CS771: Machine Learning: Tools, Techniques, Applications
Assignment 3 : Decision trees and forests, ensembles

Divyanshu Shende
Roll No.: 13264

April 3, 2016

1 Decision Tree Classifier

Here are the result obtained on using Decision Tree classifier on MNIST data. The score was on different runs is as follows.

Score = $\frac{correct\ predictions}{total\ predictions}$

Run	Score
1	score = 0.741
2	score = 0.742
3	score = 0.748
4	score = 0.75
5	score = 0.741

Average Score = 0.7444

2 Random Forest Classifier

Random Forest Classifier was trained with varying number of trees and fixed max depth (= 4 for comparison with Adaboost). The results are as follows. Score = $\frac{correct\ predictions}{total\ predictions}$

Number of Trees	Score
50	0.779
100	0.793
200	0.787
300	0.783
400	0.788
500	0.785

Best Classifier : Uses 100 trees in forest with Accuracy = 0.793.

3 Ensemble Methods : Adaboost

Adaboost Classifier was used with varying number of Estimators. Here are the results. Score = $\frac{correct\ predictions}{total\ predictions}$

Number of Estimators	Score
50	0.866
100	0.899
200	0.907
300	0.921
400	0.925
500	0.927

Best Adaboost Classifier : Uses 500 estimators with Accuracy = 0.927
Conclusion : Adaboost classifier outperforms Random Forest Classifier for fixed depth. However, I observed that if the depth for Random Forest is unbounded, then it slightly outperforms Adaboost.