CSE 574 Introduction to Machine Learning

Handwritten Digits Classification

Team Number - 69

Aneesh Bhatnagar- 50208162

Moonis Javed - 50208261

Surabhi Singh Ludu - 50207139

Choosing the hyper-parameters for Neural Network

Accuracy : 95.52

Number of Hidden Nodes : 70

Regularization Hyper Parameter(λ) : 5

Plot : Lambda vs accuracy

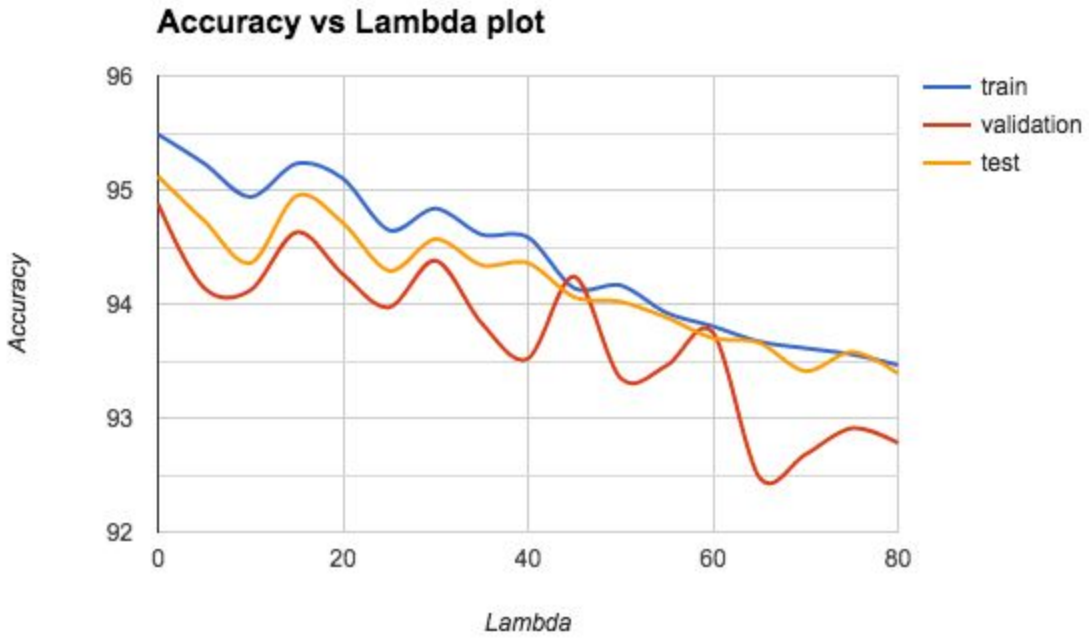

Fig 1. Accuracy vs lambda plot at 50 hidden layers

Plot: Hidden nodes vs run time

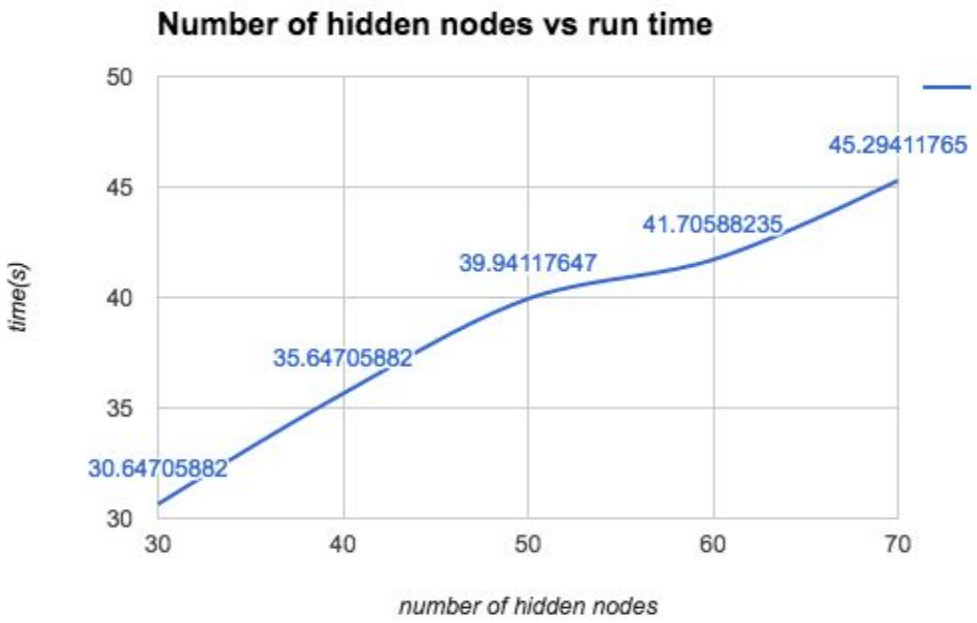

Fig 2 Number of hidden node vs run time averages for all values of lambda

Plot : Hidden nodes vs accuracy

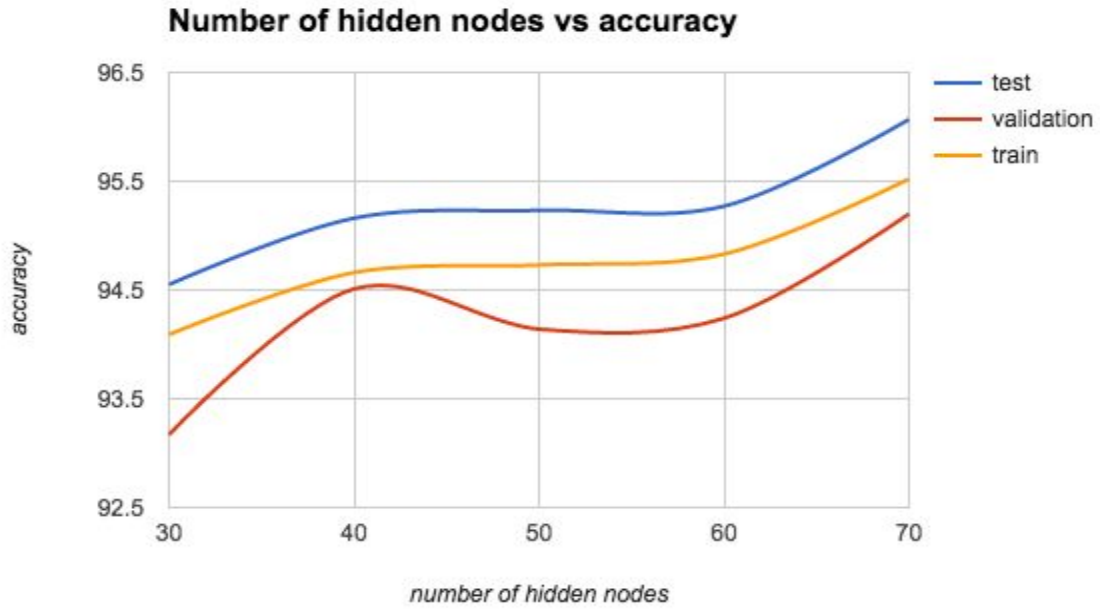

Fig 3. Number of hidden node vs accuracy plot at lambda=5

Explanation :

From the above three graphs we can see that the number of hidden nodes is directly proportional to the prediction accuracy and inversely proportional to time. The increase in the number of hidden nodes in the neural network increased the accuracy of model from 92.6% at 30 hidden nodes to 95.52% at 70 hidden nodes for the test dataset.

The time taken also increased from 30.64 seconds to 40.29 seconds as we increased the number of hidden nodes as the size of the weight matrices increased(time taken is averaged for a fixed number of hidden nodes for varying values of lambda). Thus we can fix the optimal number of hidden nodes at 70 where we achieved the highest accuracy consistently for all three datasets.

In neural networks, regularization is used to reduce over-fitting and increase accuracy by making the model more general. After our experimentation with different regularization values(lambda), we concluded non-regularized runs will usually take large number of iterations compared to regularized runs. In these non regularized runs, the weight vector grows by a large margin compared to regularized runs. For tuning regularization parameter and number of hidden nodes, we observed that the gap between testing and training accuracy is larger when hidden nodes are smaller and the gap reduced for the same regularization parameter, when the hidden nodes were increased. Thus lambda was chosen by testing with different parameters values and we chose the one which gives maximum accuracy in all the datasets result.

