Assignment 4

Documentation:
These are some of the properties of our model:
svm_type c_svc
kernel_type rbf
gamma 1e-06
nr_class 2
total_sv 477
rho 9.51586
label 1 0
probA -2.53718
probB -0.118697
nr_sv 240 237
SV

Kernel Choice : RBF Kernel
C value : 512
Gamma : 0.000001

Why we chose this kernel ?
Upon researching for the guides to choosing a kernel based on the number of feature values, we got to learn that RBF kernel is best suited in the case when the number of features is small and we have very large number of instances of training data.

C and Gamma values ?
We tried the accuracy results using cross validation for a wide range of C and Gamma values. Accuracy stayed Maximum and constant on hitting the Gamma : 10^-5 to 10^-7 range and C : 300-600 range. (Accuracy : 99.0556%)

Procedure :

1. Convert Dataset CSV to SVM format
2. Train and save the SVM Model using the converted dataset with the chosen parameters
3. Run prediction on Test CSV file and generate Result CSV file