

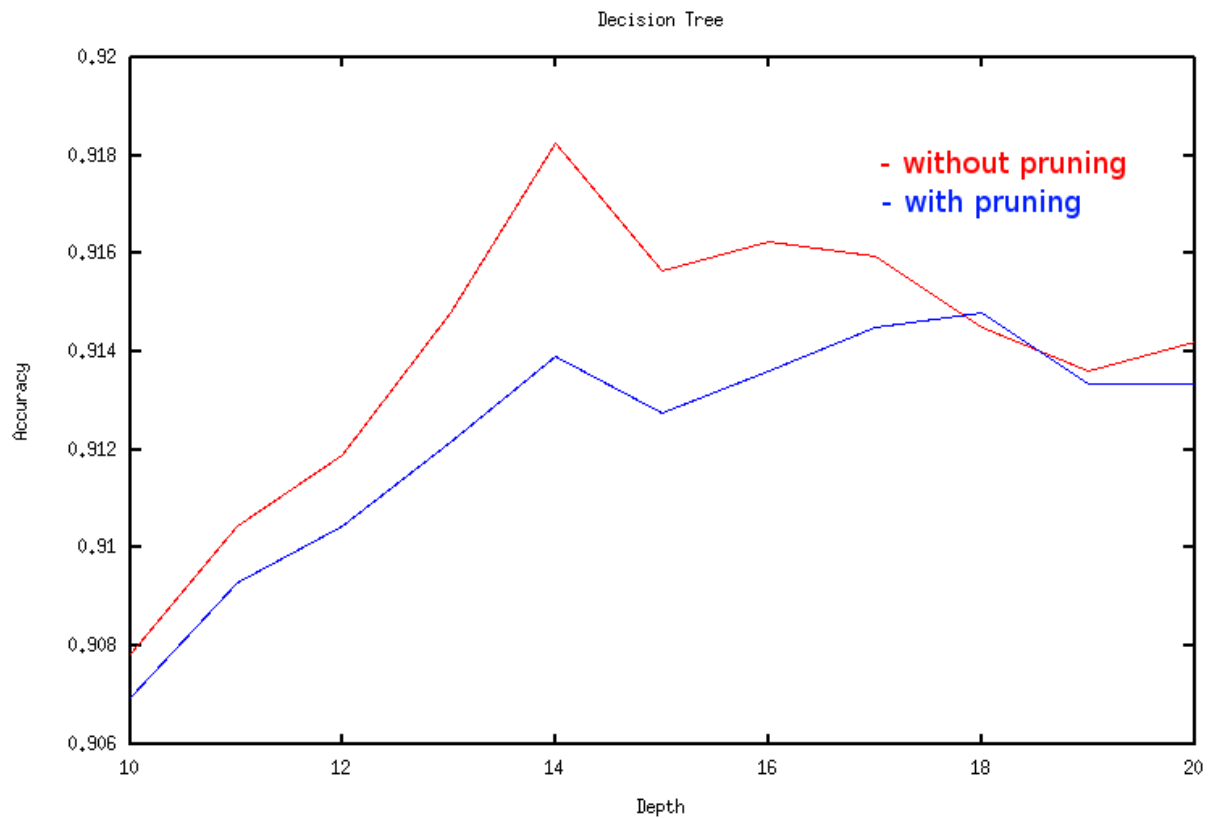
Decision Tree

A regular decision tree with or without pruning based on the input parameters.

Features: Optional Pruning

Parameters Used: depth 11

Performance: 94.6%



Here is a graph with 11 data points. The accuracy increases as we increase the search depth.

X-squared pruning doesn't seem to be working optimally in our tests.

