Experiment No.6

**Aim:** To study and implement K-means algorithm (Weka tool)

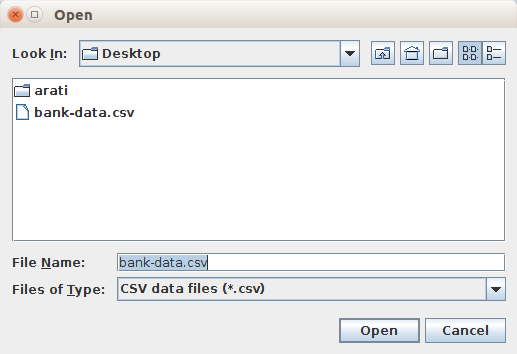
**Theory:**

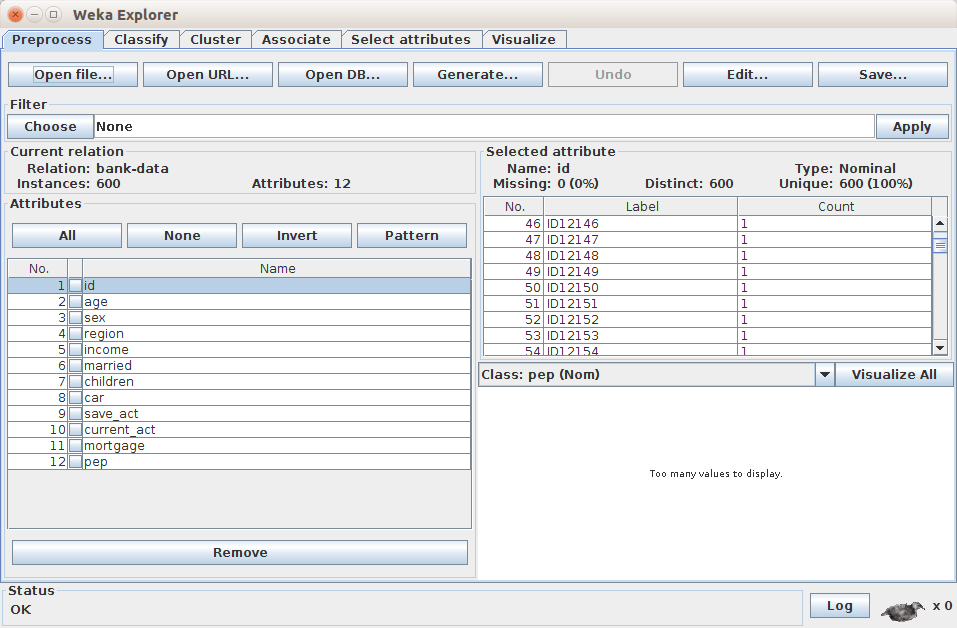
This experiment illustrates the use of simple k-mean clustering with Weka explorer. The sample data set used for this example is based on the bank data available in ARFF format. This document assumes that appropriate preprocessing has been performed. This bank dataset includes 600 instances.

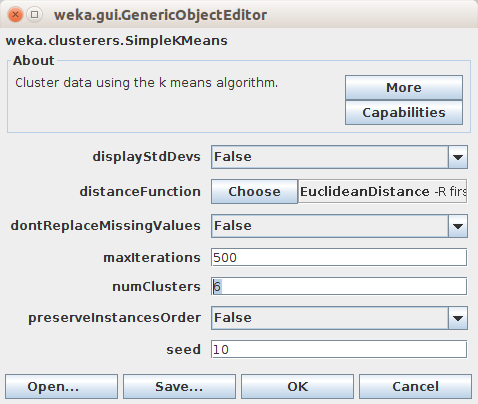
**Steps involved in this Experiment:**

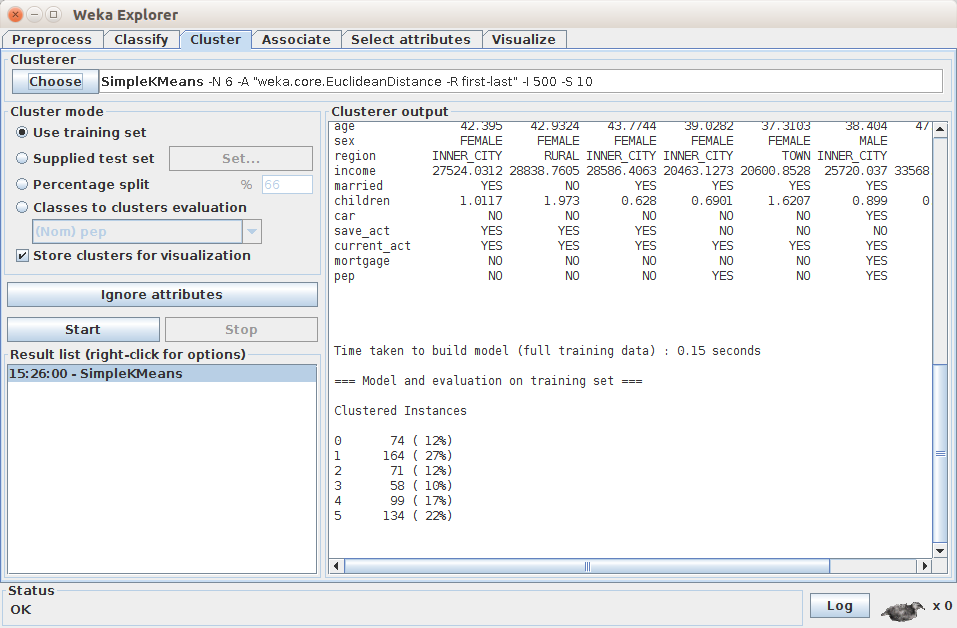
1. Run the Weka explorer and load the data file bank.arff in preprocessing interface.
2. In order to perform clustering select the 'cluster' tab in the explorer and click on the choose button. This step results in a dropdown list of available clustering algorithms.
3. In this case we select 'simple k-means'.
4. Next click in the text button to the right of the choose button to get popup window. In this window we enter six on the number of clusters and we leave the value of the seed on as it is. The seed value is used in generating a random number which is used for making the internal assignments of instances of clusters.
5. One of the option have been specified. We run the clustering algorithm. There we must make sure that they are in the 'cluster mode' panel. The use of training set option is selected and then we click 'start' button/ This process and resulting window are shown in the following screenshots.
6. The result window shows the centroid of each cluster as well as the statistics on the number and the percent of instances assigned to different clusters. Here clusters centroid means vectors for each clusters.
7. Another way of understanding characteristics of each cluster through visualization, we can do this, try right clicking the result set on the result. List panel and selecting the visualize cluster assignments.

