COL-774 Assignment-2

Q1)

a)

For Linear Kernel:

Q(i,j) = -1/2\*Y(i)\*Y(j)\*X(i,:)\*(X(j,:)');

For Guassian Kernel:

Q(I,j) = -1/2\*Y(i)\*Y(j)\* (exp (-gamma\* (X(i,:)-(X(j,:)))\* (X(i,:)-(X(j,:)))' ) );

Number of Support vectors are printed to be 151 for linear Kernel.

Number of Support vectors are printed to be 536 for Guassian Kernel.

b)

Using formulas derived in class w and b vector is calculated.

Accuracy of linear Kernel = 91.3% and b is stored in file ‘adata.dat’

c)

Accuracy of linear Kernel = 89.3% and b is stored in file ‘a\_data\_guassian.dat’

d)

Accuracy of linear Kernel = 91.3% and no. of support vectors = 151

Accuracy of linear Kernel = 89.3% and no. of support vectors = 534

Q2)

a) Image can be displayed upon running the script.

b) Stopping criterion is when change in error term becomes less than 0.0001 or number of iterations becomes more than 6\*m == 70K order

c) Accuracy comes to be 94.43%.

d)

No. of output units taken are 10 i.e. corresponding to each digit.

Accuracy over test set = 87.3% after 73208 no. of iterations

Q3)

a) Upon executing ‘q3.py’ file, corresponding to no. of nodes explored errors of each test set, validation set, and training set get plotted.

b) The Information gain performs better than previous one. As the above one having large no. of nodes tend to over-fit.

c) Graph obtained after pruning is plot-ed after obtaining data in q31.py file as pacc array variable. Clearly the error decreases after pruning of nodes.

d) The epsilon error can be placed to rectify so or some kind of averaging can be done.