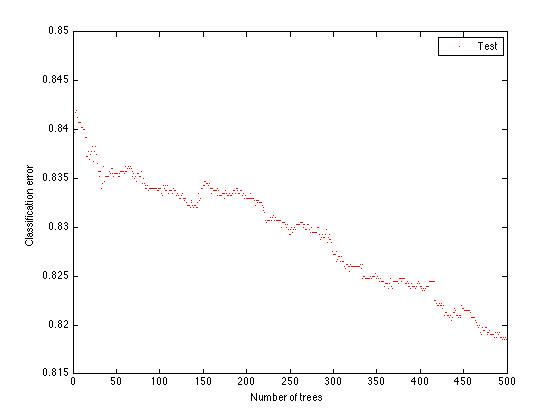
**Milestone 2 Report**

Bagging is known to reduce the variance of an ensemble, while boosting reduces the bias. In training an ensemble of boosted weak learners, we wanted to investigate the phenomenon of the reduction in bias was observed on the CIFAR-10 large scale image classification dataset.

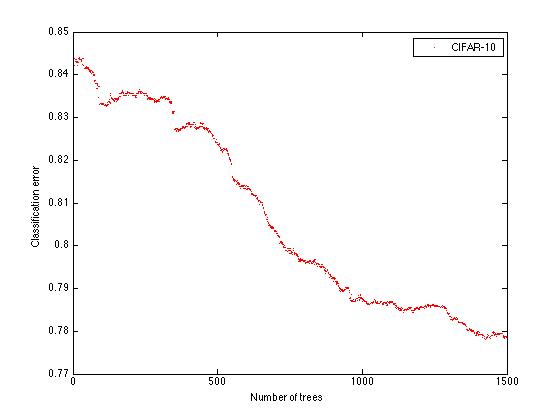
Code based on matlab tutorial on ensemble of trees:

<http://www.mathworks.com/help/stats/ensemble-methods.html#bsvk5ux>

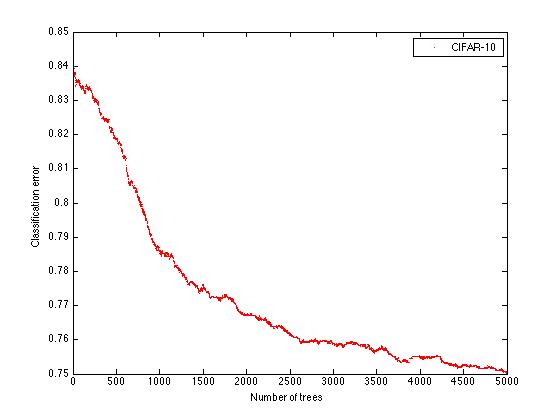
Testing 500 trees as weak learners, and combining via Adaboost, took a training time of 120 minutes, producing an 5-fold cross-validation error rate that decreases linearly with an increasing number of trees.



Repeating the process for 1500 trees as weak learners, and combining via Adaboost took training time of 197 minutes, with 5-fold cross-validation error rate that continued to decrease linearly.



Training on 5000 weak learners took 12.93 hours, and produced the 5-fold cross-validation error of 0.75.



In lecture 8, Neural Networks and Deep Belief networks were examined and found to perform well on noisy data. [ref]. Deep Belief networks work by having each layer’s output be consecutively fed as features to the next layer in the neural network. Convolutional Neural Network was briefly touched in lecture 8 as employing 2 types of units. These are the detection units that learn to recognize some features, and pooling units that combine outputs from spatially close detection units. [ref]. These thus model the retina in the eyes, where there is combination of spatially close detection cells via further specialized cells in the eyes. One state of the art method employs a combination of these techniques, with a further monitoring of the learning that goes on in the different layers of the neural network. To that end, the paper builds an additional metric of companion loss, which is a regularization term that accompanies each layer of the deep belief network.

<http://vcl.ucsd.edu/~sxie/2014/09/12/dsn-project/>