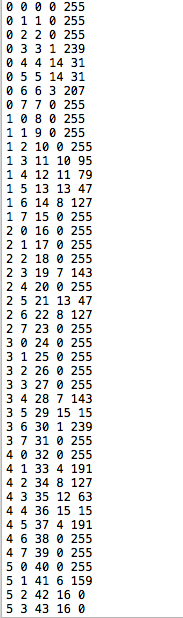
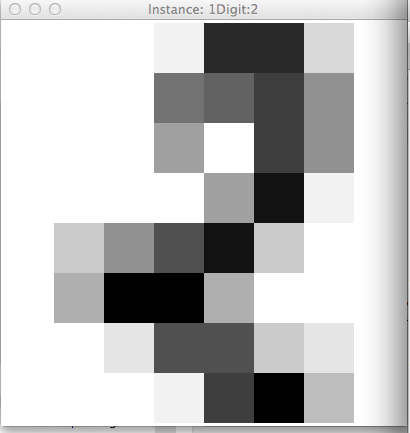
Assignment 3

1. √
2. 

Relationship:

Column1: row index

Column2: column index

Column3: sequence

Column4: 1-16

Column5: dark degree.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | avg\_acc | Digit 8 precision | Digit 8 recall | F-score |
| 1 | 97.941% | 0.995 | 0.974 | 0.979 |
| 2 | 97.3845% | 0.995 | 0.982 | 0.974 |
| 3 | 97.8297% | 0.997 | 0.968 | 0.978 |
| 4 | 97.7741% | 0.997 | 0.971 | 0.978 |
| 5 | 97.941% | 0.997 | 0.961 | 0.979 |
| 6 | 97.662% | 0.995 | 0.968 | 0.977 |
| 7 | 97.5515% | 0.995 | 0.966 | 0.976 |
| 8 | 97.5515% | 0.975 | 0.914 | 0.976 |
| 9 | 97.8854% | 0.97 | 0.925 | 0.979 |
| 10 | 97.5515% | 0.964 | 0.92 | 0.976 |
| 11 | 97.7741% | 0.97 | 0.925 | 0.978 |

