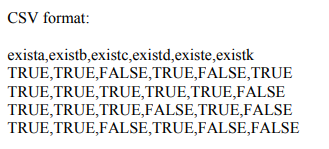
**Experiment 5**

Name: Sneha Savarkar  
UID: 2019120055  
TE EXTC  
Subject : Data Analytics

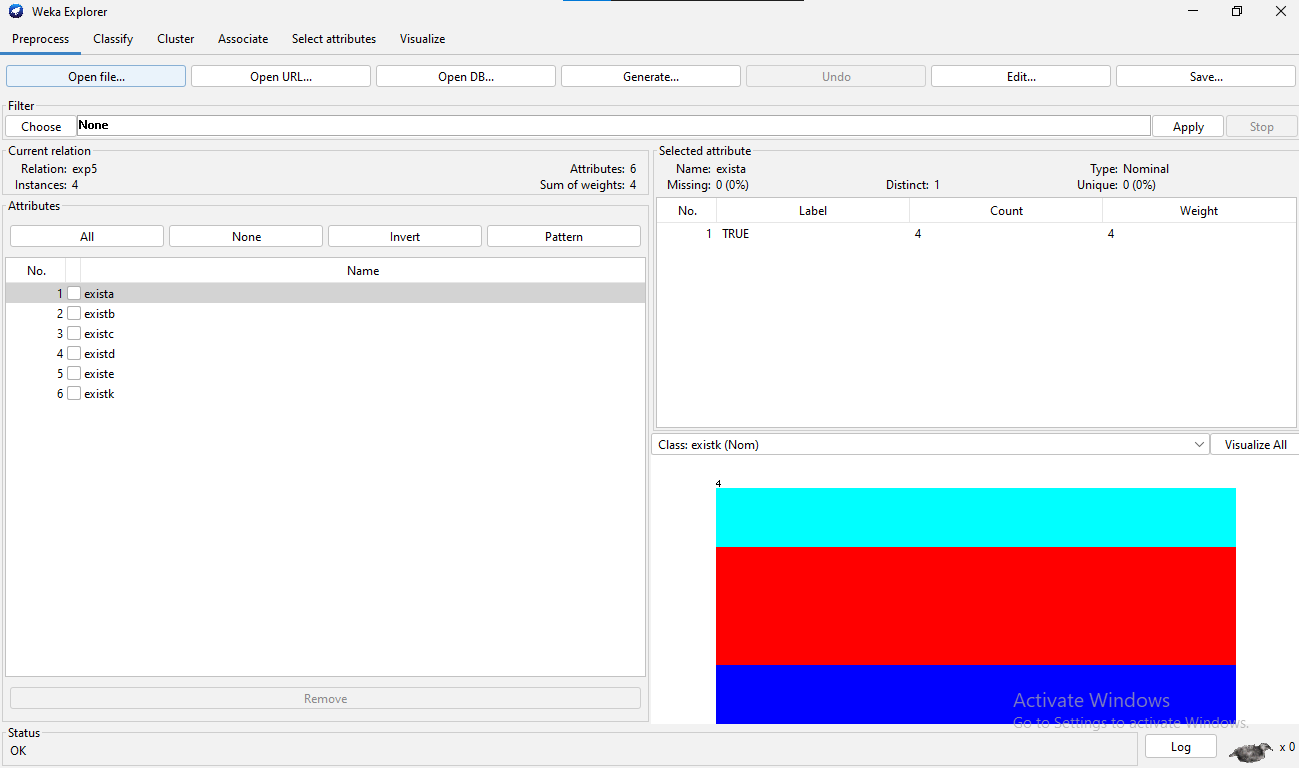
Aim: To apply Apriori algorithm to given dataset Association Rule mining with WEKA

Procedure:

