Exercise 4:

GENERATING ASSOCIATION RULES ON DATASET USING APRIORI AND FP-GROWTH ALGORITHMS WITH WEKA TOOL.

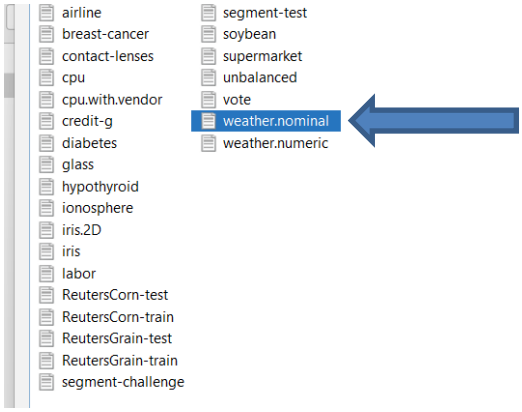
Objectives:

- Generate Association Rules for the Weather.Nominal dataset using the Apriori Algorithm.

Procedure:

1. Setting Up the Environment:

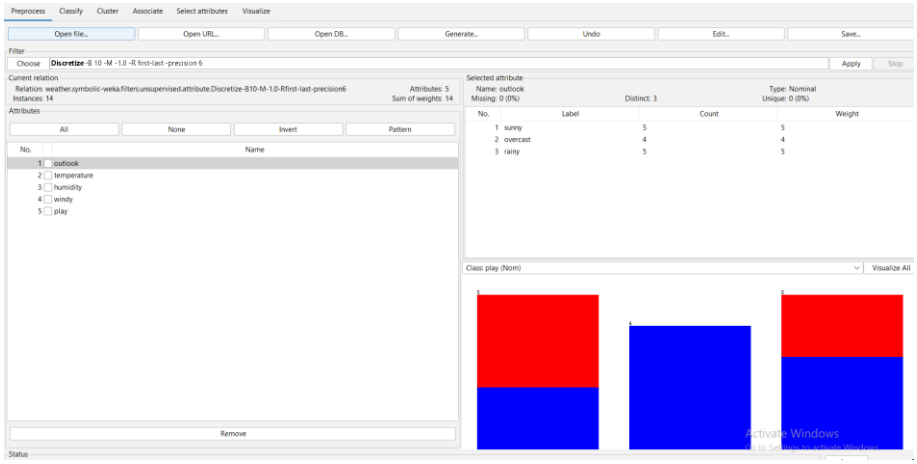
- Install Weka tool on your system if not already installed.
- Load the Weather.Nominal dataset into Weka.



2. Association Rules using Apriori Algorithm:

a) Generating Rules with Default Parameters:

- Open Weka and load the Weather.Nominal dataset.

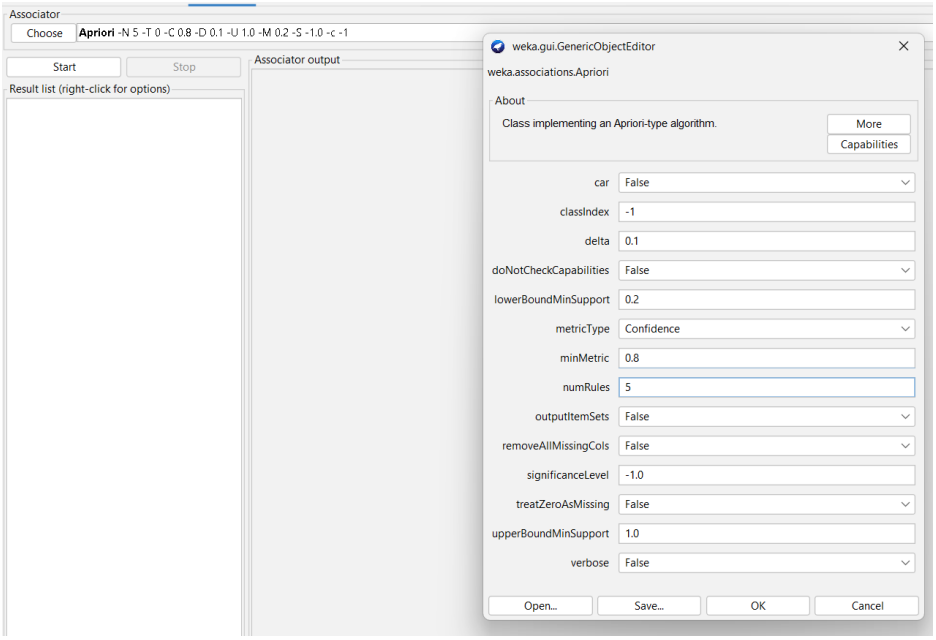


- Apply the Apriori algorithm with default parameters.
- Record and analyze all generated association rules.

b) Adjusting Parameters:

• Set the minimum support range from 20% to 100% with an incremental decrease factor of 10%.

