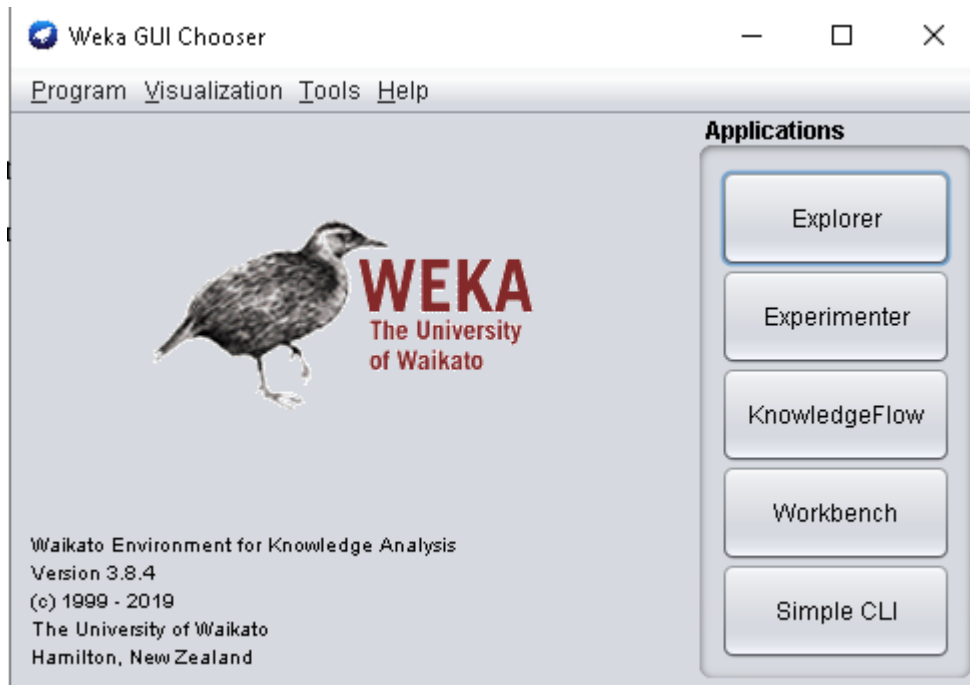


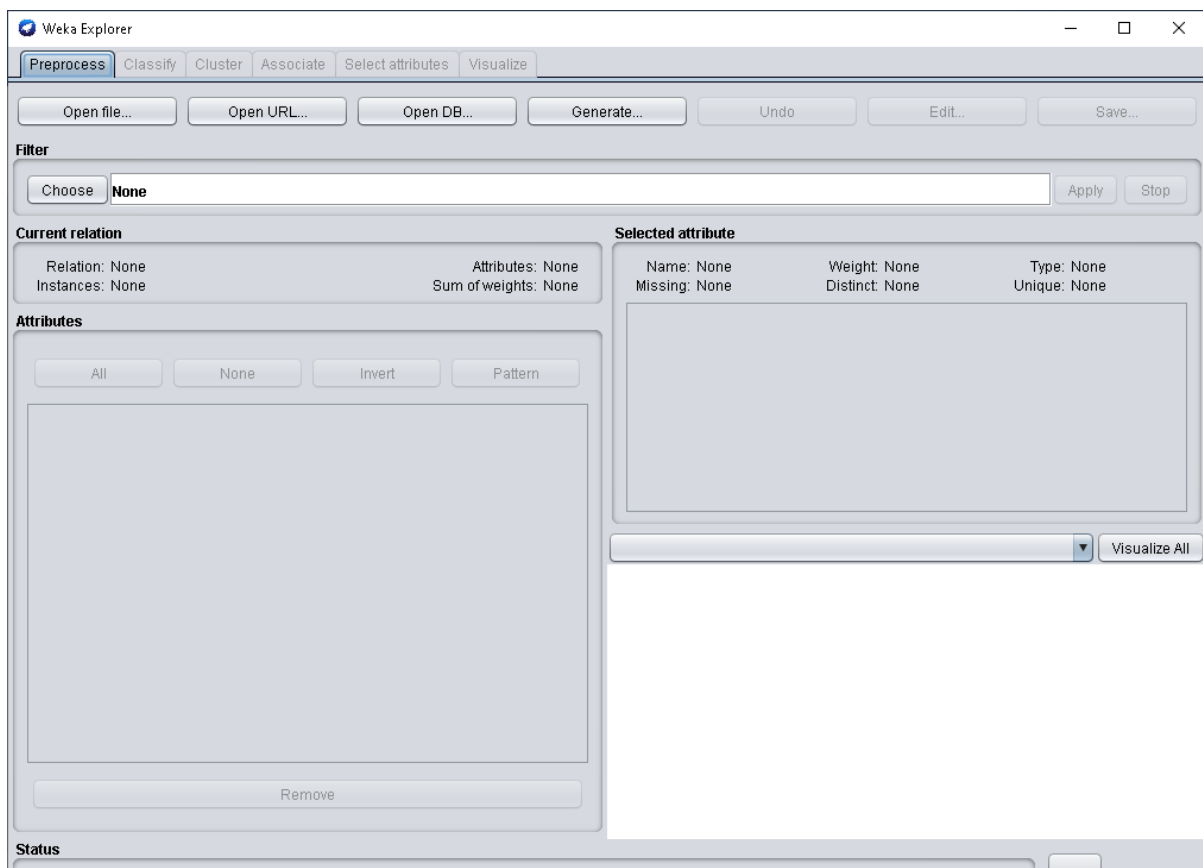
Practical – 6

AIM : To perform hand on experiments with sample data sets on weka.

Step 1: Open tool.



Step 2: Click on Explorer button.



Step 3: To import dataset click on Open file button. I have generated datasets from generator button.

The screenshot shows the Weka Explorer window with the 'Generate...' button highlighted. The 'Current relation' is 'weka.datagenrators.classifiers.classification.Agrawal-S_1_-n_100_-F_1_-P_0.05'. The 'Attributes' list includes salary, commission, age, elevel, car, zipcode, hvalue, hyears, loan, and group. The 'Selected attribute' section shows statistics for 'salary' (Name: salary, Missing: 0 (0%), Distinct: 97, Type: Numeric, Unique: 96 (96%)). A bar chart visualizes the distribution of the 'group' attribute, showing counts for various categories.

Step 4: Use classification algorithm on this dataset here random forest is used and click on start button.

