**Practical 9**

|  |
| --- |
| **Use WEKA to implement Association Mining land clustering** |
| **K-means, Agglomerative and Divisive Clustering** |

**Title**

Introduction to the Weka machine learning toolkit

**Aim**

To learn to use the Weak machine learning toolkit for Association mining and clustering

**References**

* http://people.sabanciuniv.edu/berrin/cs512/lectures/WEKA/WEKA%20Explorer%20Tutorial-REFERENCE.pdf
* Witten, Ian and Eibe, Frank. *Data Mining: Practical Machine Learning Tools and Techniques.* Springer.
* wekadataset: <https://storm.cis.fordham.edu/~gweiss/data-mining/datasets.html>
* https://www.cs.ubc.ca/labs/beta/Projects/autoweka/datasets/

**Task:**

**1) Association mining on grocery .arff file (attach screen shots of output)**

**2) Cluster mining -k-means clustering scheme using customers.arff (attach screen shots of output)**

**Post labs**

How do you load Weka?

1. What options are available on main panel?
2. What is the purpose of the the following in Weka:
   1. The Explorer
   2. The Knowledge Flow interface
   3. The Experimenter
   4. The command-line interface
3. Describe the arff file format.
4. Press the Explorer button on the main panel and load the weather dataset and answer the following questions
   1. How many instances are there in the dataset?
   2. State the names of the attributes along with their types and values.
   3. What is the class attribute?
   4. In the histogram on the bottom-right, which attributes are plotted on the X,Y-axes? How do you change the attributes plotted on the X,Y-axes?
   5. How will you determine how many instances of each class are present in the data
   6. What happens with the Visualize All button is pressed?
   7. How will you view the instances in the dataset? How will you save the changes?
5. What is the purpose of the following in the Explorer Panel?
   1. The Preprocess panel
      1. What are the main sections of the Preprocess panel?
      2. What are the primary sources of data in Weka?
   2. The Classify panel
   3. The Cluster panel
   4. The Associate panel
   5. The Select Attributes panel
   6. The Visualize panel.