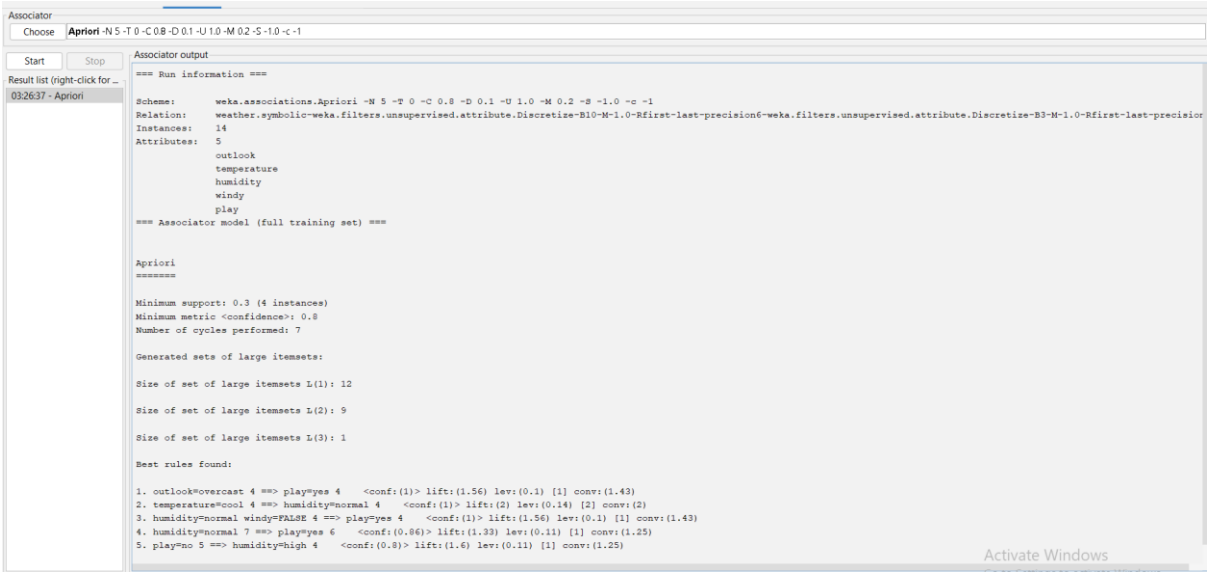
• Set the confidence factor to 80%.



i) **Generating 5 Rules:**

• Apply the Apriori algorithm with the specified parameters.

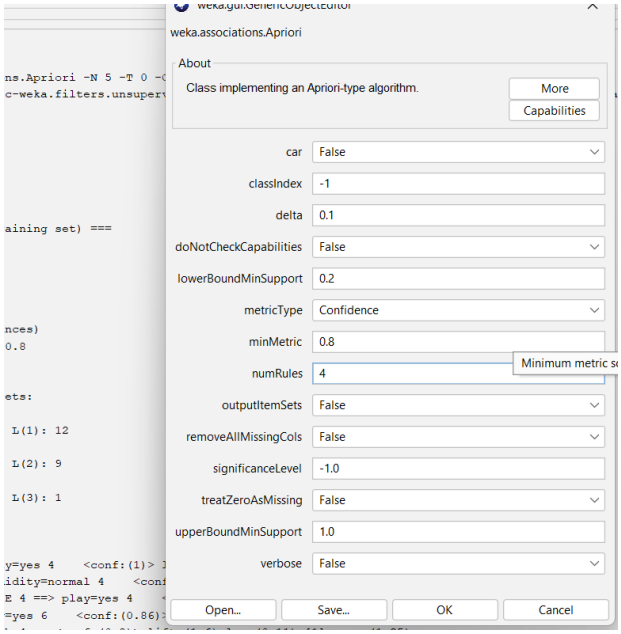
• Note down the generated rules and analyze them.



