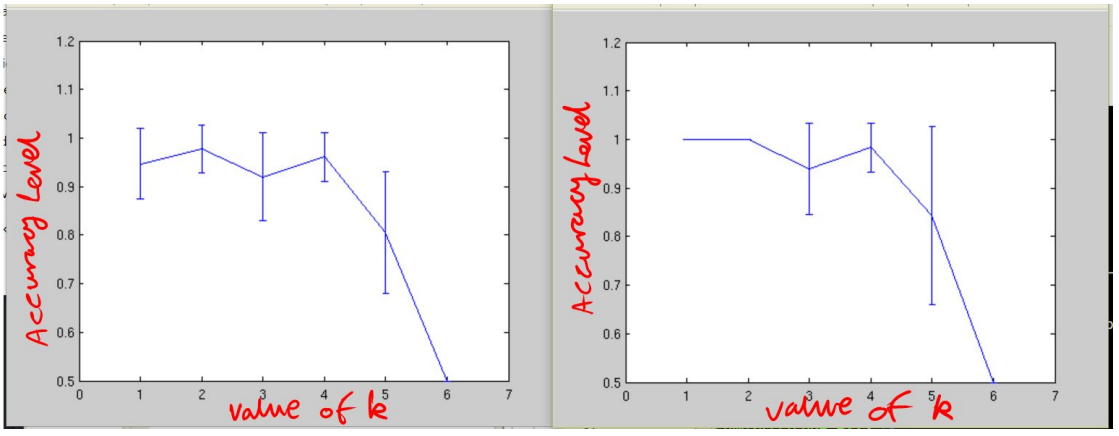
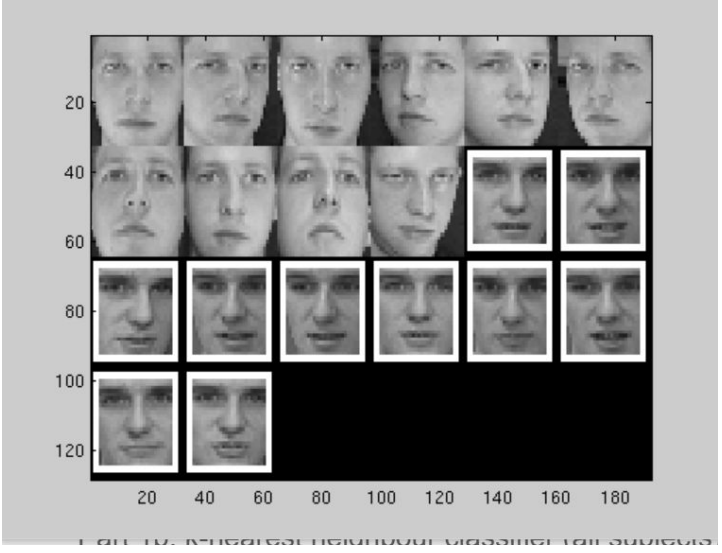
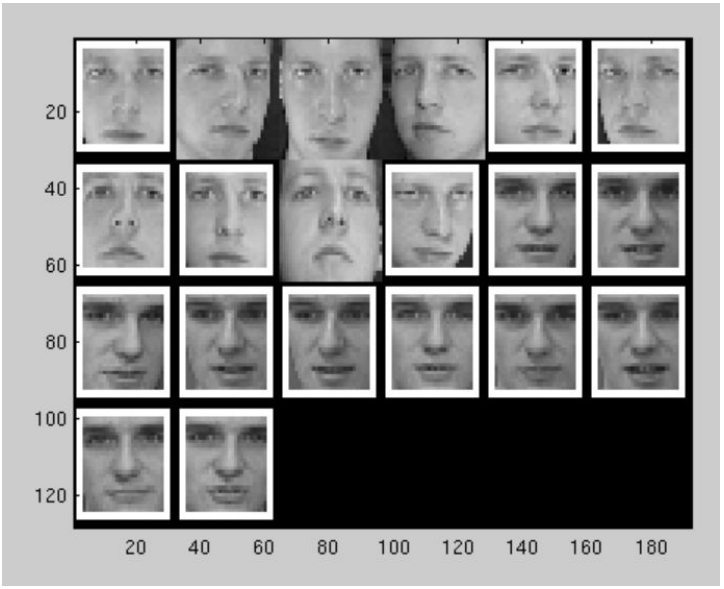
Report for ex2 - 10136960

Part 1a: k-nearest neighbour classifier (Subjects 1 and 30)

1. 100% training accuracy when k = 1. Since only 1 nearest neighbour, the nearest neighbour of a testing point will therefore be itself, and by definition always in the same class.
2. Very similar behaviour over the training and testing sets, but testing graph never reached 100%.
   - 

3. With binary classification, it is a bad idea to have k as an even number. If this happens, the classifier is unsure which class to assign a test point to, if there are k/2 neighbours each from class X and Y.
   Otherwise, it is acceptable to have k as an even number.
4. Subject 30 harder to classify - it has more incorrect classifications on the graph.
   Below shows training data results then testing data results.
   - 

- 

Part 1b: k-nearest neighbour classifier (all subjects)

1. No, because not every single subject can be ensured to be sampled.
2. Subject 30 was the most difficult to recognise - it had more incorrect classifications (denoted by white boxes)

Part 2a: Linear Least Squares classifier (Subjects 1 and 30)

1. Minimum: 57%, Maximum 100%
    - It will always be above 50% due to the nature of the data conveniently sorted for the linear least squares line
2. Test subject 30 is harder to classify because more classifications are labelled wrong (white boxes)
    -