



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

Experiment Number: 9					
Date of Performance:					
Date of Submission:					
Program Execution/ formation/ correction/ ethical practices (07)	Documentation (02)	Timely Submission (03)	Viva Answer to sample questions (03)	Experiment Total (15)	Sign



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

Experiment 9

Aim: Implementation of Association Rule Mining algorithm (Apriori)

Lab outcome: CSL 503.4: Implement Association rule mining and web mining algorithms.

Problem Statement: To implement Association Rule Mining algorithm.

Theory:

Association rule mining finds interesting associations and relationships among large sets of data items. This rule shows how frequently a itemset occurs in a transaction. A typical example is a Market Based Analysis.

Market Based Analysis is one of the key techniques used by large relations to show associations between items. It allows retailers to identify relationships between the items that people buy together frequently.

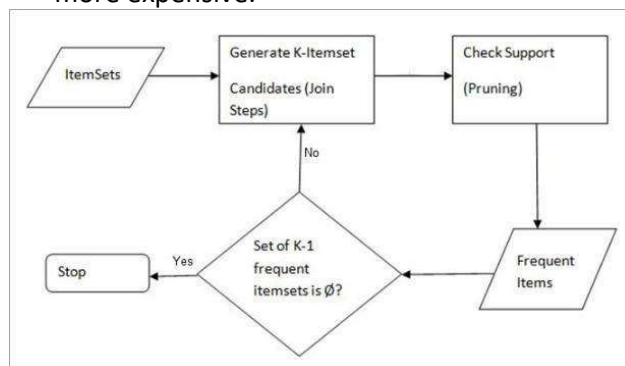
Apriori algorithm is given by R. Agrawal and R. Srikant in 1994 for finding frequent itemsets in a dataset for Boolean association rule. Name of the algorithm is Apriori because it uses prior knowledge of frequent itemset properties. We apply an iterative approach or level-wise search where k-frequent itemsets are used to find k+1 itemsets.

Advantages of Apriori Algorithm

- It is used to calculate large itemsets.
- Simple to understand and apply.

Disadvantages of Apriori Algorithms

- Apriori algorithm is an expensive method to find support since the calculation has to pass through the whole database.
- Sometimes, you need a huge number of candidate rules, so it becomes computationally more expensive.



Algorithm:

- Step 1. Computing the support for each individual item
- Step 2. Deciding on the support threshold
- Step 3. Selecting the frequent items
- Step 4. Finding the support of the frequent itemsets
- Step 5. Repeat for larger sets
- Step 6. Generate Association Rules and compute confidence
- Step 7. Compute lift



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

Program Listing and Output:

Code:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from apyori import apriori
store_data = pd.read_csv("store_data.csv", header=None)
print(store_data.shape)
records = []
for i in range(1, 7501):
    records.append([str(store_data.values[i, j]) for j in range(0, 20)])
print(type(records))
# <class 'list'>
association_rules = apriori(records, min_support=0.0045, min_confidence=0.2,
min_lift=3, min_length=2)
association_results = list(association_rules)
print("There are {} Relation derived.".format(len(association_results)))
for i in range(0, len(association_results)):
    print(association_results[i][0])
for item in association_results:
    # first index of the inner list
    # Contains base item and add item
    pair = item[0]
    items = [x for x in pair]
    print("Rule: " + items[0] + " -> " + items[1])

    # second index of the inner list
    print("Support: " + str(item[1]))

    # third index of the list located at 0th
    # of the third index of the inner list

    print("Confidence: " + str(item[2][0][2]))
    print("Lift: " + str(item[2][0][3]))
    print("=====")
```

Output:-



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**

Chembur, Mumbai - 400 088

UG Program in Cyber Security

```
(7501, 20)
<class 'list'>
There are 48 Relation derived.
frozenset({'light cream', 'chicken'})
frozenset({'mushroom cream sauce', 'escalope'})
frozenset({'escalope', 'pasta'})
frozenset({'ground beef', 'herb & pepper'})
frozenset({'ground beef', 'tomato sauce'})
frozenset({'whole wheat pasta', 'olive oil'})
frozenset({'shrimp', 'pasta'})
frozenset({'light cream', 'chicken', 'nan'})
frozenset({'chocolate', 'frozen vegetables', 'shrimp'})
frozenset({'ground beef', 'cooking oil', 'spaghetti'})
frozenset({'mushroom cream sauce', 'escalope', 'nan'})
frozenset({'pasta', 'escalope', 'nan'})
frozenset({'ground beef', 'frozen vegetables', 'spaghetti'})
frozenset({'frozen vegetables', 'milk', 'olive oil'})
frozenset({'mineral water', 'frozen vegetables', 'shrimp'})
frozenset({'frozen vegetables', 'olive oil', 'spaghetti'})
frozenset({'frozen vegetables', 'shrimp', 'spaghetti'})
frozenset({'frozen vegetables', 'tomatoes', 'spaghetti'})
frozenset({'ground beef', 'grated cheese', 'spaghetti'})
frozenset({'ground beef', 'mineral water', 'herb & pepper'})
frozenset({'ground beef', 'nan', 'herb & pepper'})
frozenset({'ground beef', 'herb & pepper', 'spaghetti'})
frozenset({'ground beef', 'milk', 'olive oil'})
frozenset({'ground beef', 'nan', 'tomato sauce'})
frozenset({'ground beef', 'shrimp', 'spaghetti'})
frozenset({'milk', 'olive oil', 'spaghetti'})
frozenset({'mineral water', 'soup', 'olive oil'})
frozenset({'whole wheat pasta', 'nan', 'olive oil'})
frozenset({'shrimp', 'nan', 'pasta'})
frozenset({'pancakes', 'olive oil', 'spaghetti'})
frozenset({'chocolate', 'frozen vegetables', 'shrimp', 'nan'})
frozenset({'ground beef', 'cooking oil', 'nan', 'spaghetti'})
frozenset({'ground beef', 'frozen vegetables', 'nan', 'spaghetti'})
frozenset({'mineral water', 'frozen vegetables', 'milk', 'spaghetti'})
frozenset({'frozen vegetables', 'milk', 'nan', 'olive oil'})
frozenset({'mineral water', 'frozen vegetables', 'shrimp', 'nan'})
frozenset({'frozen vegetables', 'nan', 'olive oil', 'spaghetti'})
```



Mahavir Education Trust's
**SHAH & ANCHOR KUTCHHI ENGINEERING
COLLEGE**
Chembur, Mumbai - 400 088
UG Program in Cyber Security

```
=====
Rule: mushroom cream sauce -> escalope
Support: 0.005733333333333333
Confidence: 0.30069930069930073
Lift: 3.7903273197390845
=====
Rule: escalope -> pasta
Support: 0.005866666666666667
Confidence: 0.37288135593220345
Lift: 4.700185158809287
=====
Rule: ground beef -> herb & pepper
Support: 0.016
Confidence: 0.3234501347708895
Lift: 3.2915549671393096
=====
Rule: ground beef -> tomato sauce
Support: 0.005333333333333333
Confidence: 0.37735849056603776
Lift: 3.840147461662528
=====
Rule: whole wheat pasta -> olive oil
Support: 0.008
Confidence: 0.2714932126696833
Lift: 4.130221288078346
=====
Rule: shrimp -> pasta
Support: 0.005066666666666666
Confidence: 0.3220338983050848
Lift: 4.514493901473151
=====
```

Conclusion: Here we implemented Association Rule Mining algorithm (Apriori)