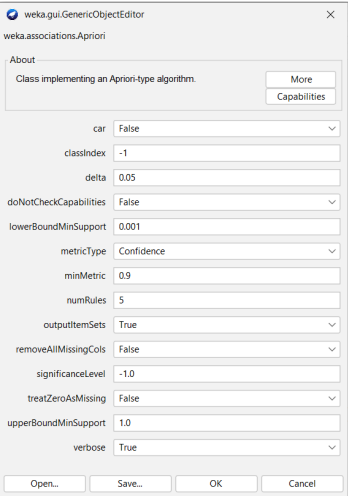
1. Make a CSV File



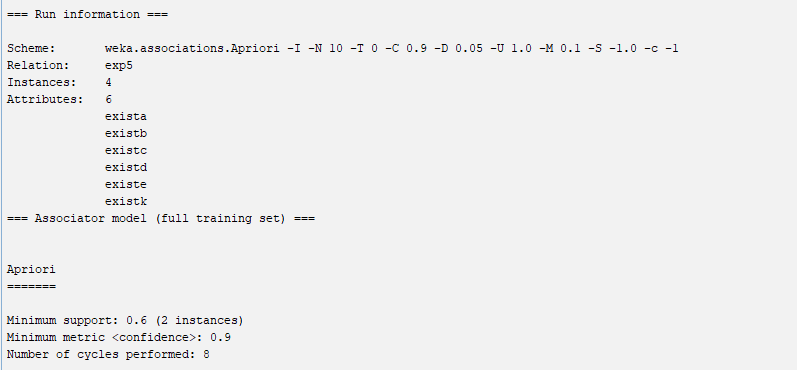
1. In WEKA, open the CSV file in question. After opening, it will look like this

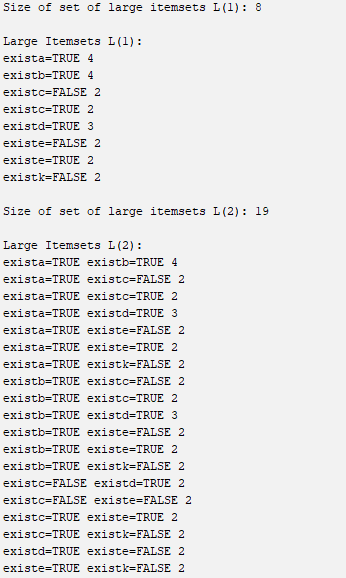


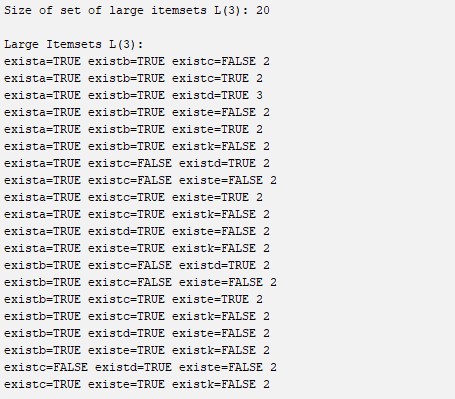
1. After hitting the 'Choose' button, go to the Associate tab and choose 'Apriori' from the drop down menu.
2. Choose Apriori algorithm from the drop-down menu.
3. Double-click the apriori algorithm to bring up an option menu where you may set appropriate values.



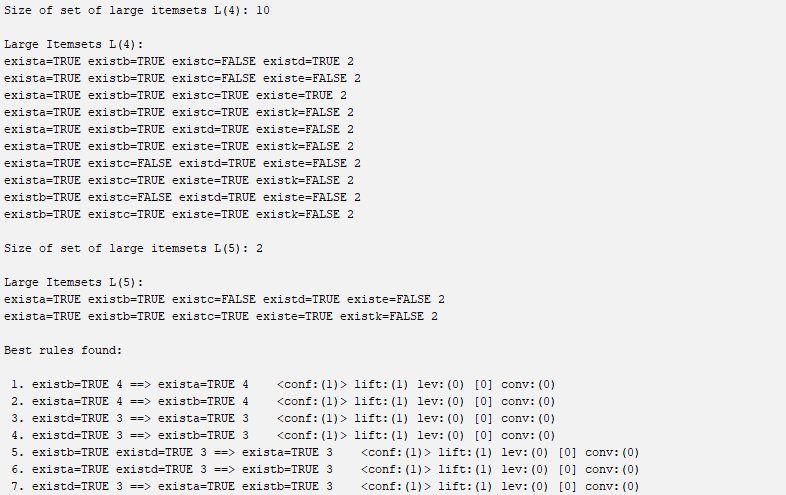
1. Now press start, and WEKA will process the data for us.

