Experiment No.09

A.1 Aim: Study and implementation of APRIORI algorithm using WEKA tool.

PART B

(PART B: TO BE COMPLETED BY STUDENTS)

(Students must submit the soft copy as per the following segments within two hours of the practical. The soft copy must be uploaded on the Blackboard or emailed to the concerned lab in charge faculties at the end of the practical in case there is no Blackboard access available)

Roll No. 50	**Name:** AMEY THAKUR
Class: Comps TE B	**Batch:** B3
Date of Experiment: 28/04/2021	**Date of Submission:** 28/04/2021
Grade:	

B.1 Software Code written by a student:

(Paste your problem statement related to your case study completed during the 2 hours of practice in the lab here)

@relation car

@attribute Type {Subcompact, Compact, Sedan, Luxury}
@attribute Color {Red,Silver,Black}
@attribute Fuel {Petrol,Diesel,CNG,Electric}
@attribute Economic {Yes,No}

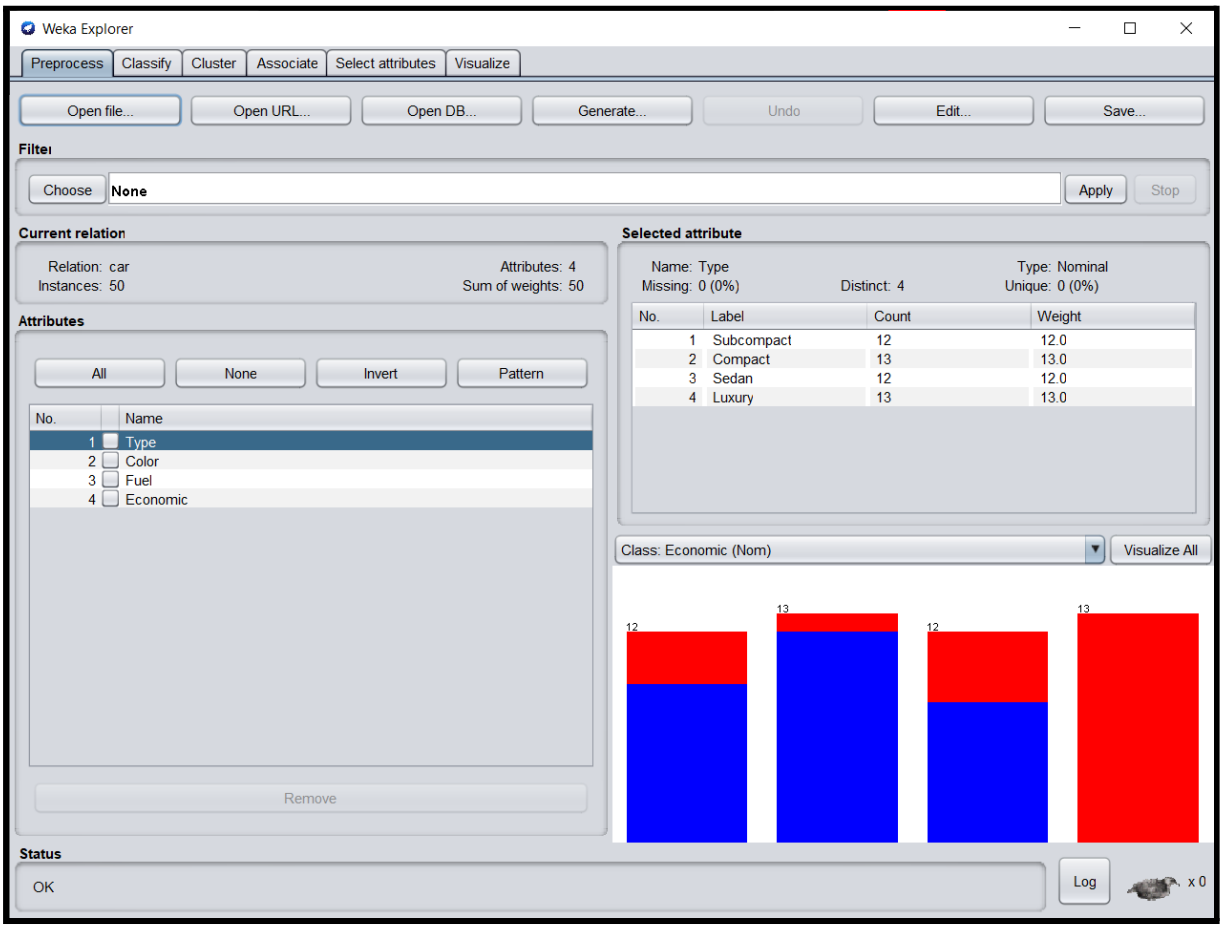
@data

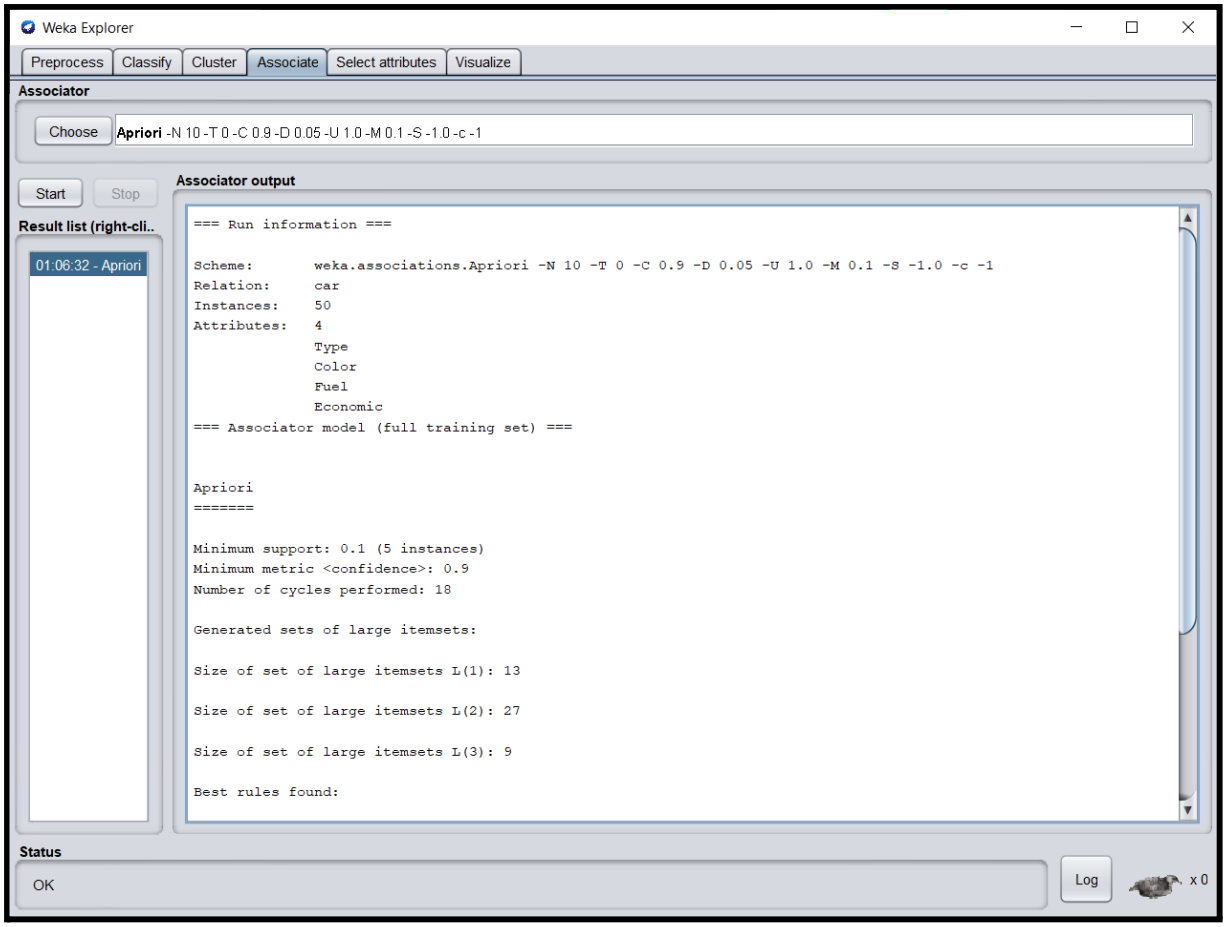
Subcompact	Red	Petrol	Yes
Compact	Black	Diesel	Yes
Compact	Silver	Petrol	Yes
Luxury	Red	Electric	No
Subcompact	Silver	CNG	Yes
Luxury	Red	Petrol	No
Sedan	Silver	Electric	No
Sedan	Black	Diesel	Yes
Subcompact	Black	Electric	Yes
Compact	Red	CNG	Yes
Sedan	Silver	Petrol	Yes
Luxury	Red	Diesel	No
Luxury	Silver	Electric	No