Original Author’s

k = 1 : 98.00

k = 2 : 97.38

k = 3 : 97.83

k = 4 : 97.61

k = 5 : 97.89

k = 6 : 97.77

k = 7 : 97.66

k = 8 : 97.66

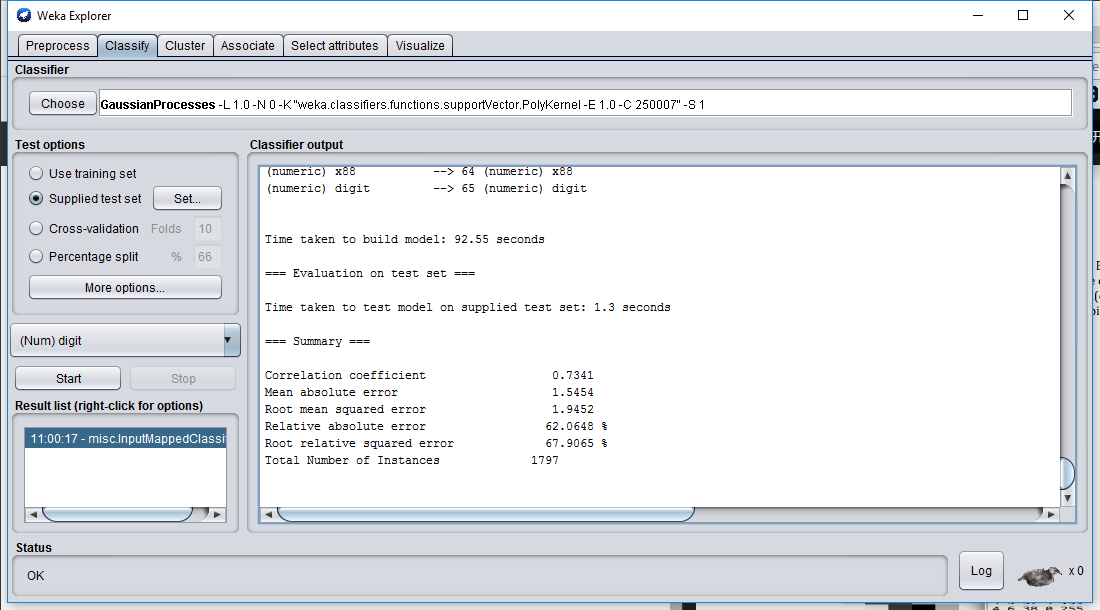
k = 9 : 97.72

k = 10 : 97.55

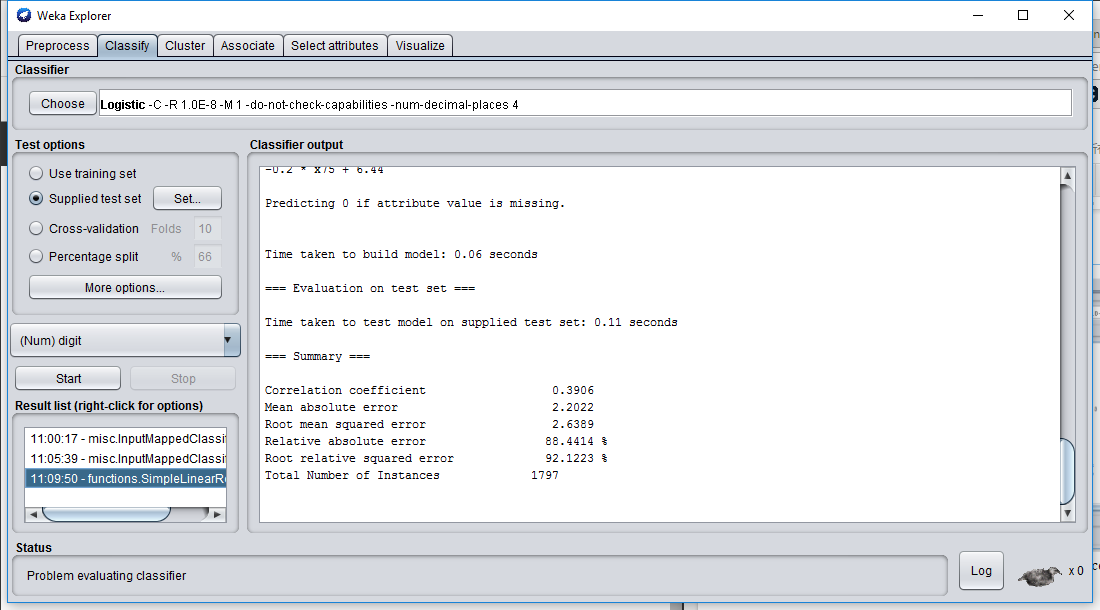
k = 11 : 97.89

there are only minor differences

d)



