Lab Assignment – 6

1) Download the following dataset using the link given below:
   a) Iris
   b) Wine
   c) Breast cancer

Reduce the following dataset into k (specified below) dimensions using LDA and on the reduced dataset perform classification using SVM and Random Forest for different values of k. Generate the confusion matrix for each classification and calculate the classification accuracy, precision, recall, and F1-score using confusion matrix for each value of k.

   a) Iris for k = 2,4
   b) Wine for k = 4, 6, 10, 13
   c) Breast cancer for k = 5, 10, 15, 20, 32

Plot the classification region of Iris dataset using k=2 for both SVM and Random Forest.

Also, perform the classification task on the reduced dataset using PCA for each dataset using same k. Compare the results of PCA and LDA, and submit a report containing the following:

   a) The difference between PCA and LDA.
   b) Comparison of classification results using PCA and LDA for each dataset.

Note:

1) Use of inbuilt LDA and PCA function is not allowed.
2) For SVM and Random Forest, you may use library functions.
3) You may use PCA function that you coded in assignment 5.
4) For classification, divide the dataset into 70% for training and 30% for testing.