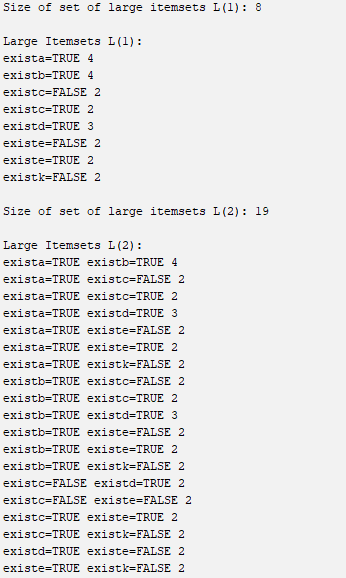


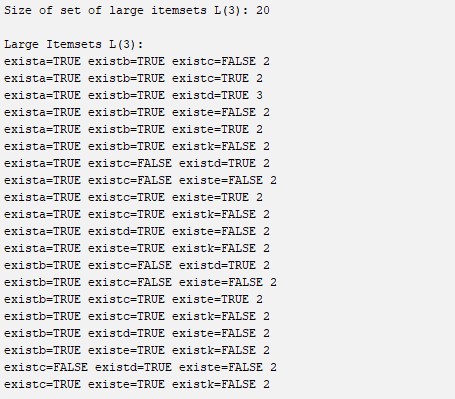




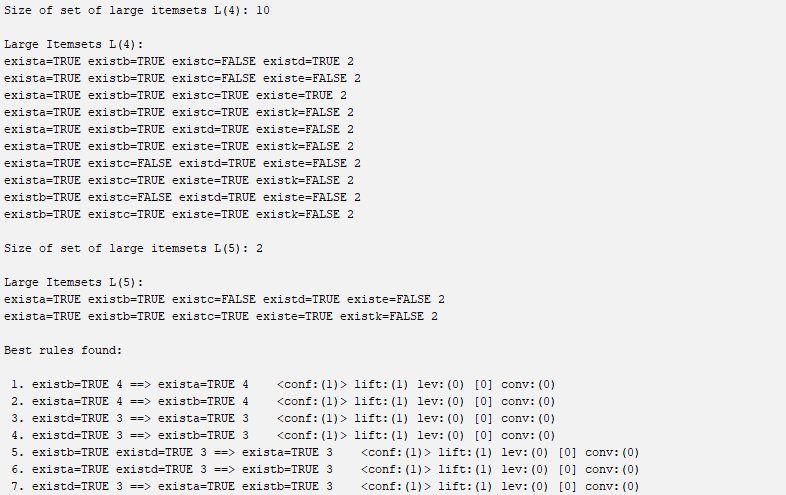
1. The minimum support is 0.6 and minimum confidence is 0.9

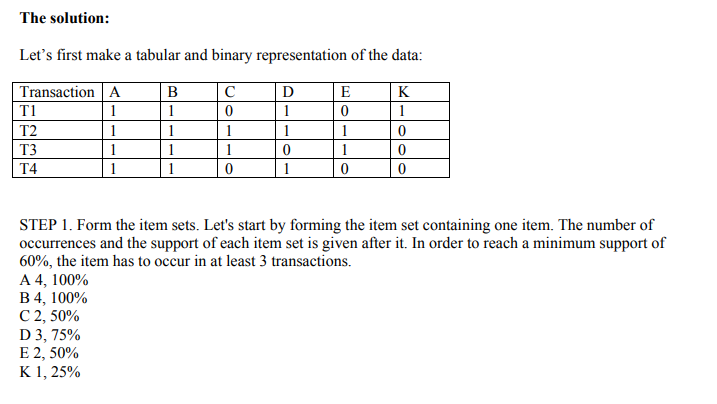


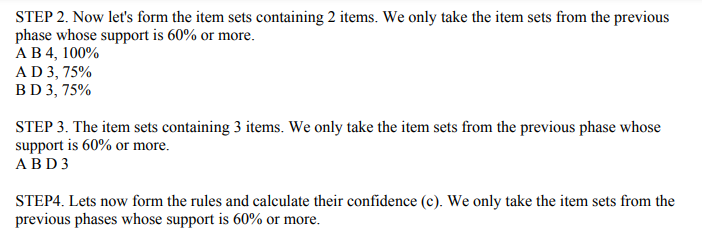


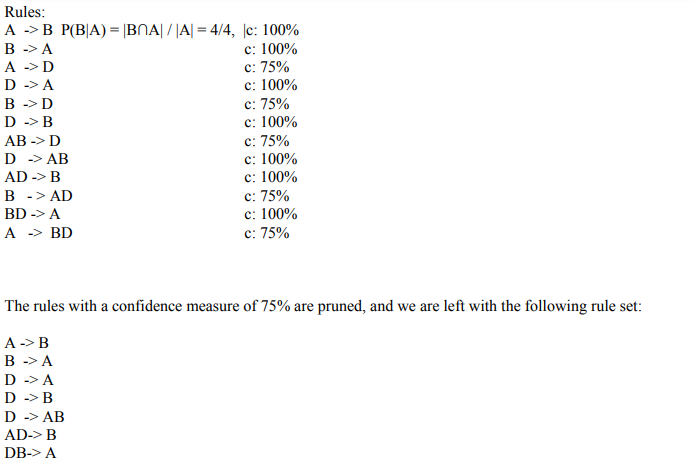


1. After performing all the steps of Apriori we can find out the Best rules









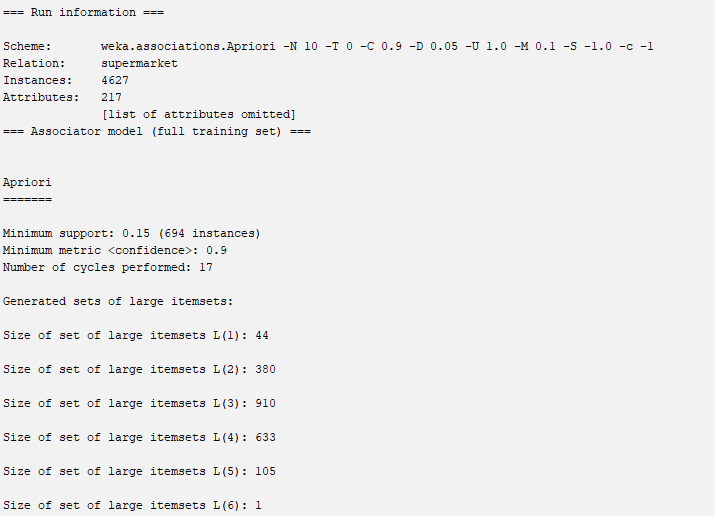
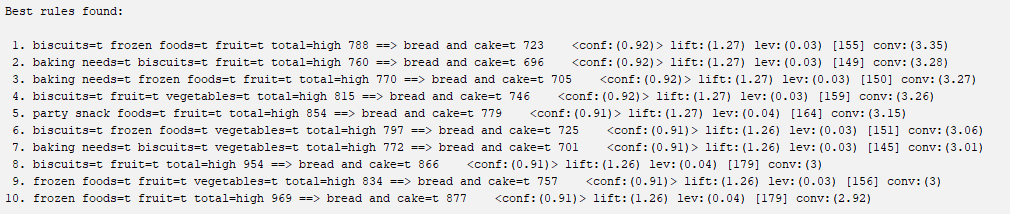
Interpretation:

We can observe that the best rules determined by the manual solution and WEKA are identical. As a result, we can conclude that both answers are accurate and that Apriori has been used.

