PROJECT PART-2 REPORT

(rsanikom@asu.edu)

Introduction: SVM
Using supervised machine learning, the support vector machine (SVM) technique may be applied to both classification and regression tasks. By increasing the margin between classes, the algorithm, which is a discriminative classifier, determines a decision border between them. In large dimensional spaces, SVMs perform better and may be used to invert nonlinear-functions.

Problem Statement:
We must use multi-Class SVM (libSVM) to categorize these 50 Categories.

Input:
The 50 categories/classes, 3 Features, train data of 4786 samples, test data of 1833 samples.

Output:
Accuracy measured when training the data.

Method:
In this project we use libSVM library for implementation in python.
svmtrain trains the svm based on the parameters given and gives an output.
svmpredict returns the output by taking the inputs, outputs and some parameters.

Step 0:
Part-1: Calculated the accuracy of 3 features.
Used the svmprob, svmparameter APIs and testdata to test the model and obtain accuracies.

Accuracy of X1	Accuracy of X2	Accuracy of X3
10.7807%	16.6755%	8.92193%

Part-2: Calculated the accuracies using the posterior probability, and after training the data with parameters tested the data for accuracies.

Accuracy of X1	Accuracy of X2	Accuracy of X3
27.881%	27.5093%	28.7839%

Step-1:
Posterior Probability = (posterior Probability X1+ posterior Probability X2+ posterior Probability X3)/3

Calculated the average of all probabilities obtained at step-0 for all features and chose the class that has largest posterior probability as the class of the training sample.
Accuracy: 44.71587%

Step-2:
Calculated the accuracy by all the values of X1, X2, X3 and trained the SVM using the parameter "-c 10 -t 0"
Accuracy: 37.0685%

Conclusion: The average of the posterior probability for each feature produces the results with the maximum accuracy.