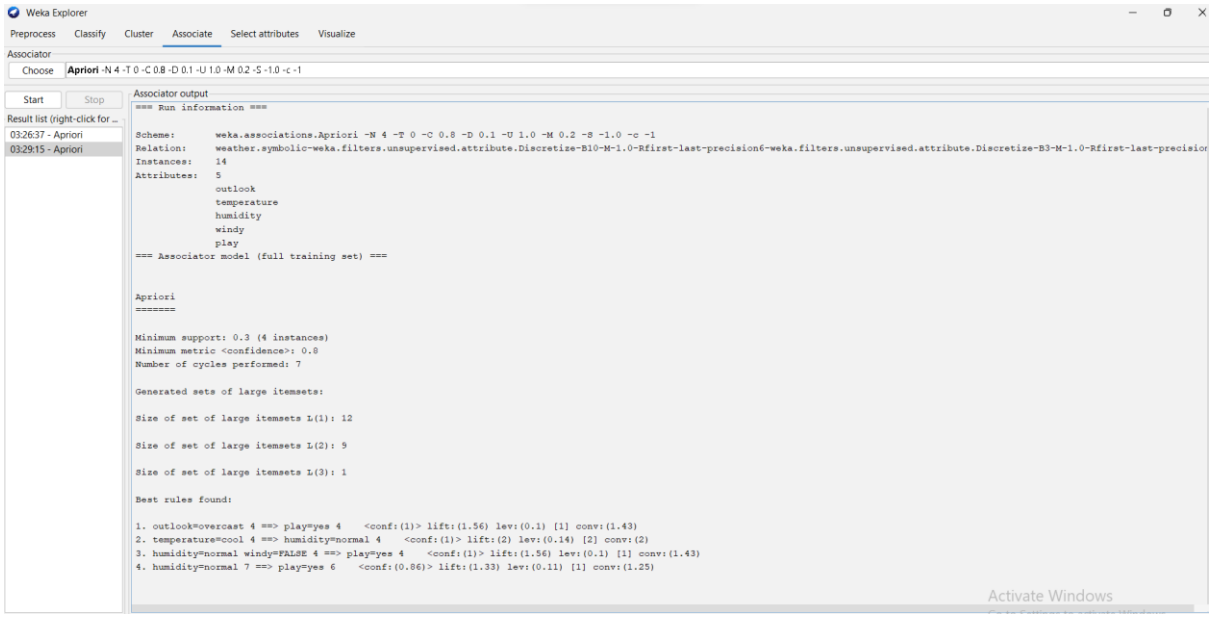
ii) **Generating 4 Rules:**

• Apply the Apriori algorithm with the specified parameters.

• Note down the generated rules and analyze them.



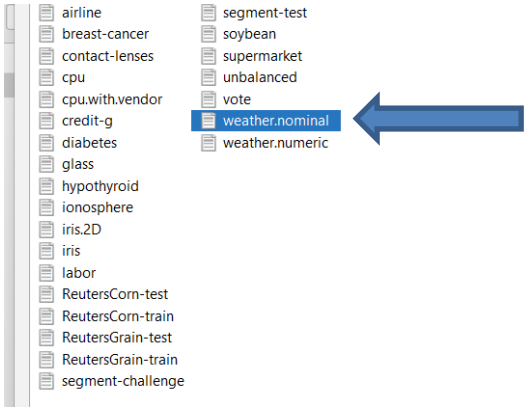
After Applying the Apriori algorithm with the specified parameters when generating 4 Rules:



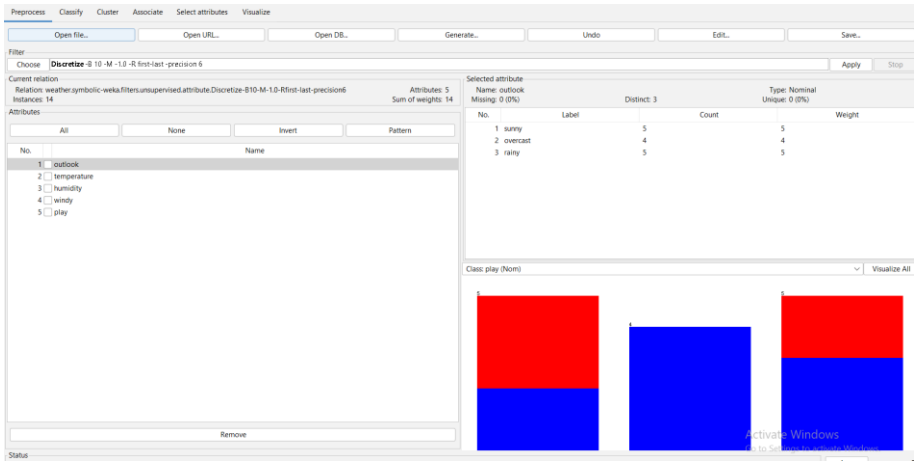
3. Association Rules using FP-Growth Algorithm:

a) Generating Rules with Default Parameters:

• Open Weka and load the Weather.Nominal dataset.



• Apply the FP-Growth algorithm with default parameters.



