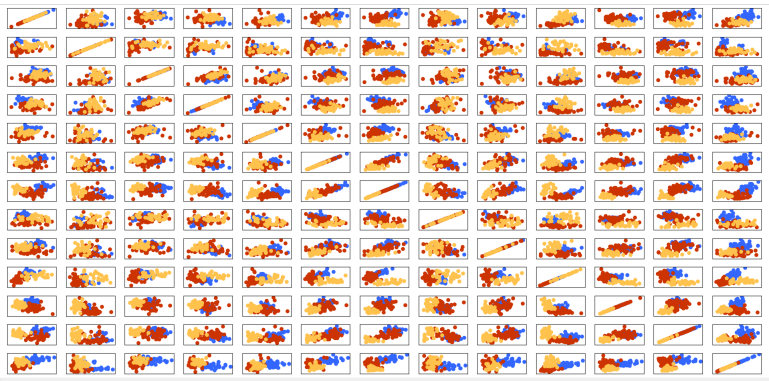
SPS Coursework 2 Report

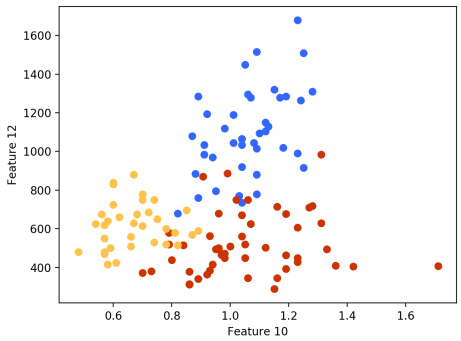
Michael Mafeni and Ainesh Sevellaraja

1 Feature Selection

We first plotted a 13x13 scatter plot for all pairwise combinations of features of the training data set, as seen below.



Based on the plot above, we decided to select features 10 and 12 as the data separate into their classes fairly well. This can be seen in the scatter plot of feature 12 against feature 10 below. We then reduced the train set and test set to only contain features 10 and 12.



2 K-Nearest Neighbours

3 Alternative Classifier

4 Use Three Features

5 Principal Component Analysis (PCA)