

## Experiment no. 9

Aim: To implement Association Rule Mining Algorithm (Apriori).

Requirements: Windows OS and Weka Tool.

Problem Statement: To implement Apriori algorithm on Supermarket data set using Weka Tool.

Theory:

Apriori Algorithm:

The Apriori algorithm uses frequent item sets to generate association rules, and it is designed to work on the databases that contain transactions. With the help of these association rule, it determines how strongly or how weakly two objects are connected. This algorithm uses a breadth-first search and Hash Tree to calculate the item set associations efficiently. It is the iterative process for finding the frequent item sets from the large dataset.

This algorithm was given by the R. Agrawal and Srikant in the year 1994. It is mainly used for *market basket analysis* and helps to find those products that can be bought together. It can also be used in the healthcare field to find drug reactions for patients.

What is Frequent Item set?

Frequent item sets are those items whose support is greater than the threshold value or user-specified minimum support. It means if A & B are the frequent item sets together, then individually A and B should also be the frequent item set.

Suppose there are the two transactions: A= {1,2,3,4,5}, and B= {2,3,7}, in these two transactions, 2 and 3 are the frequent item sets.

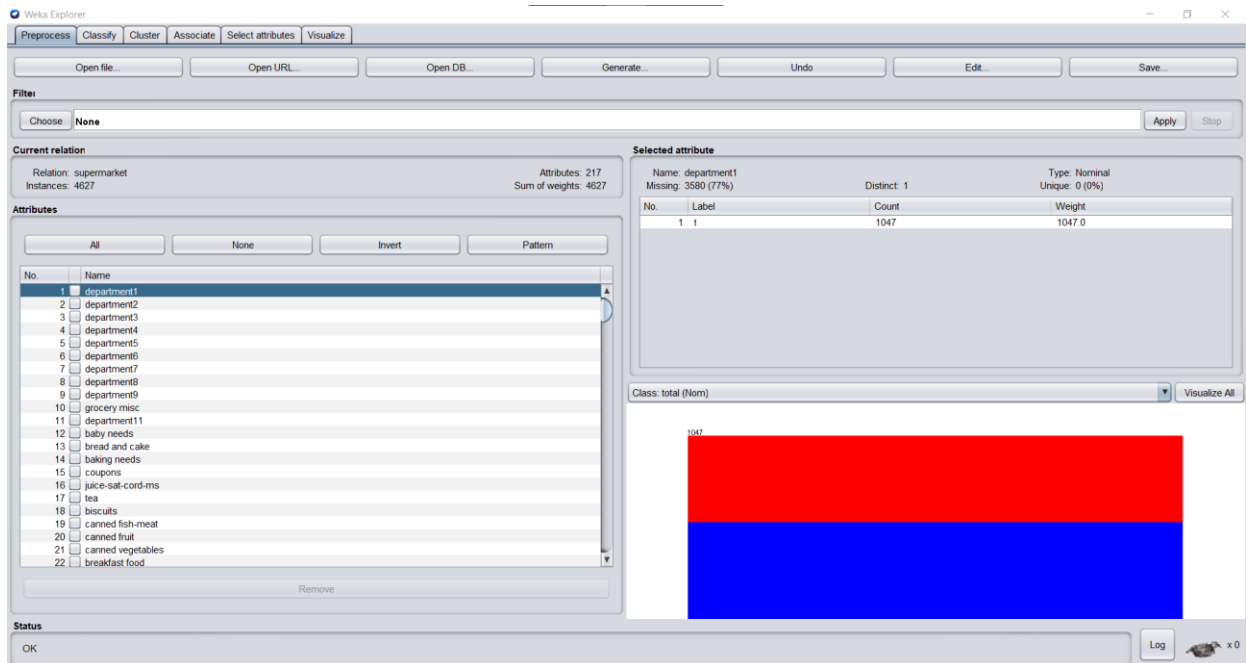
Steps for Apriori Algorithm:

Step-1: Determine the support of item sets in the transactional database, and select the minimum support and confidence.

Step-2: Take all supports in the transaction with higher support value than the minimum or selected support value.

Step-3: Find all the rules of these subsets that have higher confidence value than the threshold or minimum confidence.

Step-4: Sort the rules as the decreasing order of lift.

Output:Supermarket data set:Parameter of Apriori algorithm:

weka.gui.GenericObjectEditor

weka.associations.Apriori

**About**

Class implementing an Apriori-type algorithm.

More

Capabilities

car: False

classIndex: -1

delta: 0.05

doNotCheckCapabilities: False

lowerBoundMinSupport: 0.1

metric Type: Confidence

minMetric: 0.9

numRules: 10

outputItemSets: False

removeAllMissingCols: False

significanceLevel: -1.0

treatZeroAsMissing: False

upperBoundMinSupport: 1.0

verbose: False

Open... Save... OK Cancel

Result:

The screenshot shows the Weka Explorer interface with the Apriori algorithm selected in the 'Associate' tab. The 'Result list (right-click...)' on the left shows '23:23:12 - Apriori' selected. The 'Associator output' pane displays the following results:

```

Apriori
=====
Minimum support: 0.15 (694 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17

Generated sets of large itemsets:

Size of set of large itemsets L(1): 44
Size of set of large itemsets L(2): 380
Size of set of large itemsets L(3): 910
Size of set of large itemsets L(4): 633
Size of set of large itemsets L(5): 105
Size of set of large itemsets L(6): 1

Best rules found:

1. biscuits=t frozen foods=t fruit=t total=high 788 ==> bread and cake=t 723    <conf:(0.92)> lift:(1.27) lev:(0.03) [155] conv:(3.35)
2. baking needs=t biscuits=t fruit=t total=high 760 ==> bread and cake=t 696    <conf:(0.92)> lift:(1.27) lev:(0.03) [149] conv:(3.28)
3. baking needs=t frozen foods=t fruit=t total=high 770 ==> bread and cake=t 705    <conf:(0.92)> lift:(1.27) lev:(0.03) [150] conv:(3.27)
4. biscuits=t fruit=t vegetables=t total=high 815 ==> bread and cake=t 746    <conf:(0.92)> lift:(1.27) lev:(0.03) [159] conv:(3.26)
5. party snack foods=t fruit=t total=high 854 ==> bread and cake=t 779    <conf:(0.91)> lift:(1.27) lev:(0.04) [164] conv:(3.15)
6. biscuits=t frozen foods=t vegetables=t total=high 797 ==> bread and cake=t 725    <conf:(0.91)> lift:(1.26) lev:(0.03) [151] conv:(3.06)
7. baking needs=t biscuits=t vegetables=t total=high 772 ==> bread and cake=t 701    <conf:(0.91)> lift:(1.26) lev:(0.03) [145] conv:(3.01)
8. biscuits=t fruit=t total=high 954 ==> bread and cake=t 866    <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(3)
9. frozen foods=t fruit=t vegetables=t total=high 834 ==> bread and cake=t 757    <conf:(0.91)> lift:(1.26) lev:(0.03) [156] conv:(3)
10. frozen foods=t fruit=t total=high 969 ==> bread and cake=t 877    <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(2.92)

```

The 'Status' bar at the bottom shows 'OK' and a 'Log' button.

Conclusion: We have successfully implemented Apriori algorithm on supermarket data set using Weka Tool.