Artificial Intelligence

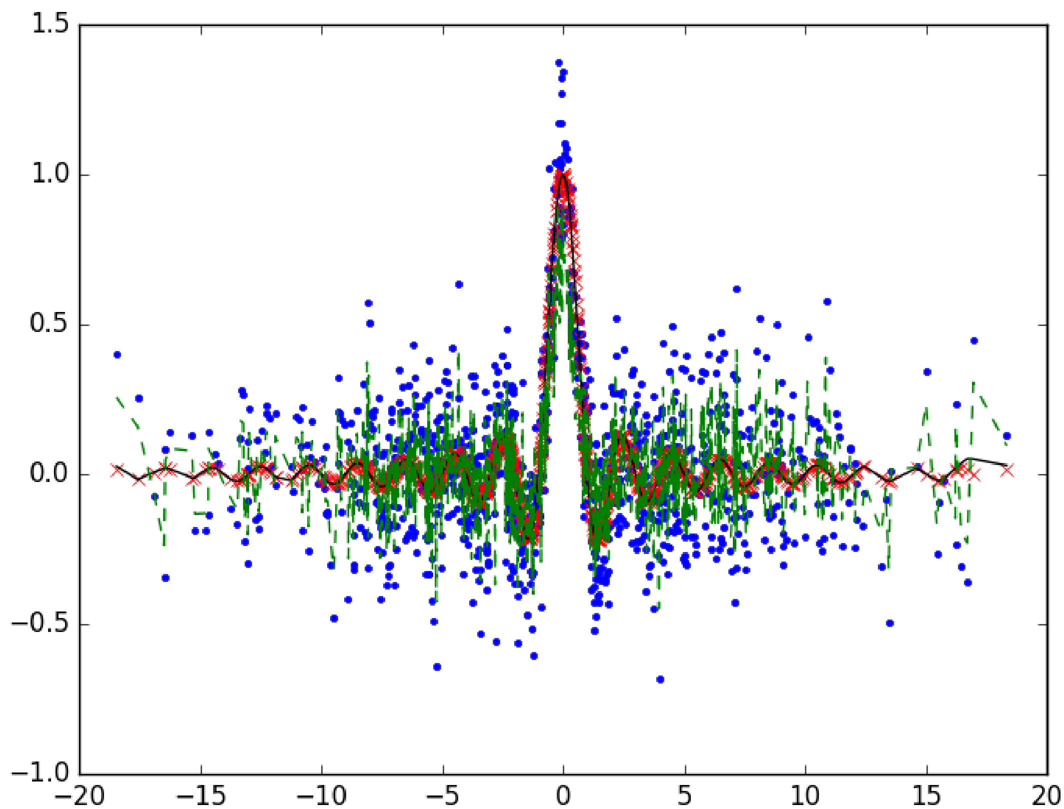
Final Exam

Pooja Jadhav Eshwarlal

1(c). For 1000 samples, the best value of C is 8, gamma is 0.5 and best CV is -0.08.

```
C:\AI\Python27>finalSVMsinc.py
C 0.03125, epsilon 0.0, gamma 3.0517578125e-05. Testing set CV score: -0.373303
C 0.03125, epsilon 0.0, gamma 2.0. Testing set CV score: -0.081988
C 8.0, epsilon 0.0, gamma 0.5. Testing set CV score: -0.081011
[LibSVM]...
Warning: using -h 0 may be faster
*.............................................
Warning: using -h 0 may be faster
*
optimization finished, #iter = 49547
obj = -5.229379, rho = -0.031092
nSV = 852, nBSV = 64
Training set score: 0.999999
Testing set score: 0.999919
```

The sinc function plot for 1000 samples is



1(d) The epsilon Ɛ value effects the number of support vectors used for regression function. For large values of epsilon Ɛ fewer support vectors are selected. For example, if a dataset is correctly classified then epsilon Ɛ will be zero. Epsilon can create a realistic fit when data is not precise and noisy. It is always a better choice to opt for lower values of epsilon Ɛ to achieve high accuracy over the training set.

The value of C indicates the number of mismatches that the SVM can ignore while learning. In search of optimal margin if we introduce too much correction then many exceptions are accepted which creates an overfit on the data. On the other hand, if C is too small, SVM tries to search for a hyper plane that completely separates the data points resulting in an suboptimal hyperplane whose margin is too small. Yes there is great penalty and there is no room for error without penality.

