CSE 474 Introduction to Machine Learning

Programming Assignment I

Project by: Corey T Ropell, Jasmeet Kaur Chawla, Joshua T Suttenberg

Introduction

In this assignment, our task was to implement a Multilayer Perceptron Neural Network and evaluate its performance in classifying handwritten digits. We were also able to use the same network to analyze a more challenging hand-drawn images dataset and compare the performance of the neural network against a deep neural network library using TensorFlow.

In this report, we will discuss our code along with accuracy of classification method on the handwritten digits test data along with on the AI quick draw data set and the comparison of the accuracy and training time of deep neural network (using TensorFlow) with different number of layers.
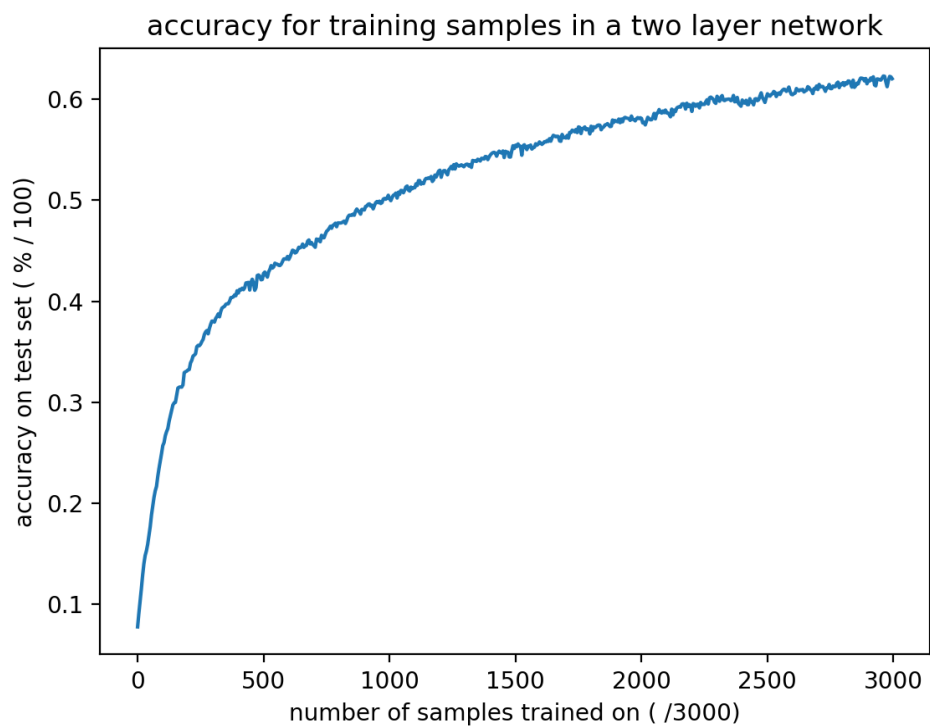
Hidden Layer Testing

AI_quick_draw Data Set with TensorFlow

The Number of Hidden Layer

The introduction of number of hidden layers makes it possible for the neural network to exhibit non-linear behavior. Therefore, to get optimal results, with the data set provided to us we implemented the default(2), 3, 5 and 7 layers.

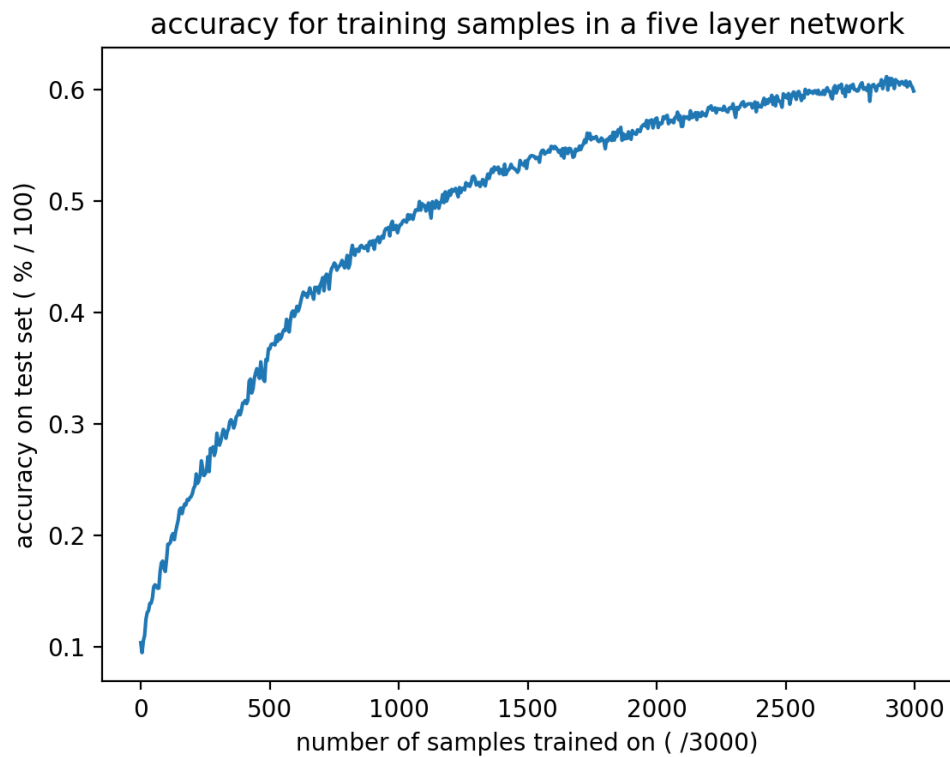We use cross – Validation to test the accuracy of the data set.