Random Forest

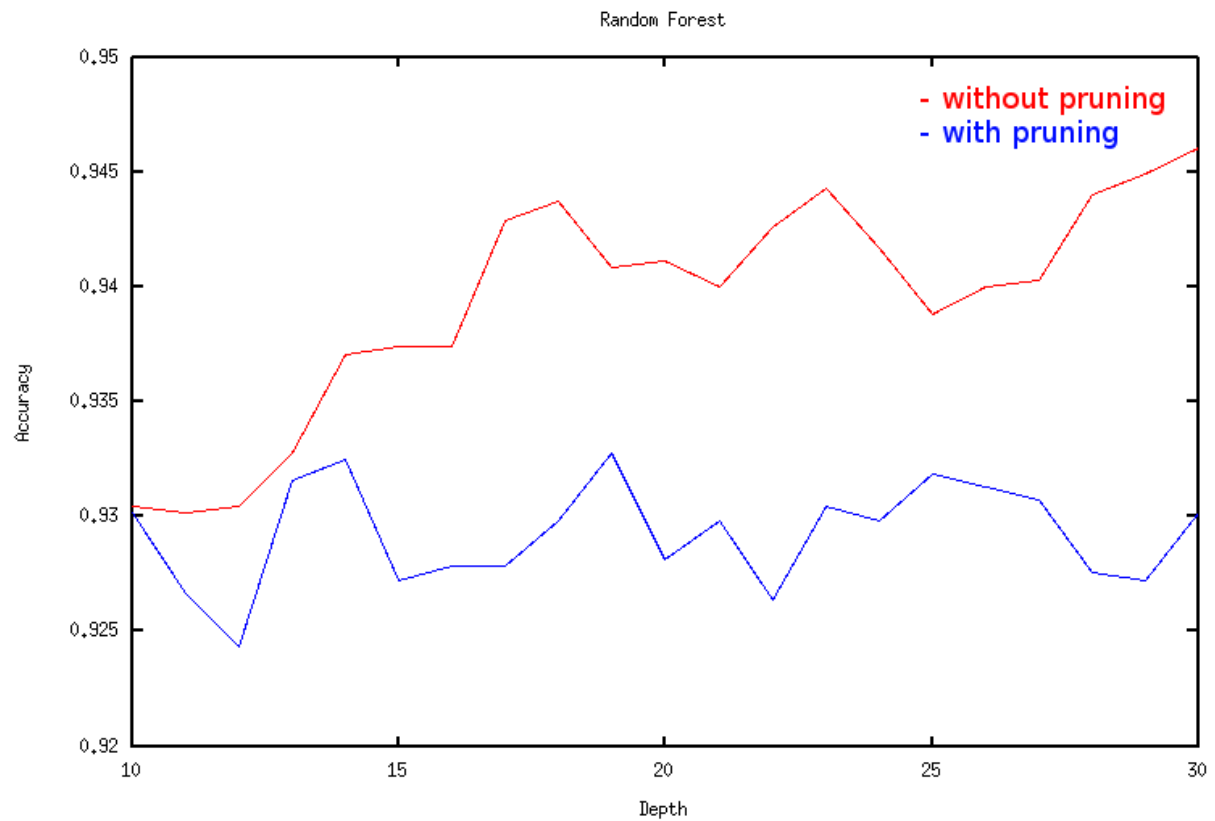
A random forest chooses to split on the best feature from a random subset of features at each depth. It also trains with a random subset of training data for each tree.

Settings: - random (ceiling of $57^{0.5}$) ~ 8 features each round

 - random 66.7% of training data for each tree

Parameters Used: depth 27

Performance: 95.1%



Here is a graph with 21 data points. We used 15 random trees. The accuracy increases as we increase the search depth.

X-squared pruning doesn't seem to be working optimally in our tests.

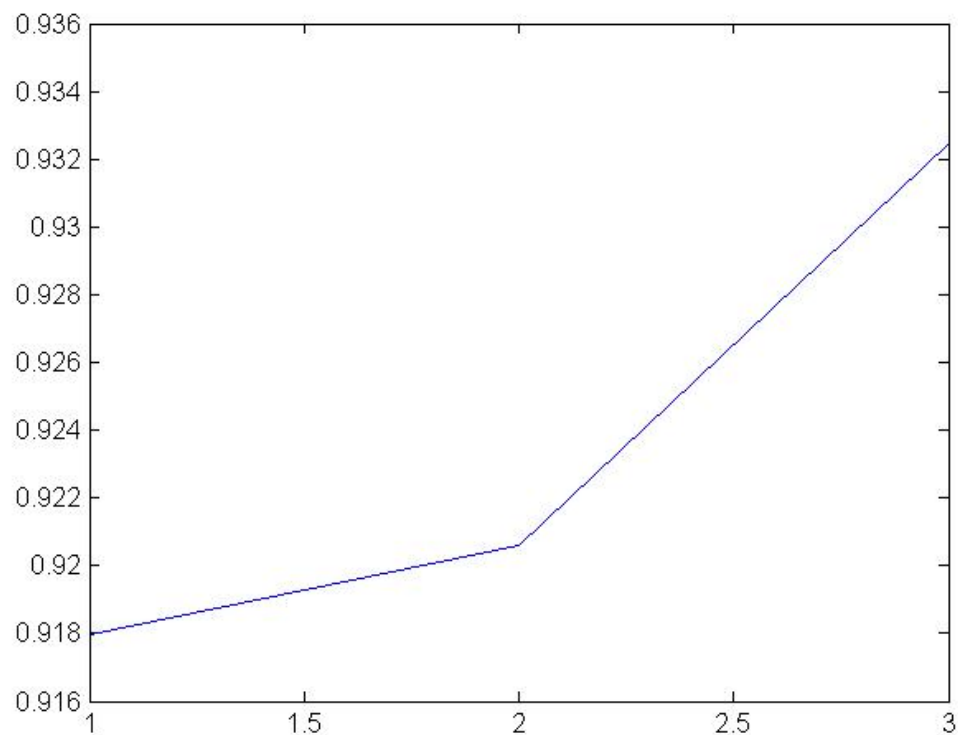
AdaBoost

t weak decision trees by calling dTrees and updating the the sample significance using the error rate of the $i-1$ weak learner to build the i th learner.

Accuracy at depth 1: 91.8%

Accuracy at depth 2: 92.1%

Accuracy at depth 3: 93.3%



Citations:

Most of the implementation decisions made are based on the papers given in the reading directory of this assignment package.

The extra sources include:

Random Forests Leo Breiman and Adele Cutler

http://www.stat.berkeley.edu/~breiman/RandomForests/cc_manual.htm