**Supermarket.arff**

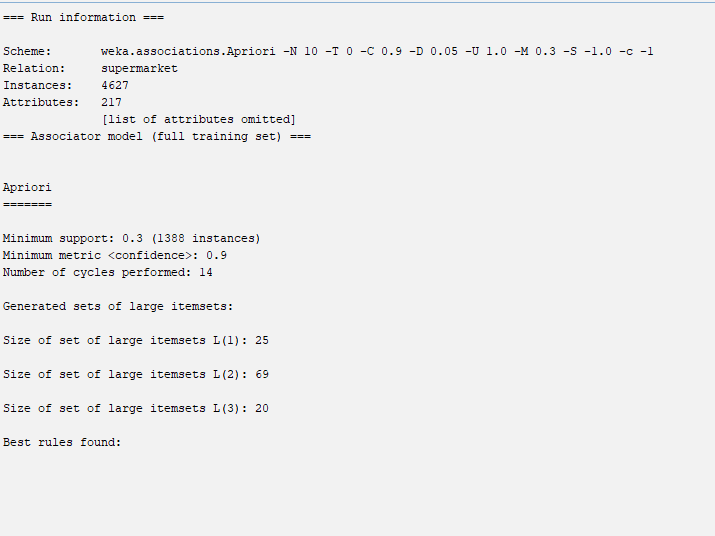
The Apriori Algorithm was run for an inbuilt dataset called supermarket.arff

Case1: The minimum support is 0.15 and confidence is 0.9

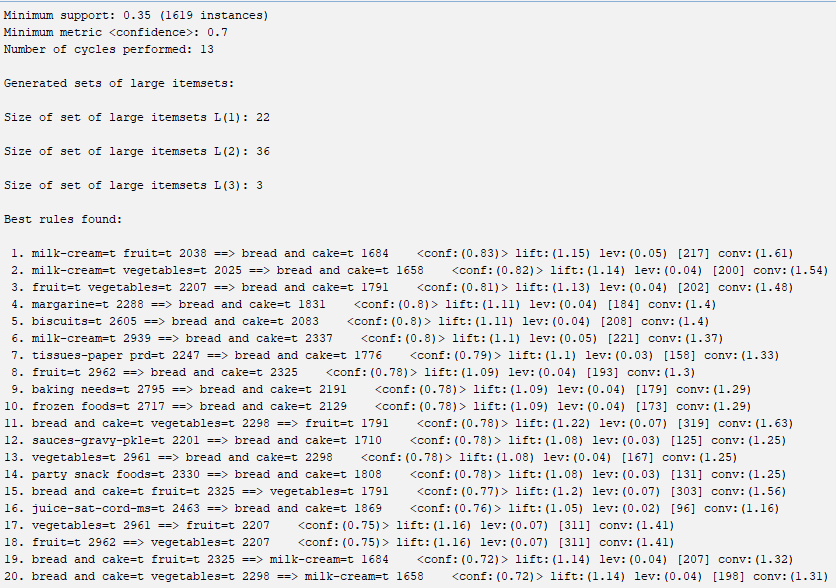
We can see that in this case 10 rules are generated all with the confidence of 0.9 or higher

Case 2: The minimum support is 0.3 and confidence is 0.9.



