Experiment 4

Association Rule Mining on Employee Dataset

Aim: To create an employee.arff dataset and demonstrate Association rule process on it using apriori algorithm

Tasks:

1. Create employee.arff dataset and load it into Weka.
2. Apply Apriori algorithm with default parameters.
3. Change the parameters and observe the results

Task 1: Create employee.arff dataset and load it.

Create employee.arff with following categorical attributes and load it in to Weka.

|  |  |
| --- | --- |
| **Attribute** | **States** |
| Designation | 1. Manager 2. Developer 3. Tester |
| Beneficiary | 1. Yes 2. No |
| GPF | 1. Yes 2. No |
| Salary | 1. Low 2. Medium 3. High |
| CreditRating | 1. Poor 2. Fair 3. Excellent |
| BankLoan | 1. Yes 2. No |

Task 2: Apply Apriori algorithm with default parameters

Association Rule Mining is a process that finds features which occur together or features that are correlated. Popular applications are Market Basket Analysis and Cross Marketing.

Association rules are mined out after frequent itemsets in a big dataset which can be found using algorithms such as Apriori and FP Growth.

Frequent Itemset mining mines data using support and confidence measures.

Apriori Rule Learner in Weka implements Apriori algorithm. It iteratively reduces the minimum support from its upperBound until (i) it finds the required number of rules or the minimum support reaches lowerBound.

Default values of the some important parameters:

lowerBoundMinSupport = 0.1 (10%)

upperBoundMinSupport = 1.0 (100%)

metricType = Confidence

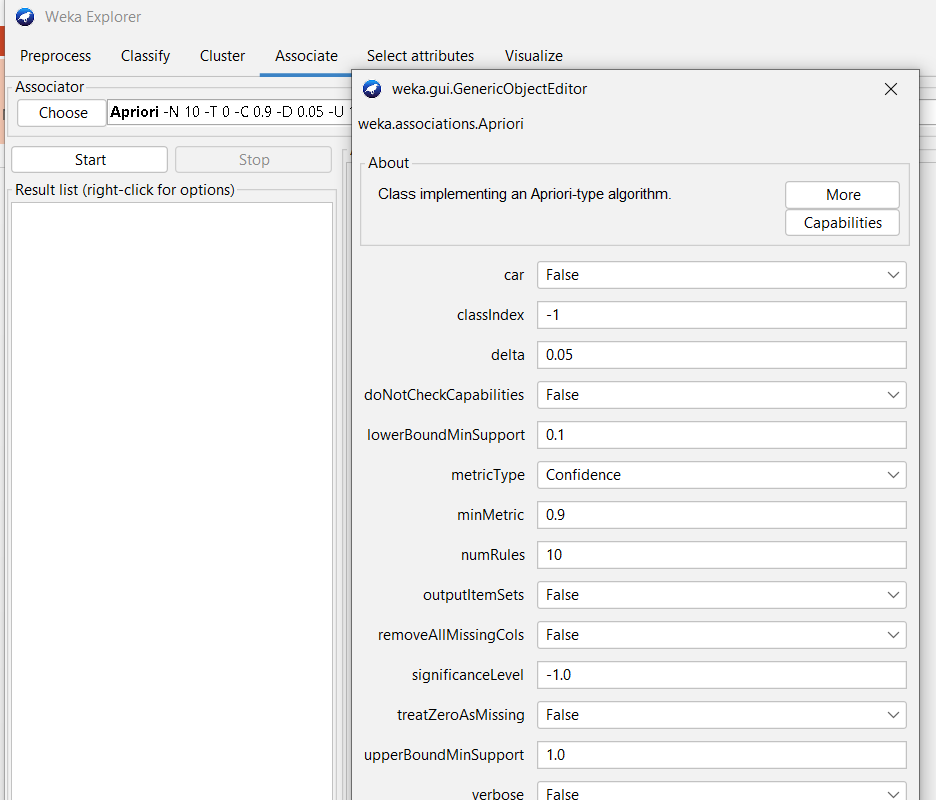
minMetric = 0.9 (90%)

numRules = 10

Steps:

* + - 1. Select Associate 🡪 Choose 🡪 associations 🡪 Apriori.
      2. Observe the default parameter values.
      3. Click on Start.





Observations:

|  |  |
| --- | --- |
| **Default Parameters** | **Observations** |
| lowerBoundMinSupport =  upperBoundMinSupport =  metricType =  minMetric =  numRules = | Minimum support =  Minimum Metric <Confidence> =  Number of cycles performed =  Best rules found with confidence:  1.  2.  3.  4.  5.  6.  7.  8.  9.  10. |

Task 3: Apply Apriori algorithm with required parameters

Observations: Change the default parameter values and perform the experiment

|  |  |
| --- | --- |
| **Parameters** | **Observations** |
| lowerBoundMinSupport =  upperBoundMinSupport =  metricType =  minMetric =  numRules = | Minimum support =  Minimum Metric <Confidence> =  Number of cycles performed =  Best rules found with confidence: |

Conclusion: