CS771A: Machine Learning Techniques

Assignment #2

Instructor: Prof. Harish Karnick

Deepak Kumar(12228)

1 Question 1

This problem is based on Hinge Loss which is defined as

$$L(y) = max(0, 1 - t.y) \tag{1.1}$$

here t = +1 or -1 : the intended output
y is the actual raw output from the decision function
NOTE : Hinge Loss is convex function

1.1 Result

Performance with various values of C (Tuning C) :

Value of C	**Performance (in %)**
0.01	98.96
0.1	98.44
0.5	98.27
1, 10, 100, ..., 10000000	98.27

Hence I choose the value C=0.01 for the 'Hinge Loss' and C=1000000 for standard formulation of SVM (Large Margin Classifier)

5-Fold cross validated result for hinge loss in the SVM Classifier over the Spam dataset
Library used : sklearn.svm.LinearSVC

Validation Set	**Perf (in %) Binary Hinge**	**Perf (in %) Hinge Loss**	**Perf (in %) Binary Standard**	**Perf (in %) Standard Formulation**
1	99.654	99.308	99.481	99.481
2	99.138	99.448	99.483	99.793
3	99.962	98.789	99.135	98.270
4	98.937	98.791	99.136	98.273
5	99.481	99.308	99.308	98.962
Average	**99.434**	**99.129**	**99.309**	**98.956**

2 Question 2

Dataset : Connect Four Game positions with final outcome (win/loss/draw) for the first player.
Train SVM over the given dataset and report the results for
1. One-Vs-Rest
2. One-Vs-One

2.1 Result

5 Fold Cross Validated results are as follows : OVO : My Implementation without Multiclass SVM Use (Classification based on most occuring class)
OVR : My Implementation without Multiclass SVM Use (Classification based of largest distance from margin plane)
SVM OVO : sklearn.svm.SVC(kernel='linear')
SVM OVR : sklearn.svm.LinearSVC(kernel='linear')
Total Number of Training Vectors = 67557

Validation Set	**Num Train**	**Num Test**	**OVR**	**SVM OVR**	**OVO**	**SVM OVO**
1	54046	13511	76.863	76.863	76.826	76.693
2	54046	13511	81.267	81.267	80.949	81.141
3	54045	13512	70.648	70.648	70.722	70.878
4	54046	13511	76.116	76.116	76.153	76.316
5	54045	13512	68.058	68.058	67.784	68.014
Average	——	——	**74.590**	**74.590**	**74.487**	**74.608**