Assignment 4 Report

Part 1



*Figure 1. Initial neural net*



*Figure 2. Initial neural net using 20 epochs*



*Figure 3. Final neural net [784,20,10], epochs = 20, eta = 5, mini batch size = 20*

The initial neural net for part 1 had the neural net structure [784,10,10], it ran at 10 epochs, with a mini-batch size of 10 and a learning rate (eta) of 0.1. The accuracy of the final epoch for the initial neural net was 84.58% (Figure 1). My first instinct was to increase the number of epochs to 20, to see how this net will perform given more iterations to learn. This approach increased the accuracy to 89.59% (Figure 2), and the net accuracy seemed to peak around that point. Therefore, I decided next to increase eta, to ensure that the net was not landing in a local minimum for optimization. After experimenting with several values between 0.1 and 10, I landed at eta = 5 as the optimal learning rate. Then, I increased the mini-batch size because my computer can handle the increased processing requirements. This approach got me to around 94% accuracy. To get that last bit of accuracy, I increased the number of hidden nodes in the neural net to 20 to allow for more complexity in the functions that the neural net is processing. These hyperparameters led to the final epoch my best neural net for part 1 to have an accuracy of 95.02% (Figure 3). The following is the code exerts of me creating and training the net:

net = network.Network([784,20,10])

net.SGD(trainingData, 20, 20, 5, test\_data = testingData)

When later evaluating the test data on the trained net, the accuracy was 95.32%. The script fine\_failed\_tests.py found all the tests that the neural net failed. I have included three of them below with the image, the test index, the number guessed, and the actual number.



*Figure 4. MNIST test data index 3503. The neural net guessed that this was a one. It is actually a nine.*

*Figure 5. MNIST test data index 6744. The neural net guessed that this was a six. It is actually a two.*



*Figure 6. MNIST test data index 9893. The net guessed that this was an eight. It is actually a two.*

Part 2

Part 3

initial params: net[9,10,2]

sgd (10,10,0.1)

Middle params: net[9,20,2]

sgd (20,20,1)

Final params: net[9,20,2]

sgd (100,30,0.3)