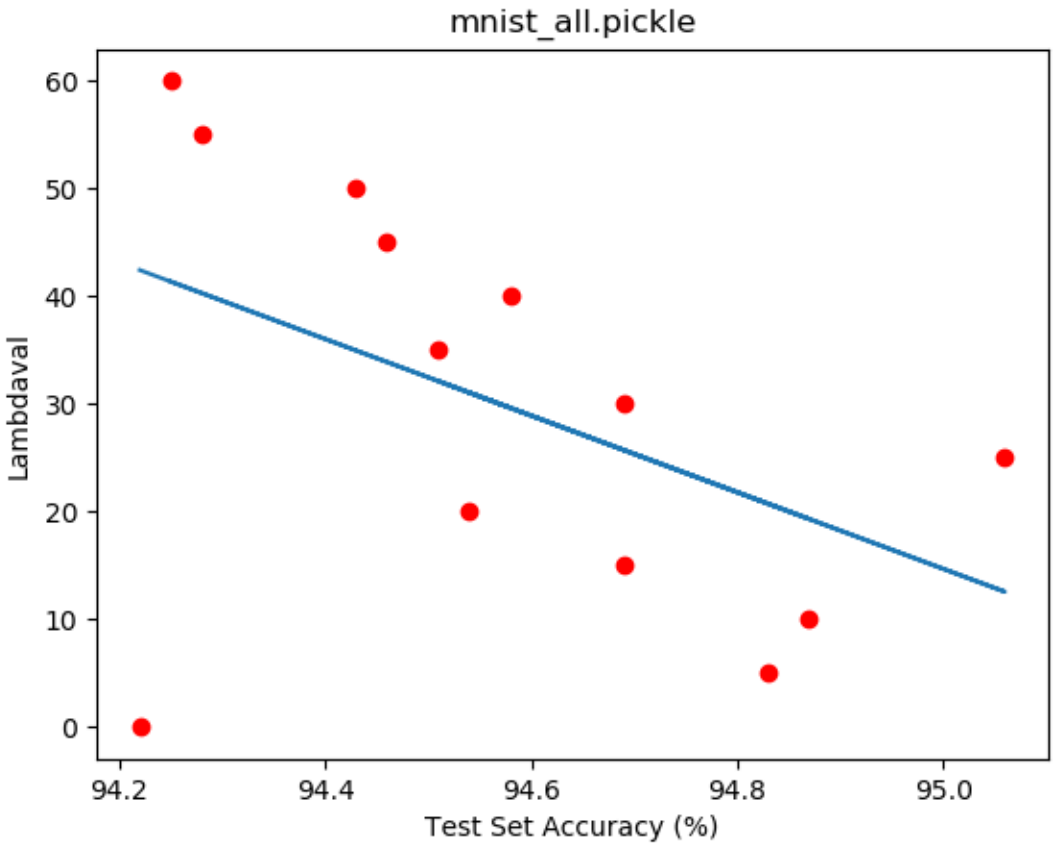
CSE 474 Programming Assignment 2 – Neural Networks

Group 37: Joseph Dubill, Lawreen Latif, Andrew Oslica

Optimal Hyper-parameter

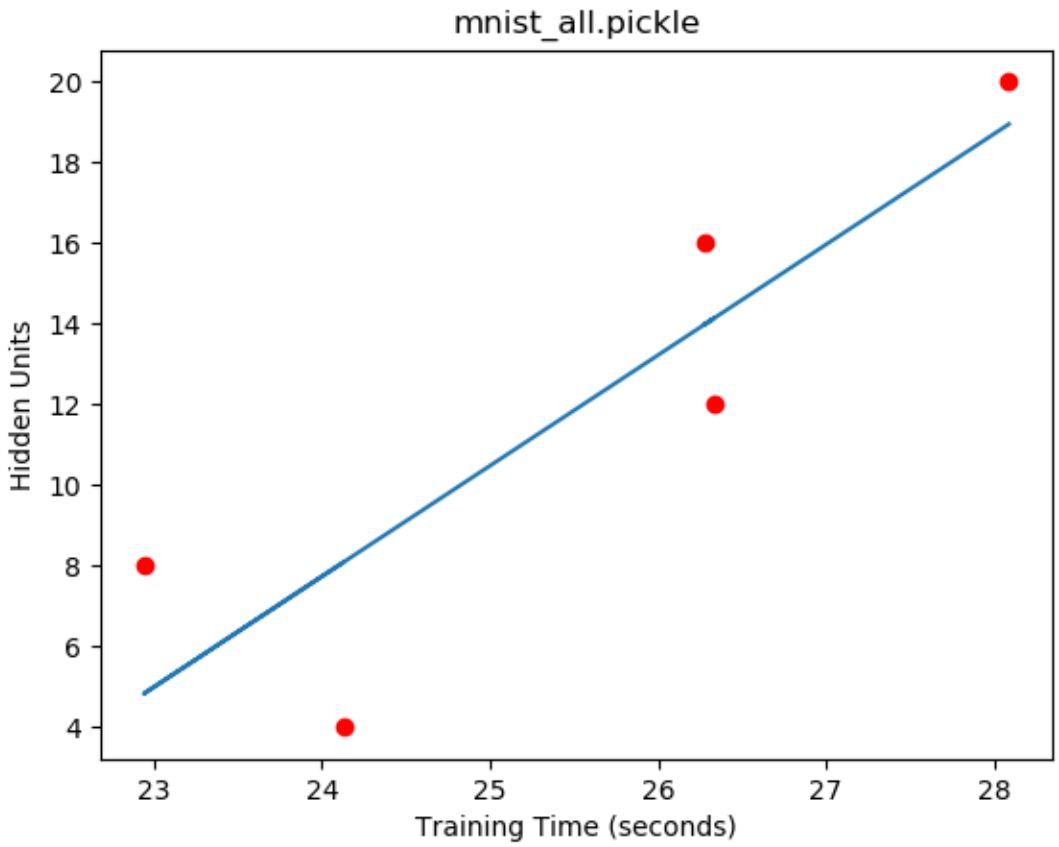
We tested the effects our hyper-parameter has on the accuracy of our test set by fixing the hidden units to 50 and running the neural network a total of 13 times, each time adding 5 to λ, starting at λ = 0 and finishing up at λ = 60.



As we can see, the optimal λ ended up somewhere in the middle with λ = 25 achieving the highest accuracy on the set, although the accuracy only varied by around 1%. The effects of increasing λ began to drop off the higher we went and we began to see worse results in comparison to lower λ values.

Effects of Hidden Units on Training Time

We intuitively thought that as we add more hidden units to our network, the time it will take to train our weights would increase basically linearly. We first found an optimal λ before running our neural network with 4, 8, 12, 16, and 20 hidden units.



Our results proved our intuition correct. Training time simply rose linearly with the number of hidden units being used.

Accuracies

Our Classification Methods

Both of the following accuracies were calculated with the default nnScript.py, meaning hidden_units was set to 50 and our lambdaval was set to 0.

Handwritten digits test data: Test set Accuracy- 94.83%

AI Quick Draw data set: Test set Accuracy- 76.32%

Tensorflow Implementation

2 layers:
[2995/3000] train_loss: 1.13514, train_accuracy: 0.59500 ; test_loss: 1.15828, test_accuracy: 0.62044

Your program finished in 9 minutes 39 seconds!

3 layers:
[2995/3000] train_loss: 1.12997, train_accuracy: 0.63500 ; test_loss: 1.12489, test_accuracy: 0.62752

Your program finished in 10 minutes 45 seconds!

5 layers:
[2995/3000] train_loss: 1.15285, train_accuracy: 0.61000 ; test_loss: 1.16507, test_accuracy: 0.60468

Your program finished in 13 minutes 8 seconds!

7 layers:
[2995/3000] train_loss: 1.28236, train_accuracy: 0.55500 ; test_loss: 1.18986, test_accuracy: 0.60120

Your program finished in 16 minutes 27 seconds!

Interestingly, simply increasing the layers did not necessarily result in higher test accuracies with our highest result being on 3 layers, but it did increase our run time by roughly 3 minutes for every 2 layers we added.