CST8390 - Lab 1
Data exploration and integration with Weka - Iris dataset

Due Date: Sept 13, 17, 18 in corresponding lab sessions **Introduction**

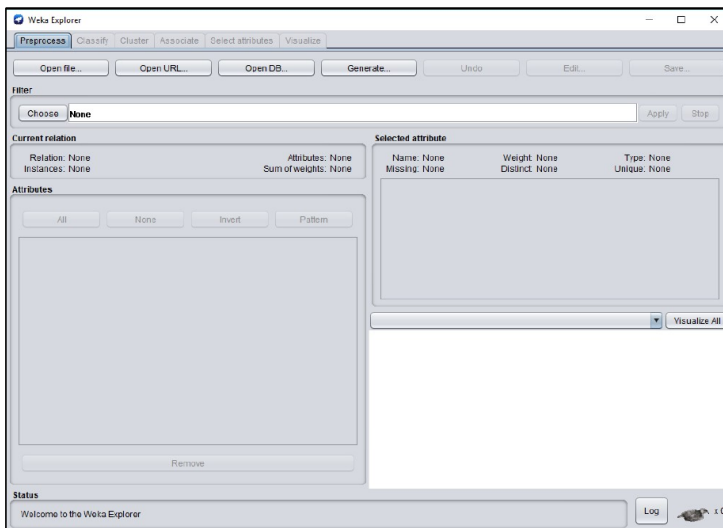The goal of this lab is to install and familiarize with Weka.

Steps:

1. Download and install Weka. You can find it here:
    http://www.cs.waikato.ac.nz/ml/weka/downloading.html

2. Open Weka and have a look at the interface. It is an open-source project written in Java from the University of Waikato.

3. Click on the Explorer button on the right side:

4. Check different tabs to familiarize with the tool.

5. Weka comes with a number of small datasets. Those files are located at C:\Program Files\Weka3-8 (If it is installed at this location. Or else, search for Weka-3-8 to fins the installation location). In this folder, there is a subfolder named 'data'. Open that folder to see all files that comes with Weka.

6. For easy access, copy the folder 'data' and paste it in your 'Documents' folder.

7. In this lab, we will work with the dataset Iris. To open Iris dataset, click on 'Open file' in the 'Preprocess tab'. From your 'data' folder, select iris.arff and hit open.

8. To know more about the iris dataset, open iris.arff in notepad++ or in a similar tool and read the comments.

9. Click on visualize tab to see various 2D visualizations of the dataset.

    a. Click on some graphs to see more details about it.
    b. In any of the graph, click one'x' to see details about that data record.

10. Fill this table:

Flower Type	Count
Iris Setosa	50
Iris Versicolour	50
Iris Virginica	50

11. Fill this table:

Attribute	Minimum	Maximum	Mean	StdDev
sepal length	4.3	7.9	5.84	0.83
sepal width	2.0	4.4	3.05	0.43
petal length	1.0	6.9	3.76	1.76
petal width:	0.1	2.5	1.20	0.76

In order to get the credit for this lab:
1. Show the loaded file in Weka (5 marks)
2. Fill in the tables for questions 10 & 11 (5 marks)