

2)

a) Sometimes the accuracy of Perceptron is better than NCC but always smaller than IDA algorithm's accuracy

b) According to given covariance matrix data is uncorrelated so LDA is not optimal for this dataset

c) If we left the data uncorrelated but make the class variances same we can get the same result

3) For each folds the function splits the data into a train and test data from the original data using a matrix contains N element. Every time it takes one row of the matrix by turns, use it as an index for forming the train data and uses the rest for test data. It does the prediction with train data and calculate accuracy for both train and test data. After computation of accuracies for all folds return them. We should look at test accuracies since we train the model based on training data test data gives the right accuracy for comparing the performance

4) For USPS dataset I would prefer LDA or Perceptron but for BCI data I would definitely prefer only LDA algorithm.