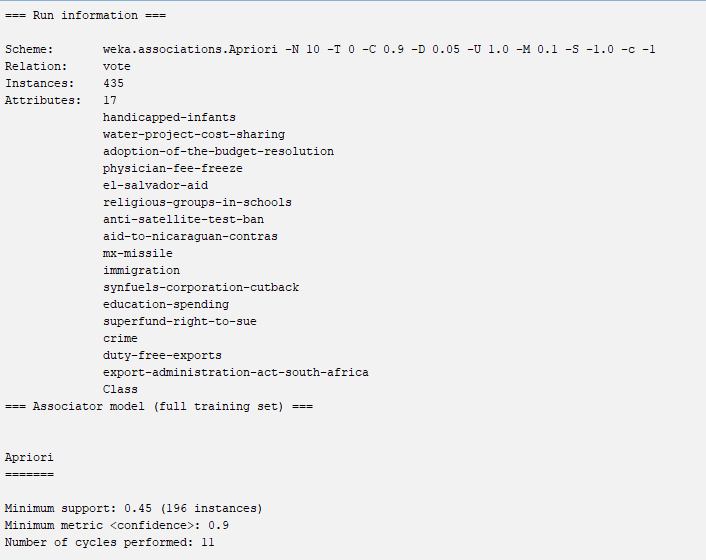
In this case we can see that no rule is generated because the minimum support is high.

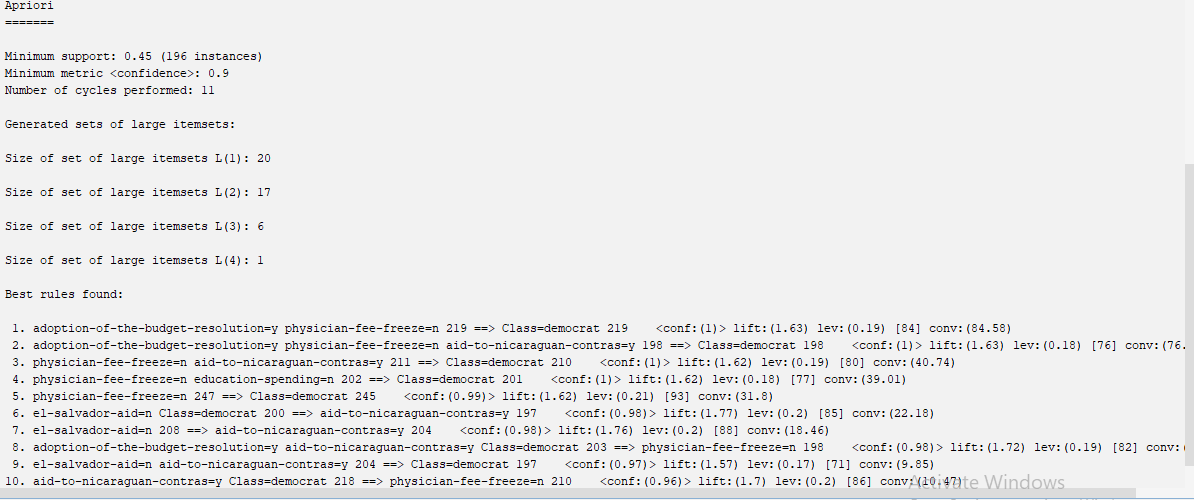
Case 3: The minimum support is 0.3 and confidence is 0.7



In the above solution, we can observe that there are 20 rules generated because the confidence is low.

**The Apriori algorithm was also run on another inbuilt Vote.arff**





The minimal support was 0.45, the confidence level was 0.9, and a total of 10 rules were generated. The democrats are all the classes that are linked here. We will notice several republican class associated rules as we increase the number of republic party entries in our dataset.

Conclusion:

Association rule mining finds new connections and linkages among vast amounts of data. This rule indicates how often an itemset appears in a transaction. We can find rules that forecast the occurrence of an item based on the occurrences of other things in the transaction given a set of transactions.

We can use the Apriori technique to mine the frequent itemset and construct association rules between them. The key constraint is the amount of time necessary to hold a large number of candidate sets with frequent item sets, low minimum support, or huge item sets, implying that it is not an efficient solution for large datasets.