The screenshot shows the Weka Explorer window with the 'Classifier' tab selected. The 'RandomForest' classifier is chosen. The 'Test options' section shows 'Percentage split' set to 66%. The 'Classifier output' section displays the following information:

```

=== Run information ===

Scheme:      weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:    weka.datagenrators.classifiers.classification.Agrawal-S_1_-n_100_-F_1_-P_0.05
Instances:   100
Attributes:  10
             salary
             commission
             age
             elevel
             car
             zipcode
             hvalue
             hyears
             loan
             group

Test mode:   split 66.0% train, remainder test

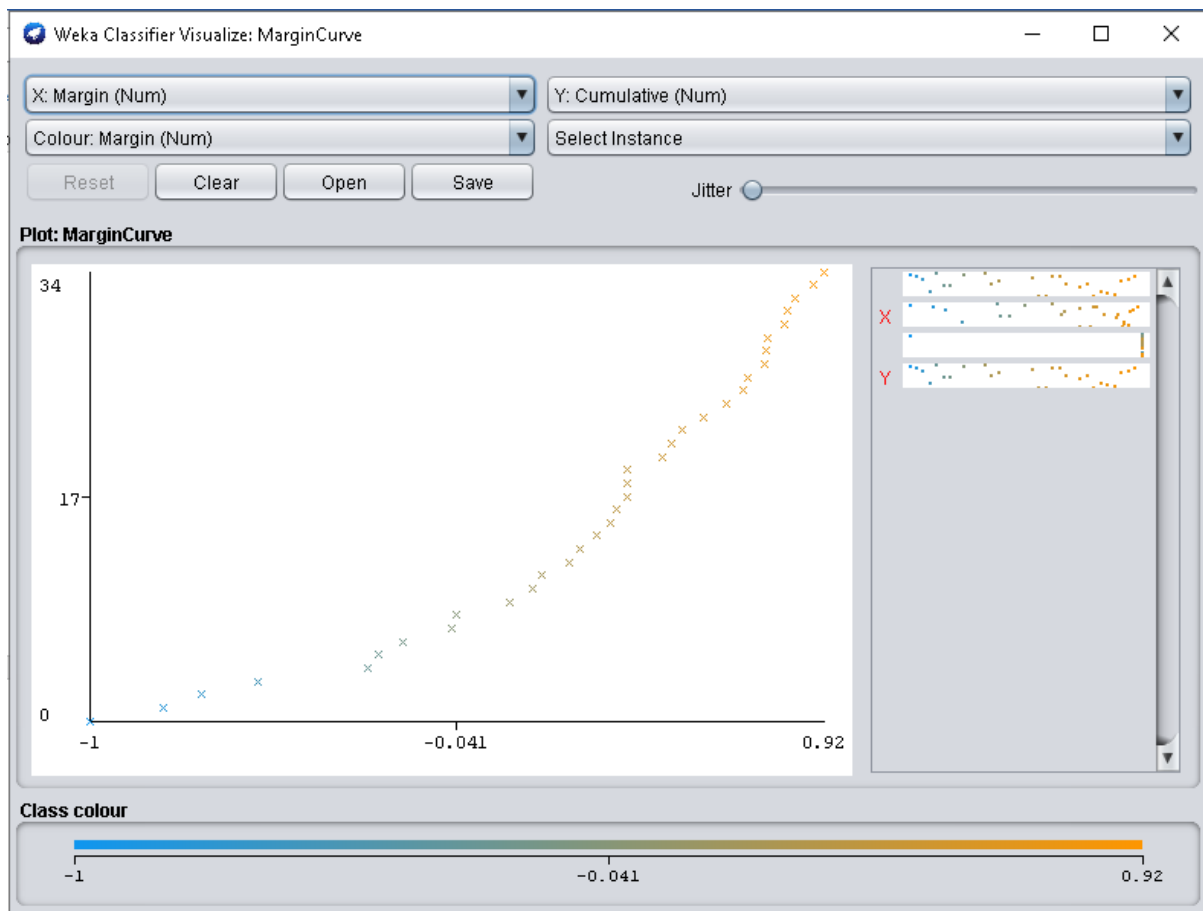
=== Classifier model (full training set) ===

RandomForest

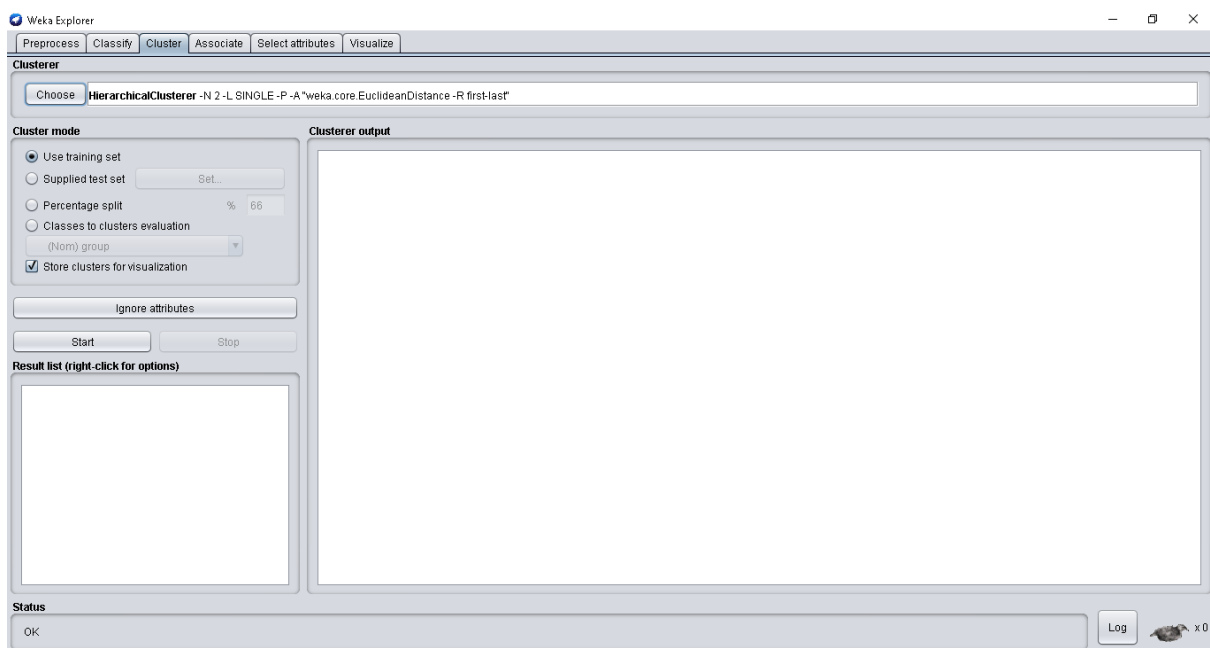
Bagging with 100 iterations and base learner
weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.23 seconds
  
```

Step 5: Right click on trees.randomforest and select visualize margin curve.



Step 6: we can use clustering algorithm in this tool here Hierarchical cluster is used.



[illegible][illegible]