Gaussian Process

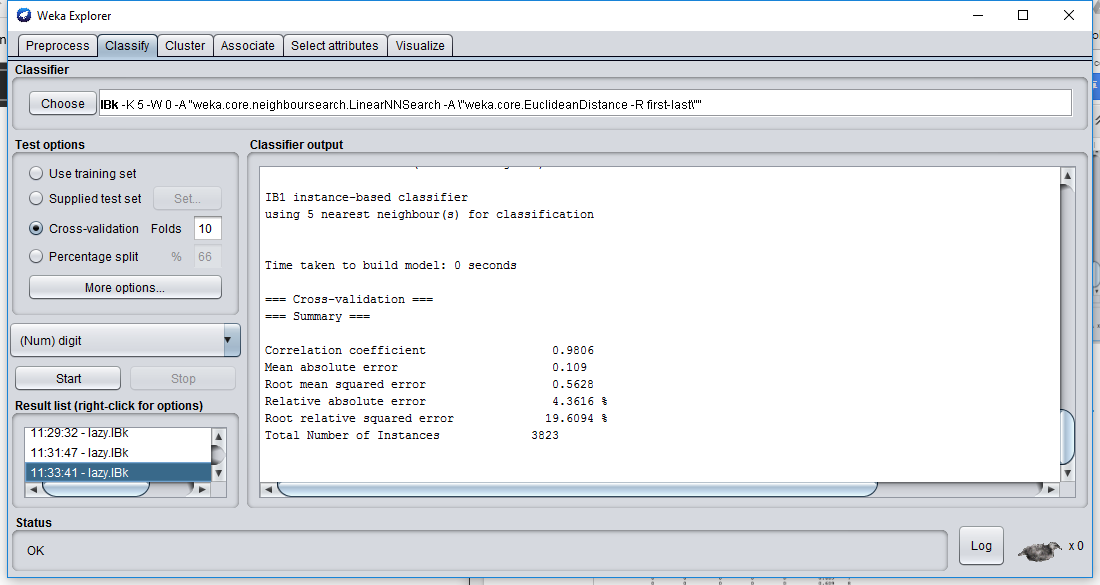


Simple Linear

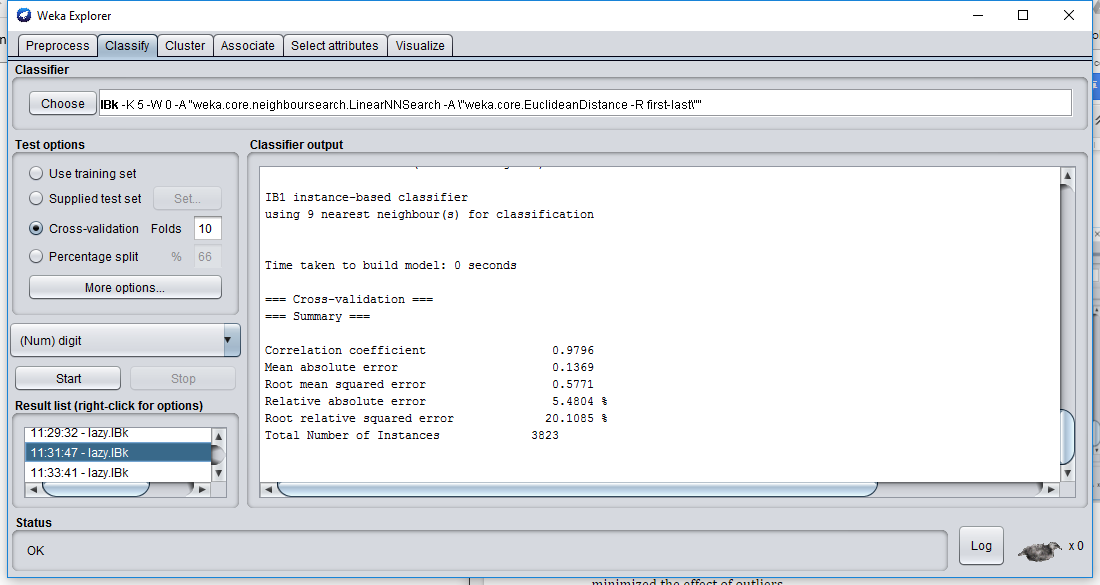
the instance based classifiers out-perform non-instance based classifiers.

e)

f)



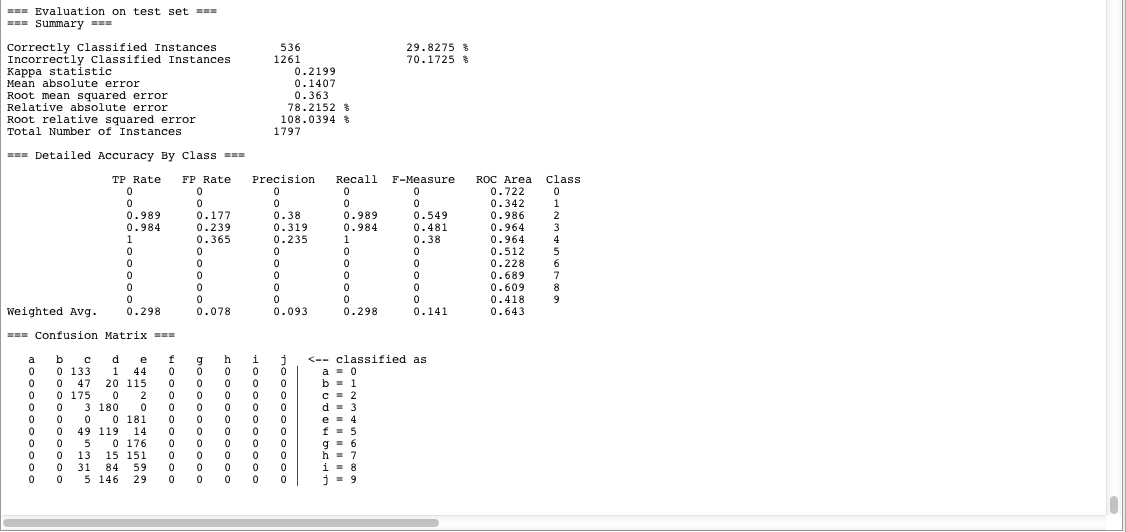
KNN=5, get better accuracy



KNN=9, get better accuracy

It might because 10-fold cross validation train and test for 10 times which minimized the effect of outliers.

g) I expect the accuracy will decrease to only 30% because the training has 3 numbers while the training data have 9. I choose the multilayer perceptron to test.



h)

Use preprocessing programs made available by NIST to extract normalized bitmaps of handwritten digits from a preprinted form.

Make sure the scanned picture (X x Y) is dividible into nonoverlapping blocks of

4 x 4. The value in the block should be in range from 0 to (X+Y)/4.