4. a)   LDA gave best accuracy of around 97.3 %.

b)   LDA gave best results in our analysis because I think the feature were less and the sample data was also small. LDA uses covariance to consider the interdependence of features and it is proven to work well.

LDA's performance highly depends on the 2 fold strategy that we used and also on the features and number of samples. It is not the case that it is the best algorithm as it highly depends on those parameters. Below are the individual reasons of the algorithms which could have affected their overall accuracy.

Naïve Baysian gave an accuracy of 96%, because it predicts the output based on the conditional probability which couldn't perform better than the LDA. Also it can't perform best when we use half the data as training set.

Linear regression gave an accuracy of 96.67% when used with linear and polynomial of 2 degree and gave an accuracy of 92% when used with polynomial of degree 3. This algorithm establishes a linear relationship among the features based on the statistical analysis which performs better with higher number of features.

kNN gave an accuracy of 94% because it finds the nearest neighbours on the plot for the prediction and can't predict accurately when there is an overlapping of value of features for different classes.

QDA had an accuracy of 96% and I think QDA perform better when data is high dimensional.