Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam – 603 110
(An Autonomous Institution, Affiliated to Anna University, Chennai)

UCS2612 Machine Learning Laboratory

Academic Year: 2023-2024 Even   **Batch: 2021-2025**
Faculty In-charges: Y.V. Lokeswari & Nilu R Salim   **VI Semester A & B**

Lab Test 1-Q5

Images of Kecimen and Besni raisin varieties grown in Turkey were obtained with CVS. A total of 900 raisin grains were used, including 450 pieces from both varieties. These images were subjected to various stages of pre-processing and 7 morphological features were extracted.

Download the dataset from the following link:

https://archive.ics.uci.edu/dataset/850/raisin

Develop a Python program to predict the type of raisin using all the classification models (LR, PLA, MLP, KNN, SVM, Naïve Bayes) you have learned. Interpret the model that works better for this dataset. Visualize the features from the dataset and interpret the results obtained by the model using the Matplotlib library. **[CO1, K3]**

Use the following steps to do implementation:
1. Loading the dataset.
2. Pre-Processing the data (Encoding, Standardization, Normalization, Handling missing values, Noisy data suitable for the dataset)
3. Exploratory Data Analysis.
4. Feature Engineering techniques.
5. Split the data into training, testing and validation sets.
6. Train the model.
7. Test the model.
8. Measure the performance of the trained model.
9. Represent the training and testing results using ROC curves. Does the model overfit. Comment on your obtained results.
10. Ignore the class label and perform clustering task. Measure the performance of the model.

▪▪

Upload the code in GitHub and include the GitHub main branch link in the google form shared.