

EE 559 Project

Classify Forest Cover

05/10/2014

Shankhoneer Chakrovarty

chakrova@usc.edu

Tools

Prtools description <Expand later>

Baseline Performance Measure:

Procedure:

To get the baseline performance measure, *qdc* classifier was used which is provided by prtools. 5 times 5-fold cross validation was performed to estimate the performance on the training set.

Baseline performance data:

Classifier	Average Error Rate	Standard Deviation
qdc	84.51	0.11

So we see that the baseline performance error rate is $84.51 \pm 0.11\%$. So the accuracy is $15.49 \pm 0.11\%$

Confusion Matrix is presented below:

True Labels	Estimated Labels							Totals
	1	2	3	4	5	6	7	
1	0	0	204	324	0	0	12	540
2	0	0	113	421	0	0	6	540
3	0	0	3	537	0	0	0	540
4	0	0	1	539	0	0	0	540
5	0	0	27	513	0	0	0	540
6	0	0	0	540	0	0	0	540
7	0	0	410	89	0	0	41	540
Totals	0	0	758	2963	0	0	59	3780

Comment on the above results:

'qdc' classifier gives really poor result on the given dataset. This could be due to following reasons:

1. Data is not normalized.
2. 54 dimension may be too 'complex' for the classifier to classify fairly accurately, we need to reduce the dimension to see which dimension gives the best performance.