Exercise 1

The aim of this exercise is to create 2 classifiers, a linear and a quadratic, with the glass data. Follow the steps below and use the lab notes.

i. Download* the glass.data and load it to Jupyter

ii. Transform data into [0 - 1]
(Optional : Remove any outliers)

iii. Split the data into a training and a validation set

iv. In 3 scatter plots visualize:
- All the instances in the dataset
- The training set
- The test set

v. Create a linear classifier (LDA) and fill table 1.

vi. Create a quadratic classifier and fill table 2.

vii. Repeat steps v & vi, using 10-fold cross validation this time. Compare the results.

viii. What are your conclusions? Which classifier is better?

10-Fold Accuracy

Table 1 – Linear Classifier	
Precision Score (weighted_ avg)	0.74
Recall Score (weighted_ avg)	0.71
F1 Score (weighted_ avg)	0.69

0.75920

Table 2 – Quadratic Classifier	
Precision Score (weighted_ avg)	0.52
Recall Score (weighted_ avg)	0.58
F1 Score (weighted_ avg)	0.52

0.66562

*Information about the dataset:
https://archive.ics.uci.edu/ml/datasets/glass+identification

Deliverables: This report along with the .ipynb file you created to get the results. Compress both of them and name the compressed file as: RN_xxxx, where RN is your registration number.

Good Luck !!!

The quadratic classifier is more accurate and produces better results in all instances.

The values on the quadratic are lower thereofore produce more accurate results than the linear. Same as in the 10 Fold Cross Validation.