GNR652 Assignment-2

Training SVM for detecting Credit Card fraud

Method Used: 200 samples were randomly taken from given data(100 each for positive and negative class of samples) out of which 80% samples are used for training and 20% for testing and SVM model was trained and tested 5 times. cvxopt.solvers.qp() method has been used for obtaining the optimised solution for values of alpha parameters corresponding to lagrange multipliers.

It was observed that the ‘time’ variable in the give data does not have much correlation with the transaction being fraud or genuine, as accuracy of the trained model seemed to increase when the time parameter was dropped. So in the program attached, time variable has been dropped before training although it can be included in training method just by commenting out lines #12 and #13 of code.

Results:- Follwing are the accuracies obtained for 5 times training and testing of the SVM model each time taking 200 samples from data randomly out of which 80% of the data was used for training and 20% for testing. Average accuracy of 92.5% was obtained in one execution of program although the average accuracy changes slightly every time the program is executed.

Obtained accuracies are as below
[92.5. 95 87.5 90 97.5]
Average accuracy:
92.5%

Note: Please copy the original data file containing credit card data in the same folder as the program with the name “creditcard.csv” before running the python program for testing.