Yang Xue
Ben Seibold
Adam Storti

Assignment 1 Writeup

Choosing the hyper-parameters for the Neural Network:

We experimented using different lambdas and hidden values for our neural network. For our lamba values, we tested every multiple of 10 from 0 to 60 and kept the number of hidden nodes at 50. Our training and test accuracy peaked at lamba = 10 with a value of 95.32% for training and 95.13% for test. After lamba = 10, our progressively higher lambas resulted in a negative trend in accuracy. All of this can be seen in Figure 1 below.

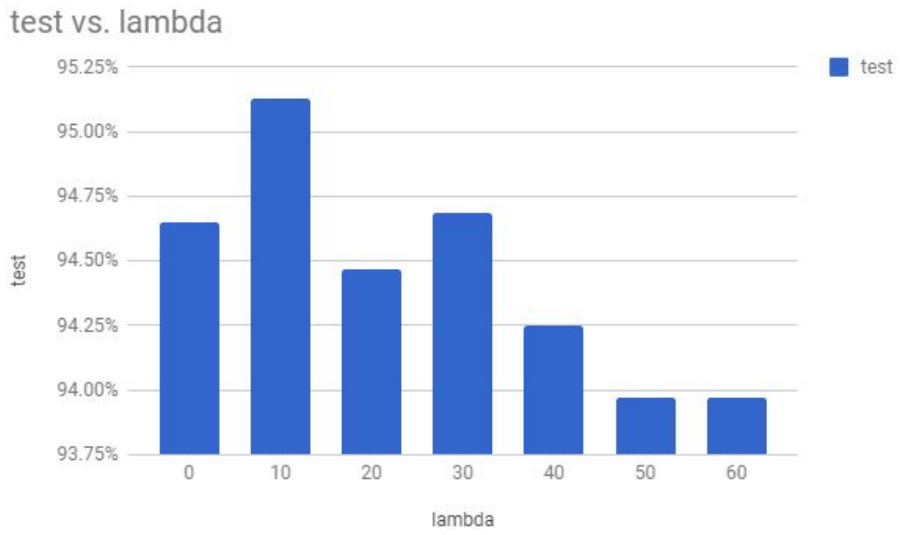

Figure 1 - *Lamba value vs accuracy percentage*

For our hidden values, we tested every number from 4 to 20 in multiples of 4, with lambda = 10. As we increased the number of nodes in the hidden layer, the accuracy increased from 75.78% at 4 to 93.24% at 20. We also did one test at 100 and got 95.34%. As expected, the more number of nodes a neural net has, the more accurate its outputs are. We chose 20 as our number of hidden nodes since it provided a high accuracy at a moderate training time.

Figure 2 *-# of Hidden units vs test accuracy*

The amount of time required to calculate the neural net increases at a much quicker rate than its accuracy does so increasing the # of nodes yields decreasing returns to scale. Thus, we chose to put the number of hidden nodes at 20, which yields 93% accuracy while running in half the time it takes the network of 100 nodes to run. This can all be seen in Figure 3 below.

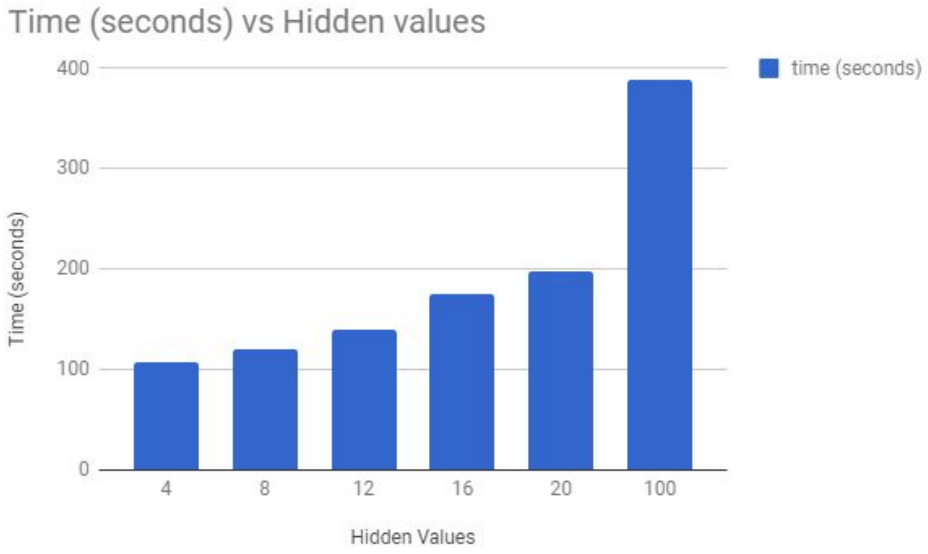

Figure 3 *-# of Hidden units vs processing time*

Accuracy of classification on handwritten digits:

For handwritten digits we had a high accuracy of over 90%. The training accuracy was 93.38%, the validation accuracy was 93.19%, and the test accuracy was 93.14%. On MNIST_sample.mat, the training time was 15.243 seconds. This time went up to 250.19 seconds on MNIST_all.mat. This is because the MNIST_sample is much smaller than the MNIST_all.

Accuracy of classification on quick draw:

For quick draw, our accuracy was much lower. The training accuracy was 67.7825%, the validation accuracy was 67.34% and the test accuracy was 66.928%. It was much lower because this is a harder problem for the neural network to learn. There is more variation in the input data, because everyone's idea of abstract objects varies much more than everyone's ideas of digits.

Tensorflow with 3, 5, and 7 layers:

For tensorflow we implemented a neural network with 3, 5, and 7 hidden layers. We expected the training time and the accuracy to increase for each hidden layer we added to the neural network. However this was not the case. The training time for 3 hidden layers was 185 minutes and 41 seconds, with a test accuracy of 62.12%. The training time for 5 hidden layers was 129 minutes and 28 seconds with an accuracy of 61.63%. And the training time for 7 hidden layers was 142 minutes and 59 seconds with an accuracy of 60.24%. The training time was sporadic and could possibly be contributed to heavy server traffic. This shows a decrease in accuracy by adding more hidden layers.