

Report - Random Forest

Number of trees	Depth	Accuracy	Training time	testing time
50	20	68.3	8s	.89s
100	40	70.83	11s	1.1s
10	10	61.93	3s	.5s
20	20	67.23	4.3s	.55s
30	30	68.29	5.05s	0.627s
40	40	72	6.4s	0.69s
50	50	73.48	7.477s	0.8s
100	50	69.45	11.5s	1.1s
100	100	68.8	12.29s	1.1s
50	30	69.777	8.3s	.83s

I would suggest Random Forest Classifier to a client because it is based on Decision Trees which can be easily explained to a client. Also, finding errors/debugging a Decision Tree is easy.

I would suggest 50 Trees which have depth 30 because it takes a reasonable amount to train and test. It also gives an accuracy of around 69-70%

Tried splitting the tree at a random value depending on the sample (as shown in the comment on line 50) but it did not give a high accuracy as splitting the data at 128.

Used gini index instead of entropy as calculating log, multiplication and addition would be slower than just multiplication and addition.