Which of the following statements about k-Nearest Neighbour (k-NN) are true in a classification setting, and for all k? Select all that apply and give explanation for each one of them (Why correct and incorrect)

1. The training error of a 1-NN will always be better than that of 5-NN.
2. The test error of a 1-NN will always be better than that of a 5-NN.
3. The decision boundary of the k-NN classifier is linear.
4. The time needed to classify a test example with the k-NN classifier grows with the size of the training set. [3 Marks – 0.75 marks for each option]

**Solution**

**1, 4 are true and 2, 3 are false**

1. This is true. If you use the training data set as the test set, then with one nearest neighbor, if given a point x, the nearest neighbor will be the exact same point and thus the error will be 0. For 5-NN, 0 is a lower bound.
2. False. This will have test error 100 for 1-nn and 0 for 5-nn
3. No. Consider two classes, one is in the shape of a moon and the other surrounds the moon. Then the decision boundary will have approximately the shape of a moon.
4. This is true. At test, KNN needs to make a full pass through the entire data set and sort points by distance. The time needed thus grows with the size of the data.