Applied Machine Learning, Module 1: A simple classification task

Import required modules and load data file

The file contains the mass, height, and width of a selection of oranges, lemons and apples. The heights were measured along the core of the fruit. The widths were the widest width perpendicular to the height.

Examining the data

Create train-test split

Create classifier object

In [8]:
```
from sklearn.neighbors import KNeighborsClassifier

knn = KNeighborsClassifier(n_neighbors = 5)
```

Train the classifier (fit the estimator) using the training data

Estimate the accuracy of the classifier on future data, using the test data

Use the trained k-NN classifier model to classify new, previously unseen objects

Plot the decision boundaries of the k-NN classifier

How sensitive is k-NN classification accuracy to the choice of the 'k' parameter?