Software Engineering Tools Lab

Assignment No-1

(Module 1- Introduction to OSS) **Due date-04/02/2022**

1. **Weka** is a GUI workbench that empowers data wranglers to assemble machine learning pipelines, train models, and run predictions without having to write code.
   Using Weka tool perform below tasks such as data preprocessing, data classification (use any appropriate ML algorithm) and data visualization efficiently on given dataset.
   Use the Iris dataset given-
   https://drive.google.com/file/d/1A3Fxsfzm6BSfhFZGDrjI47RTe45bSgYP/view

   Note-provide screen shots for every task
   Create a report which will illustrate the details of tasks performed (for e.g to perform preprocessing of data provide details of navigation and selection of appropriate parameters)

2. **Orange** is an easy to use data visualization tool with a large toolkit. In spite of being a GUI-based beginner-friendly tool, you mustn't mistake it for a light-weight one. It can do statistical distributions and box plots as well as decision trees, hierarchical clustering and linear projections.
   a. Install orange
   b. Show data distribution
   c. Show linear projection
   d. Show FreeViz

   Use dataset
   https://drive.google.com/file/d/1m6sKI1Dap0XK6Bw1edUd5PohwpPwXnd9/view

   Create a report for this task and upload screenshots for the same.

3. Differentiate in between free software, Open source software and proprietary software with respect to its properties.

4. Using **Anaconda Python** create Histogram, Scatter plot and Bar plot for the dataset given below.
   Dataset- https://drive.google.com/file/d/1i11BZFe8Xj9kNq7eeE9KOa_Iz1KhEdXJ/view

   a. Scatter plot- Scatter plot of Price Vs Age
   b. Histogram- for Kilometer and CC
   c. Bar plot- Bar plot for different fuel types

5. Enlist some examples along with its purpose and properties (at least 10) of FOSS and proprietary software with respect to database.