ASSIGNMENT #1

Construct a machine learning based model for classification using Python for the following UCI datasets:

UCI datasets (can be loaded from the package itself):

a. Iris plants dataset: https://archive.ics.uci.edu/ml/datasets/Iris/
b. Diabetes dataset: https://www4.stat.ncsu.edu/~boos/var.select/diabetes.html
c. Wisconsin Breast Cancer Dataset: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

1. Employ Naive Bayes (Gaussian, Multinomial & Bernoulli) classifier and show classification results (Accuracy, Precision, Recall, F-score, confusion matrix) with and without parameter tuning.
2. Use Decision Tree classifier for all the three datasets and show classification results (Accuracy, Precision, Recall, F-score, confusion matrix) with and without parameter tuning. Generate the decision tree images for all cases highlighting information like Gini and Entropy.

Save the assignment in a single pdf file with the naming convention "Class Roll No_Full Name.pdf" and submit the report to us through email (using "Reply All" button) by 25th August 2021 Tuesday EOD

- pksingh.it@jadavpuruniversity.in
- neelotpalc.cse.rs@jadavpuruniversity.in