• Record and analyze all generated association rules.

b) Adjusting Parameters:

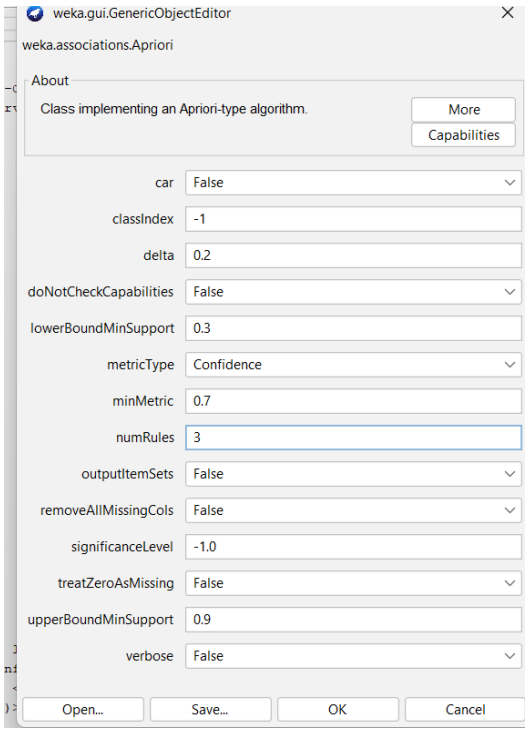
• Set the minimum support range from 30% to 90% with an incremental decrease factor of 20%.

• Set the confidence factor to 70%.

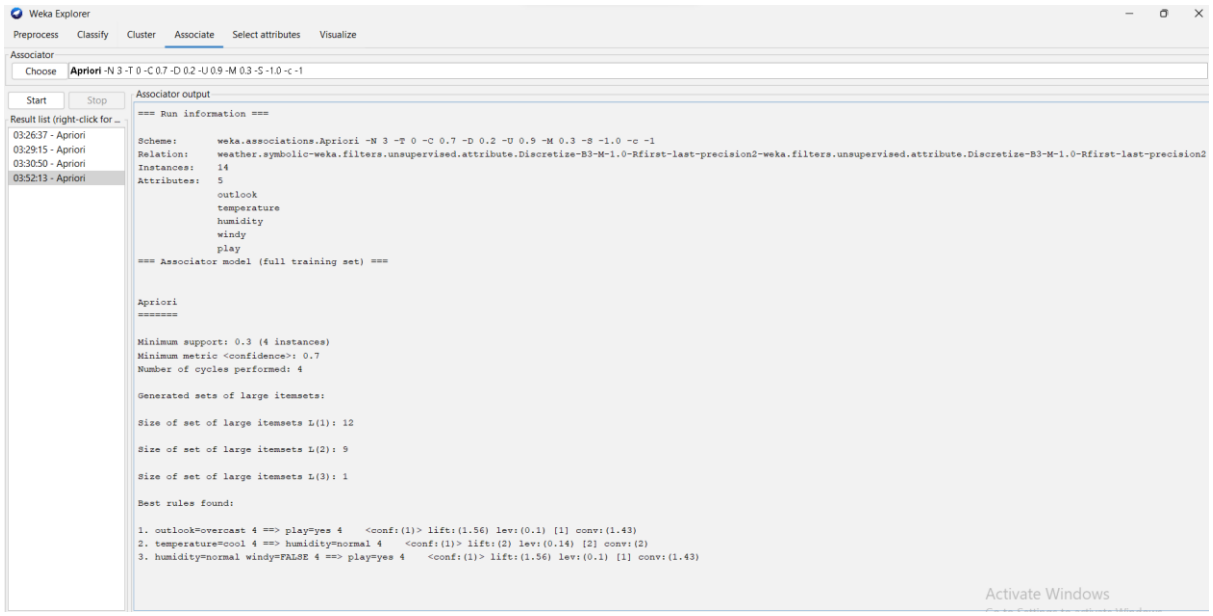
i) Generating 3 Rules:

• Apply the FP-Growth algorithm with the specified parameters.

• Note down the generated rules and analyze them.

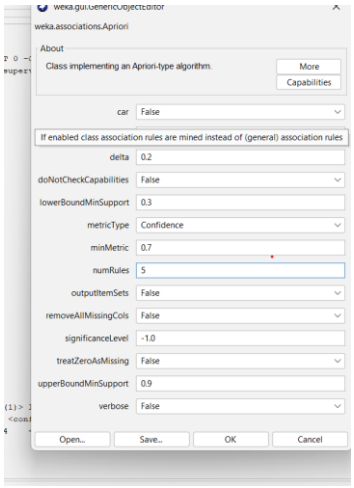


After Generating 3 Rules:



ii) Generating 5 Rules:

• Apply the FP-Growth algorithm with the specified parameters.



• Note down the generated rules and analyze them.

After Generating 3 Rules:

4. Analysis and Comparison:

• Compare the rules generated by the Apriori and FP-Growth algorithms in terms of:

• Number of rules generated.

• Support and confidence values.

• Execution time.

• Overall performance and quality of rules.

Result: