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

Predict Recurrence of Breast Cancer

This breast cancer domain was obtained from the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia. This is one of three domains provided by the Oncology Institute that has repeatedly appeared in the machine learning literature. (See also lymphography and primary-tumor.) This data set includes 201 instances of one class and 85 instances of another class. The instances are described by 9 attributes, some of which are linear and some are nominal.

Target label is Class.

https://archive.ics.uci.edu/dataset/14/breast+cancer

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

Use the following steps to do implementation:

1. Loading the dataset.
2. Pre-Processing the image data
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 assignment PDF.