KNN

Vary k: In three different executions, select three odd heroes on k, in the interval

3 to 9

Explain what difference there will be with different k

Try two runs and explain the difference in the two cases:

with non-normalized data, ie original data.

with normalized data in the interval [0,1]

In these runs change the number of training and test cases, so that: train = 75%, test = 25% train = ⅔, test = ⅓ train = 50%, test = 50%

In total, this will be 3 \* 2 \* 3 = 18 runs (Why? Because it was partly three different k, partly the data would be in original form or normalized and partly training / test data amounts would be varied in three ways.)

Present the three best executions, with respect to the performance (eg value of Accuracy) of the classification algorithm. Always show confusion matrix and accuracy for each reported run.

DECISION TREE

Try two runs and explain the difference in the two cases: with non-normalized data, ie original data. with normalized data in the interval [0,1] Tips for normalization:

In these runs change the number of training and test cases. Put random\_state = None in the DecisionTreeClassifier constructor (meaning that the test and training observations are really random every time you run) and run each point below 100 times (from start to finish) and calculate the mean value of the performance of these 100 runs: