

Random Forests

1 Accuracy of Random Forest

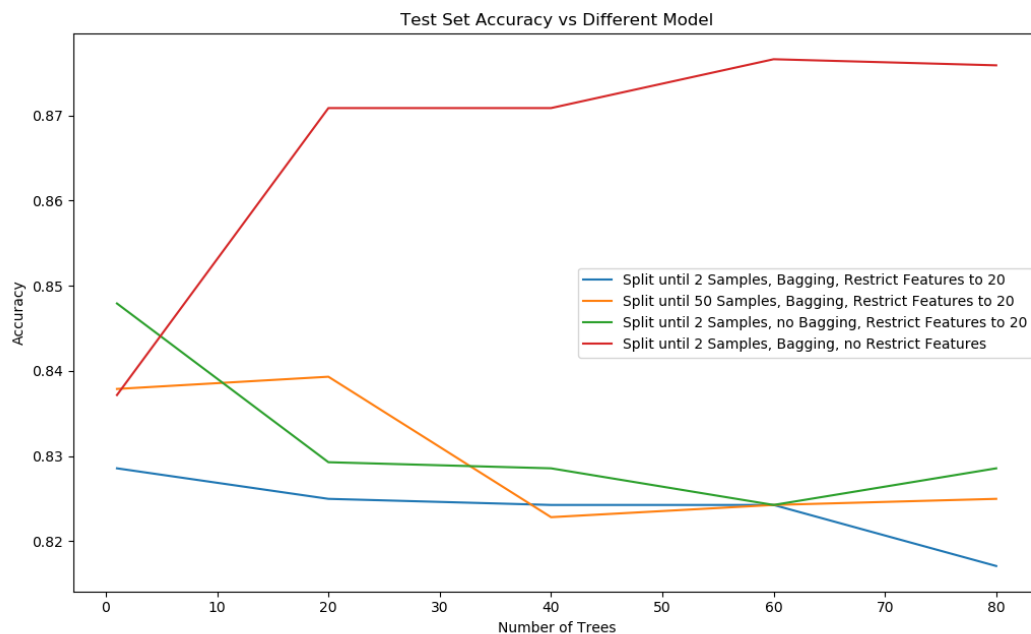
After built the model with the required parameters, we can get the accuracy of the Full Random Forest as: 84.0029%. And for each individual Decision Tree, the accuracies are:

83.7159%	83.2855%	83.2138%	81.9943%	83.5007%
83.3572%	82.5681%	84.1463%	84.5050%	84.4333%

The reason of the accuracy is below 90% is because we add the noise into the dataset. Without the noise, the accuracy of Full Random Forest can be around 97.5%.

2 Compare Different Models

By built the 4 different models, sweep the **numTrees** parameter will get the following accuracy curves. For the program raw output, please check file *Result - RandomForests.txt*.



From the data and image, we can find that:

- As the increase of **numTrees**:
 - More features will let more accuracy
 - Less features will let less accuracy
- Bagging will increase the randomness of the training dataset, this will decrease the accuracy of the model