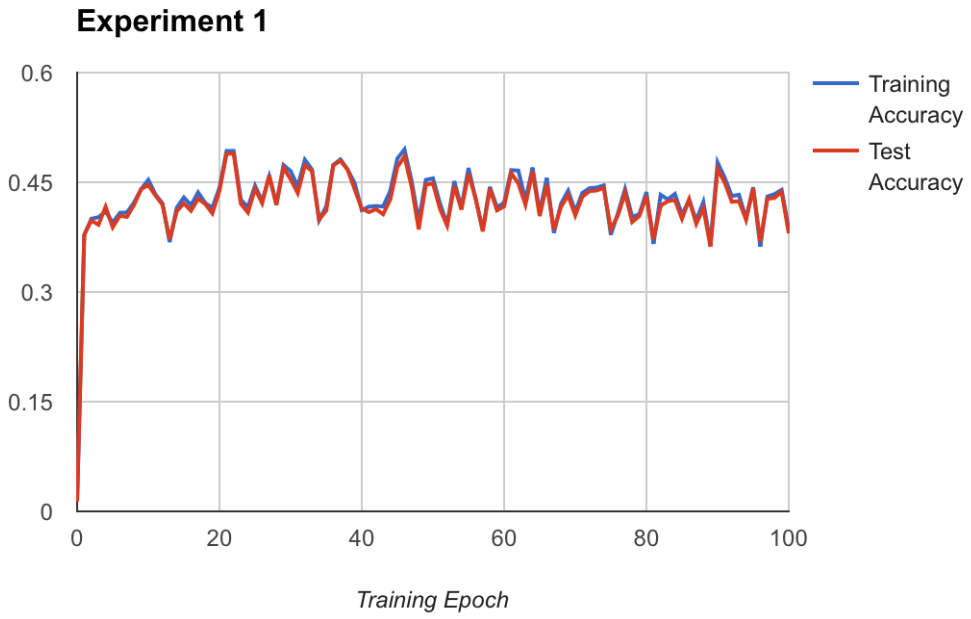
Assignment 2: Multilayer Neural Networks

Ryan Bernstein

February 2, 2016

Each experiment was run with a limit of 100 epochs.

Experiment 1: Control



There doesn't appear to be evidence of overfitting. While both training and test accuracies vary after an initial spike, they don't diverge significantly. In the event of overfitting, we'd expect the training accuracy to be much higher than the test accuracy.

Experiment 2: Effects of Varying Learning Rate

Low learning rates seem to result in higher training and test accuracies than do high ones. In neither case do we see evidence of overfitting. The tests with the lower learning rate have higher overall accuracies than do the control tests, whereas the accuracies from the tests with higher learning rates are lower.

Experiment 3: Effects of Varying Momentum

Notably, the experiments with low momentum appear to center around 45% accuracy, which is similar to the control. Raising the momentum to 0.6 causes a decrease in accuracy, rather than lowering the momentum to 0.05 causing a significant increase.

Experiment 4: Effects of Varying Hidden Layer Width

Here we see our highest divergence yet. The width of the hidden layer appears to make the largest difference in overall accuracy. While we've heard that widening the hidden layer can lead to overfitting, it seems that 8 hidden nodes is not enough to cause this phenomenon.