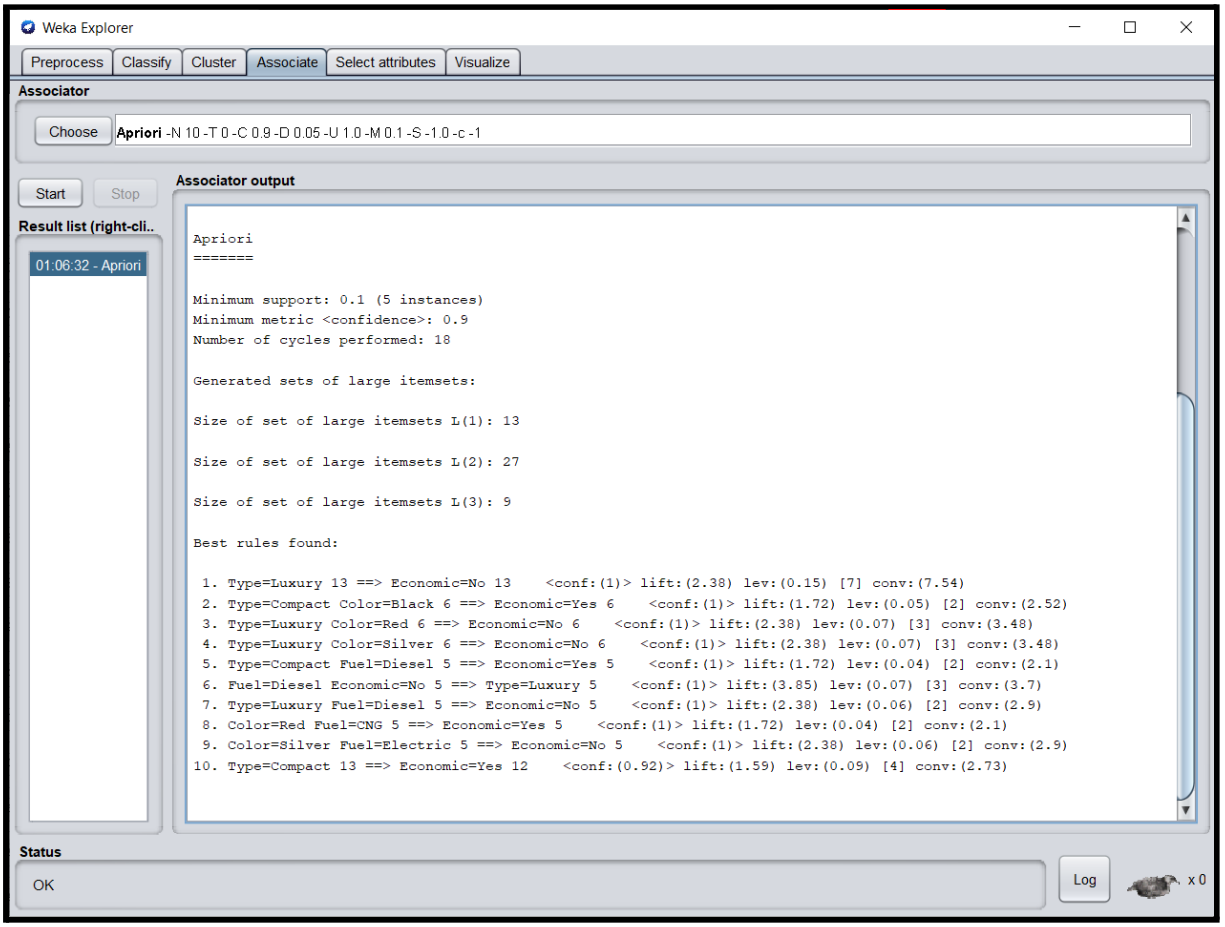
Sedan	Black	CNG	Yes
Compact	Black	Diesel	Yes
Compact	Red	Electric	No
Subcompact	Black	Petrol	Yes
Luxury	Silver	CNG	No
Sedan	Red	Diesel	Yes
Sedan	Silver	Electric	No
Subcompact	Red	CNG	Yes
Compact	Black	Petrol	Yes
Subcompact	Red	Petrol	Yes
Luxury	Silver	CNG	No
Sedan	Black	Diesel	Yes
Sedan	Red	Electric	No
Subcompact	Red	Diesel	Yes
Compact	Silver	Petrol	Yes
Luxury	Red	Diesel	No
Sedan	Black	Petrol	Yes
Subcompact	Black	Petrol	Yes
Subcompact	Red	Electric	No
Compact	Red	Petrol	Yes
Luxury	Silver	CNG	No
Compact	Black	Diesel	Yes
Luxury	Silver	Electric	No
Subcompact	Red	CNG	Yes
Luxury	Red	Diesel	No
Sedan	Black	Electric	No
Compact	Black	CNG	Yes
Compact	Red	CNG	Yes
Sedan	Silver	Petrol	Yes
Luxury	Black	Diesel	No
Subcompact	Silver	Electric	No
Sedan	Red	CNG	Yes
Compact	Black	Diesel	Yes
Luxury	Red	Petrol	No
Subcompact	Black	Electric	No
Luxury	Silver	Diesel	No
Compact	Red	Diesel	Yes

B.2 Input and Output:

(Paste your program input and output in the following format, If there is an error then paste the specific error in the output part. In case of an error with the due permission of the faculty, an extension can be given to submit the error-free code with output in due course of time. Students will be graded accordingly.)




Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Associator
Choose
Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Start
Stop
Result list (right-cli..
01:06:32 - Apriori
Associator output
=== Run information ===
Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Relation: car
Instances: 50
Attributes: 4
Type
Color
Fuel
Economic
=== Associator model (full training set) ===
Apriori
=======
Minimum support: 0.1 (5 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 18
Generated sets of large itemsets:
Size of set of large itemsets L(1): 13
Size of set of large itemsets L(2): 27
Size of set of large itemsets L(3): 9
Best rules found:
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Associator
Choose
Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Start
Stop
Result list (right-cli..
01:06:32 - Apriori
Associator output
Apriori
=======
Minimum support: 0.1 (5 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 18
Generated sets of large itemsets:
Size of set of large itemsets L(1): 13
Size of set of large itemsets L(2): 27
Size of set of large itemsets L(3): 9
Best rules found:
1. Type=Luxury 13 ==> Economic=No 13 <conf:(1)> lift:(2.38) lev:(0.15) [7] conv:(7.54)
2. Type=Compact Color=Black 6 ==> Economic=Yes 6 <conf:(1)> lift:(1.72) lev:(0.05) [2] conv:(2.52)
3. Type=Luxury Color=Red 6 ==> Economic=No 6 <conf:(1)> lift:(2.38) lev:(0.07) [3] conv:(3.48)
4. Type=Luxury Color=Silver 6 ==> Economic=No 6 <conf:(1)> lift:(2.38) lev:(0.07) [3] conv:(3.48)
5. Type=Compact Fuel=Diesel 5 ==> Economic=Yes 5 <conf:(1)> lift:(1.72) lev:(0.04) [2] conv:(2.1)
6. Fuel=Diesel Economic=No 5 ==> Type=Luxury 5 <conf:(1)> lift:(3.85) lev:(0.07) [3] conv:(3.7)
7. Type=Luxury Fuel=Diesel 5 ==> Economic=No 5 <conf:(1)> lift:(2.38) lev:(0.06) [2] conv:(2.9)
8. Color=Red Fuel=CNG 5 ==> Economic=Yes 5 <conf:(1)> lift:(1.72) lev:(0.04) [2] conv:(2.1)
9. Color=Silver Fuel=Electric 5 ==> Economic=No 5 <conf:(1)> lift:(2.38) lev:(0.06) [2] conv:(2.9)
10. Type=Compact 13 ==> Economic=Yes 12 <conf:(0.92)> lift:(1.59) lev:(0.09) [4] conv:(2.73)
Status
OK
Log
x 0

B.3 Observations and learning:

(Students are expected to comment on the output obtained with clear observations and learning for each task/ subpart assigned)

From the given supermarket dataset we have successfully implemented an apriori algorithm on it using the weka tool.

B.4 Conclusion:

(Students must write the conclusion as per the attainment of individual outcome listed above and learning/observation noted in section B.3)

Hence we've successfully implemented the Apriori algorithm in Weka.

B.5 Question of Curiosity

(To be answered by the student based on the practical performed and learning/observations)

1. What is the use of the Apriori algorithm?

Ans:

Apriori is an algorithm for frequent itemset mining and association rule learning over relational databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent itemsets determined by Apriori can be used to determine association rules which highlight general trends in the database, this has applications in many domains:

- ➔ In the business field: Market basket analysis, which is a technique that identifies the strength of association between pairs of products purchased together and identify patterns of co-occurrence
- ➔ In Education Field: Extracting association rules in data mining of admitted students through characteristics and specialities.
- ➔ In the Medical field: For example Analysis of the patient's database.
- ➔ In Forestry: Analysis of probability and intensity of forest fire with the forest fire data.
- ➔ Apriori is used by many companies like Amazon in the Recommender System and by Google for the auto-complete feature.

2. What is Support and Confidence in the Apriori algorithm?

Ans:

The parameters “support” and “confidence” are used in the apriori algorithm.

- ➔ Support refers to items’ frequency of occurrence.
- ➔ Confidence is a conditional probability. Items in a transaction form an item set.

The algorithm begins by identifying frequent, individual items (items with a frequency greater than or equal to the given support) in the database and continues to extend them to larger, frequent itemsets.

$$\text{Support}\ (A) = \frac{\text{Number of transaction in which A appears}}{\text{Total number of transactions}}$$

$$\text{Confidence}\ (A \longrightarrow B) = \frac{\text{Support}(AUB)}{\text{Support}(A)}$$