Day10 Questions

PCA & LDA

1) a) Apply PCA for the breast_cancer dataset in sklearn to reduce the number of columns to 2 and find out the accuracy score and confusion matrix. Split the data using train_test_split function for training and testing. Use Support Vector Classifier.

   b) do the same for the wine dataset

2) a) Compute the RMSE values of the ML model using SVR after applying PCA to reduce the number of columns to 3 in the diabetes dataset. Observe the effect of varying the number of principal components in the model development.
   b) do the same for boston dataset.

3) Develop an ML model for the breast_cancer dataset in sklearn using LDA.

   `(from sklearn.discriminant_analysis import LinearDiscriminantAnalysis)`

   Print the accuracy and confusion_matrix.