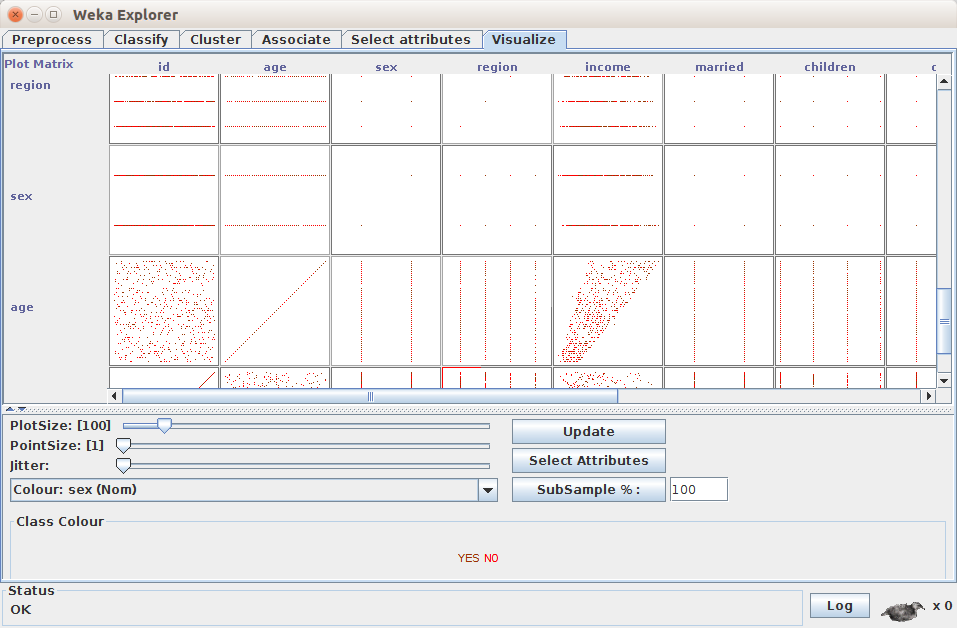
**Screenshots:**

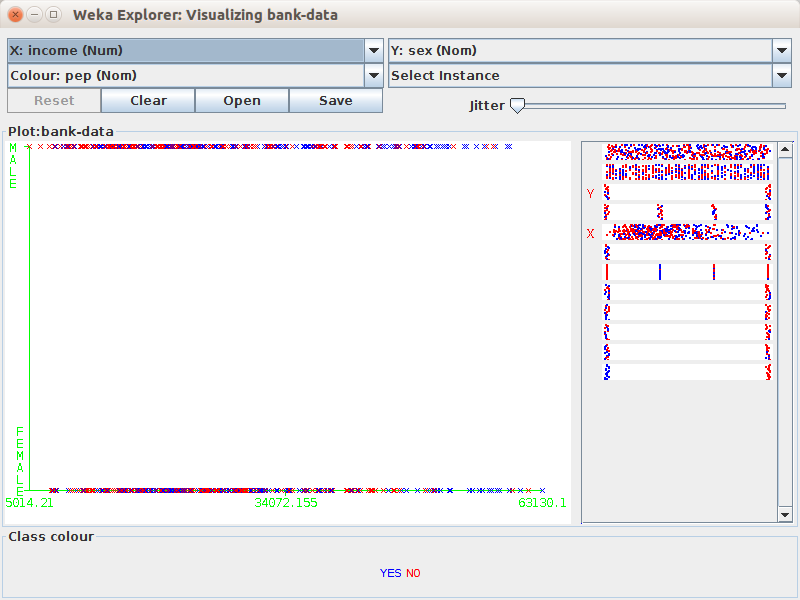












**Conclusion:**

K-means Clustering Algorithm is studied and implemented successfully using WEKA Tool.