CS 4442 Artificial Intelligence II - Assignment 1

Ian Wood - 250690592

February 12th 2017

# Question 3:

b) Tested using the full data set, not the tiny one.

|  |  |
| --- | --- |
| k Value | Error |
| 1 | 0.0835 |
| 3 | 0.0790 |
| 5 | 0.0850 |
| 7 | 0.0885 |

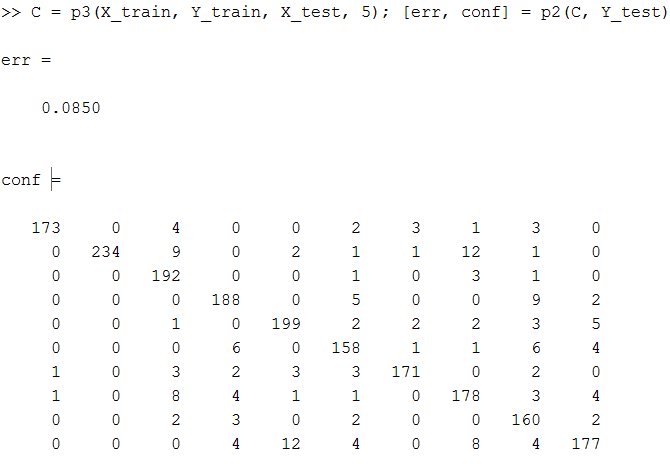
The error is consistently low, no matter the value of K. Therefore, in this test, there don’t seem to be an outstandingly large effect on the error when K is changed. This could be a consequence of the fact that error is consistently low and therefore increasing the k value can’t make a significant improvement to the accuracy of the classifier.

c) The six most common confusions are 2 classified as 8, 0 misclassified as 5, 2 misclassified as 3, 4 misclassified as 9, 8 misclassified as 3 and 8 misclassified as 2 at a rate of 12, 12, 9, 9, 8 and 8, respectively. About half of the remaining cells are 0, with the rest being below 6.

The biggest single confusions are 2 classified as 8 and 0 misclassified as 5 at 12 misclassifications each.

The biggest sum of confusions is 2 classified as 8 and 8 misclassified as 2 at a total of 20 misclassifications.

The 2 and 8 misclassifications are not symmetrical, one is 12 and the other is 8. However, it is possible that the reason that number of misclassifications are identical simply because of the distribution of data in the training set. It’s hard to say with certainty whether the difference is just noise, but there is still a reasonable difference.



# Question 5:

# Question 6:

# Question 7:

# Question 8:

# Question 10: