules
('B',)  --->  ['G']    confidence =  1.0
('G',)  --->  ['B']    confidence =  0.3333333333333333
('B',)  --->  ['E']    confidence =  1.0
('E',)  --->  ['B']    confidence =  1.0
('C',)  --->  ['F']    confidence =  1.0
('F',)  --->  ['C']    confidence =  0.5
('C',)  --->  ['D']    confidence =  1.0
('D',)  --->  ['C']    confidence =  0.5
('F',)  --->  ['D']    confidence =  0.5
('D',)  --->  ['F']    confidence =  0.5
('G',)  --->  ['H']    confidence =  0.3333333333333333
('H',)  --->  ['G']    confidence =  0.5
('G',)  --->  ['E']    confidence =  0.3333333333333333
('E',)  --->  ['G']    confidence =  1.0
('B', 'G')  --->  ['E']    confidence =  1.0
('B', 'E')  --->  ['G']    confidence =  1.0
('G', 'E')  --->  ['B']    confidence =  1.0
('C', 'F')  --->  ['D']    confidence =  1.0
('C', 'D')  --->  ['F']    confidence =  1.0
('F', 'D')  --->  ['C']    confidence =  1.0
(base) tejas-MacBook-Air:Kulkarni_Teja_MidtermProj teja$
```

CONCLUSION :

Association rule mining algorithms such as Apriori Algorithim is good way to finding simple associations between data items (Transactions). We could conclude that no rules found happens to have a very low support, so that we do not have enough information on the relationship between its items.

For a beginner, it provides an easy way to understand the association rules and quickly apply for market basket analysis. Although there are limitations such as

- Time consuming: when we have huge transactional database; program take much time; faced this problem.

- Inefficiency when memory capacity is limited with large number of transactions.

GITHUB LINK :

https://github.com/TejaKulkarni/Apriori_Algorithm