**Instructions Lab 5\_1**

Till now we have seen three different types of classifiers namely K Nearest Neighbor, Univariate Decision Trees and Bayesian (specifically Naïve Bayes Gaussian) classifiers. We will learn a few more algorithms in this course. It is a good idea to periodically check which algorithm works well for a given dataset. In this lab you will compare the performances of the three algorithms studied so far on the breast cancer dataset that you have used earlier.

1. Download the data from

<https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>

It is data regarding breast cancer diagnostics.

2. Input the data using the Pandas library after dropping the first column.

3. Perform a random 70 – 30 split of the sample for training and testing sets.

4. Build three different classifiers using the three algorithms and the training data.

5. Now test each classifier using the test data and report the corresponding confusion matrices.

6. For each training sample

For each classifier

If the sample is misclassified then

Note the index of the misclassified sample, the true class and the predicted class

7. Prepare a report of all those test samples that were misclassified by all the classifiers

*Show your work, even if it is partial, since each lab is graded based on what you achieve during the lab.*