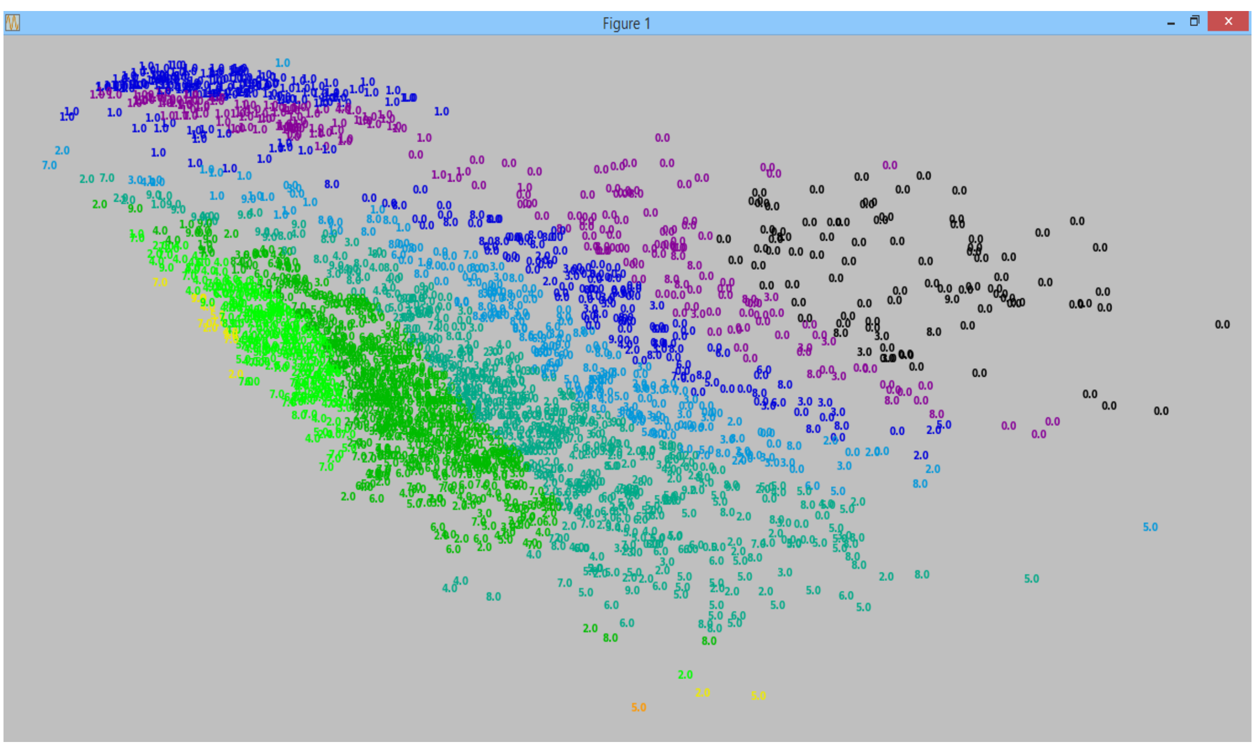
1(e) For 10,000 samples the best value of C is 0.03125, gamma is 2 and CV score is 0.2963

```
C:\AI\Python27>finalSVMsinc.py
C 0.03125, epsilon 0.0, gamma 3.0517578125e-05. Testing set CV score: -0.298391
C 0.03125, epsilon 0.0, gamma 0.0001220703125. Testing set CV score: -0.275922
C 0.03125, epsilon 0.0, gamma 0.00048828125. Testing set CV score: -0.257188
C 0.03125, epsilon 0.0, gamma 2.0. Testing set CV score: 0.296300
```

2(b). The best hyper-parameters are C=4096, gamma=0.0625 and final validation score is 0.336699. The plot we obtain is shown below.

In this image, we observe that there are groups of clustered digits. Each cluster is displayed with a distinct color. In total we observe there are 9 shades of colors. (black, purple, blue, cyan, turquoise, dark green, light green, yellow and orange.) In the figure the colors represent the prediction and the digits represent the true value.

In the black cluster and purple, the predicted digit is zero. In blue cluster,8 is predicted. In cyan, 8. In turquoise,7 . In dark green and light green 7 is predicted and orange has single number in its cluster which is 5.



2(c). In Support Vector Machines, when we have non-linearly separable data, we consider the penalty parameter C to approximate the total training error. C controls trade-off between errors on training SVM and margin maximization**.** For large values of C, large penalty is assigned to errors as a result we store more support vectors which causes an over-fit. Yes there is great penalty. No there is no room for error without penalty.