The results we got for testing with default of 2 hidden layers: *Accuracy*: 62.02%
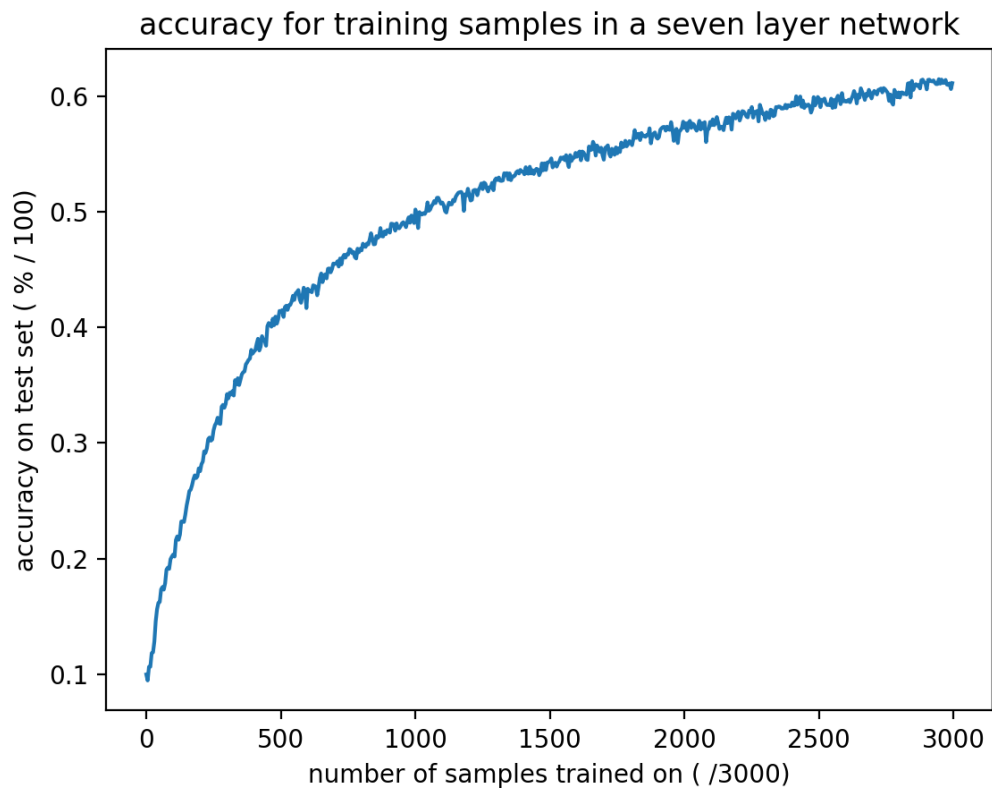
Runtime: 68 minutes 47 seconds

accuracy for training samples in a five layer network
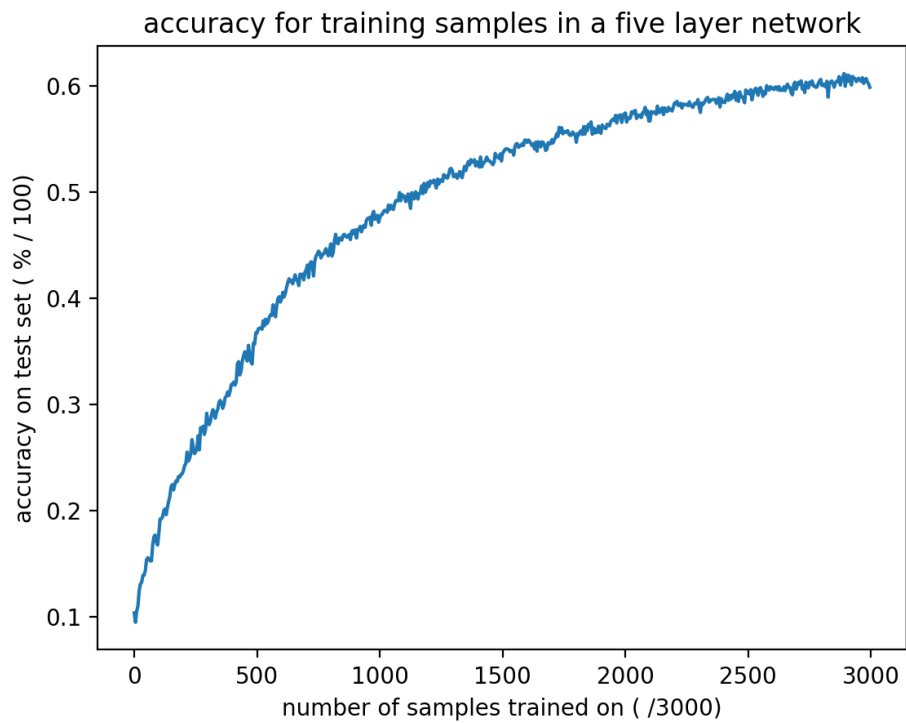
The results we get for 3 layers: *Accuracy*: 61.78% *Runtime*: 118 minutes 56 seconds

accuracy for training samples in a five layer network

The accuracy we get for 5 layers: *Accuracy*: 59.84% *Runtime*: 156 minutes 47 seconds

accuracy for training samples in a seven layer network

The accuracy we get for 7 layers: *Accuracy*: 61.15% *Runtime*: 156 minutes 20 seconds

Compare the accuracy and training time of deep neural network
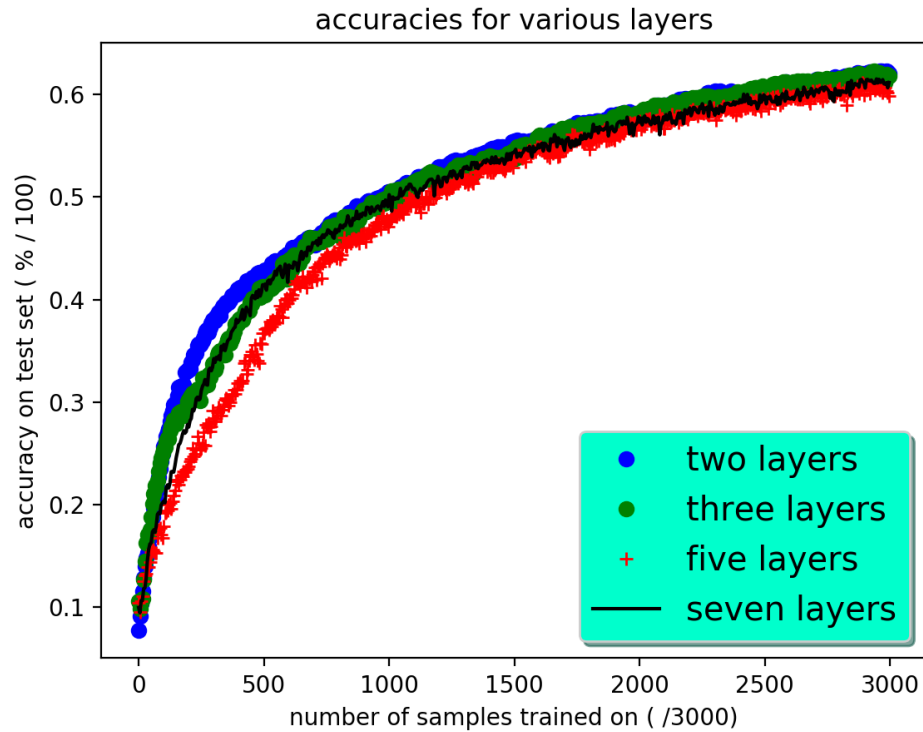
From comparing the accuracies, we can notice that the number of hidden layers seems to almost have an inverse relation with the accuracy. From the data we can see that the neural network preformed best with two layers, and second best with three. It does seem that it performed better with seven layers than with five however. The variation in the accuracies was minor with a maximum deviation of only 2.18%.

Selecting Hyper-Parameter for Neural Network (λ)

When selecting the value for λ there is a lot to be done. This parameter is used in the regularization steps of the networks calculations. This value affects the convergence of the network; thus, it has an impact on both accuracy and runtime. To find the best value for λ test runs should be performed and from the results we can determine the optimal value for λ.

Unfortunately, we were unable to run these tests to determine the accuracies and runtimes of the network for various values of λ. Upon trying to test our code on the Springsteen and Metallica servers we encountered the 'Broken Pipe' error which halted our tests. We attempted to use the 'nohup' command but even still we never seemed to get the results of the scripts.