Machine Learning for Applications in Computer Vision: Week 1

Neeraj Sujan - 03656452

23 April 2015

Exercise 2: Support Vector Machines

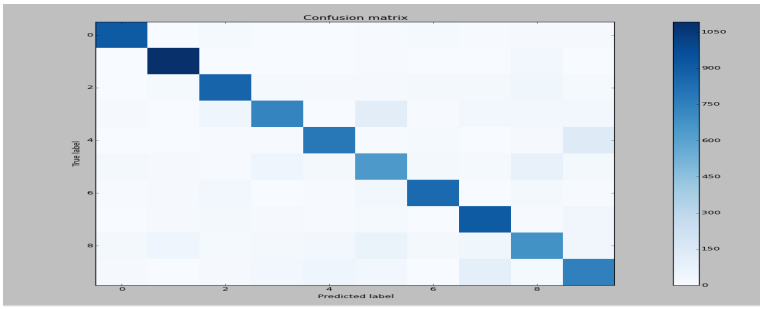
Results:

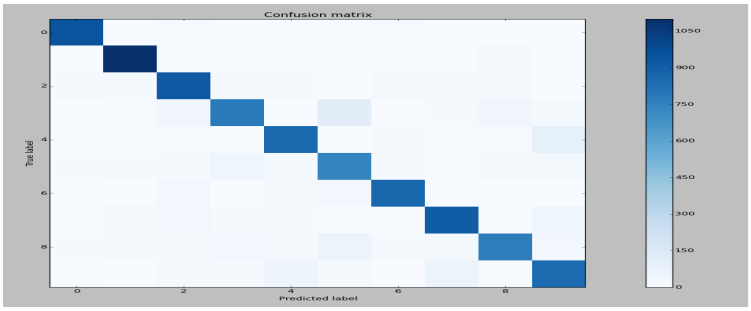
The training data set is stored in images and labels. The test data set is stored in images1 and labels1. The tests were carried out for the first 1000 samples using linear kernel,polynomial kernel and RBF kernel. The results of the tests are given below.

I got the best performance using the polynomial and linear kernel whereas for the RBF kernel, I got the worst performance. The performance metric is as displayed below

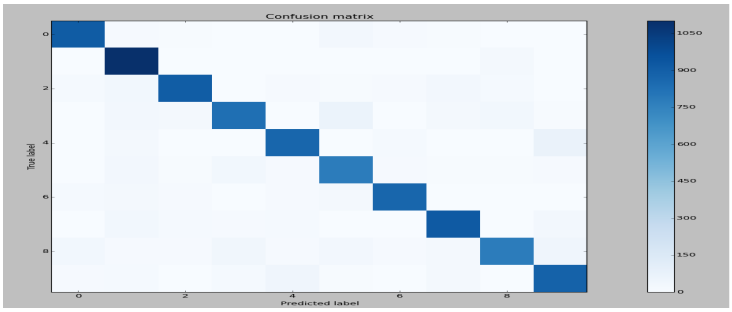
Classifier	LinearSVC	SVC,linear kernel	SVC,polynomial kernel	SVC,RBF kernel
Accuracy	0.8235	0.8758	0.8865	0.1028

Linear SVC

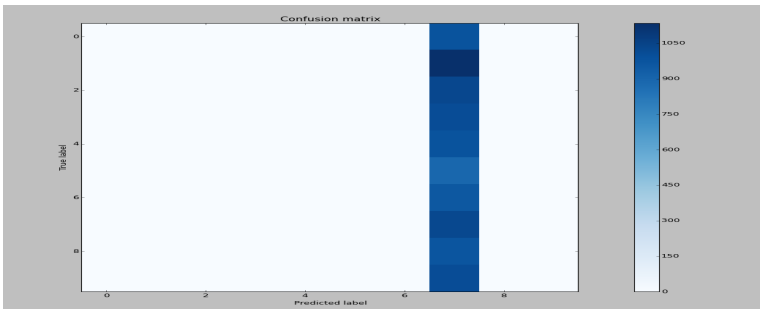




Linear Kernel
Polynomial kernel



RBF kernel



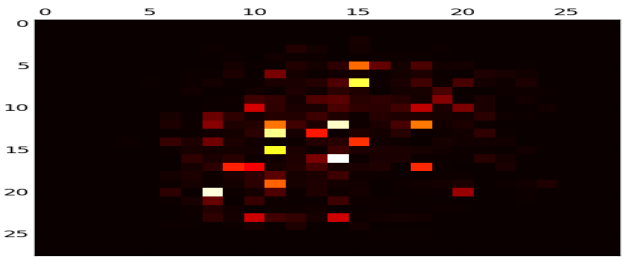
Exercise 3: Decision Tree Classifier

The accuracy of the decision tree classifier using default parameters is 0.8763. The following table gives the precision of each and every label.

Labels	Precision
0	0.92
1	0.96
2	0.87
3	0.82
4	0.88
5	0.83
6	0.89
7	0.92
8	0.82
9	0.84

For the following attributes, $criterion =' gini', maxdepth = 12, maxfeatures = 370$, I got the best performance of 0.8791.

The figure below shows the pixel importance.



Cross Validation: Using the same values, for cross-validation, I got a mean score of 0.87

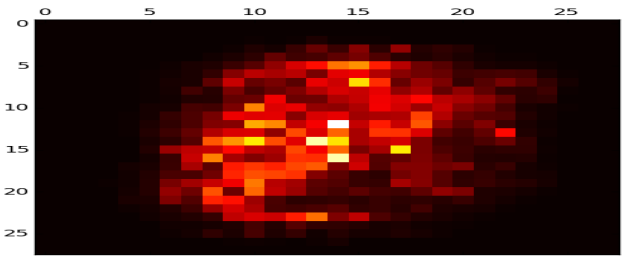
Exercise 3 : RandomForestClassifier

The accuracy of the random forest classifier using default parameters is 0.9463. The following table gives the precision of each and every label.

Labels	Precision
0	0.95
1	0.98
2	0.94
3	0.91
4	0.94
5	0.93
6	0.97
7	0.96
8	0.94
9	0.93

For the following attributes, $nestimators = 25, criterion = "gini", maxdepth = 12, maxfeatures = 30$, I got the best performance of 0.9539.

The figure below shows the pixel importance.



Cross Validation: Using the same values, for cross-validation, I got a mean score of 0.94

Conclusion: Random Forest Classifier gives a better performance in classifying the digits correctly.