After we ran tests for all the values of lambda from 0 to 80 at increments of 5 with varying number of hidden nodes, from 30 to 70, we saw an accuracy of 95.52 for the test dataset with lambda at 5 and number of hidden nodes at 70 though calculations would take more time with more nodes, better accuracy percentage offsets the increased run time. We also noticed that with a fixed number of hidden nodes, accuracy dropped slightly as the value of lambda increased.

Analysis of Deep Neural Network

Accuracy of Classification on CelebA data set

For classification of the CelebA dataset into two classes, the celebrities wearing glasses and the ones not wearing glasses we used the facennscript.py, which gave us an accuracy of **85.31%**. These results were obtained without modifying any parameters in the given script. The functions used here were identical to the ones we used for the classification of handwritten digits in the previous section.

Classification of the celebrities dataset using the deep neural network (using Tensorflow library) yielded less accuracy than the single layer neural network. The comparison between the single layer neural network and multi-layer neural networks is discussed below.

Comparison of Single and Multi-Layer Neural Networks

A tabular comparison between the single layer neural network and the multiple layer neural network, along with the execution times is shown below. All the instances of the test were run on our personal machines. The execution time is only marginally higher than what we got on Springsteen. The single layer execution was done using the facennscript whereas the other layers were tested using Tensorflow in the deepnnscript.

Number of Hidden Layers	Accuracy (%)	Execution Time (m:ss)
1	85.31	8:04
2	80.96	4:25
3	78.08	4:20
4	77.74	5:58
5	74.38	5:06
6	74.15	5:37
7	70.89	6:08

Accuracy vs Number of Hidden Layers in a Neural Network

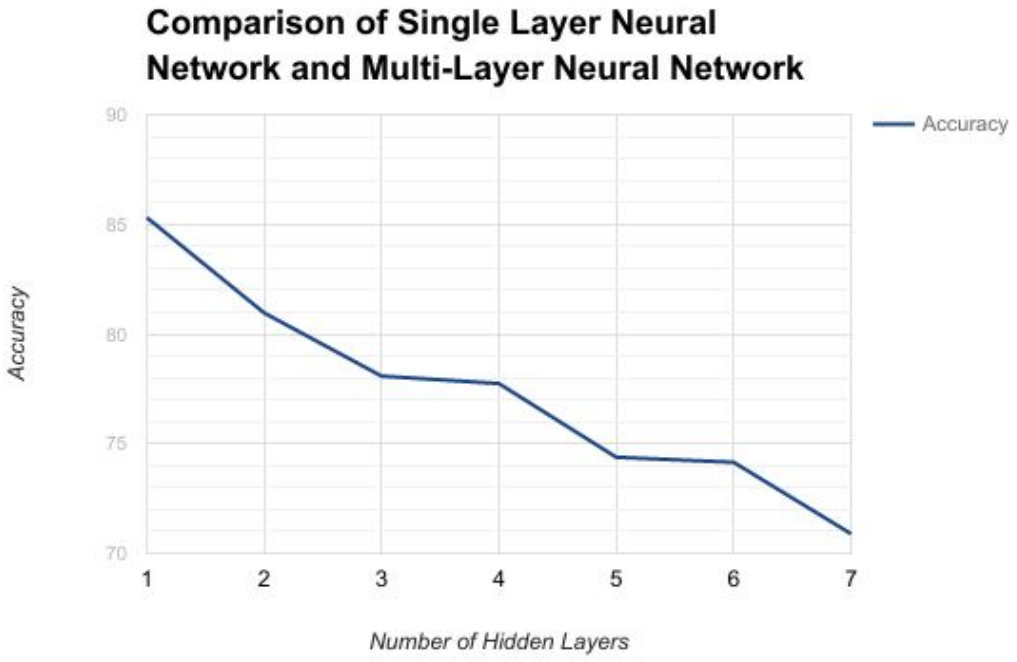

Fig 4. Accuracy vs Number of hidden layers plot

Execution Time vs Number of Hidden Layers in a Neural Network

Fig 5. Number of hidden layers vs execution time plot

Result

After experimenting with the Deep Neural Network using the Tensorflow Library and a single layer neural network we can conclude that the results of the single layer neural network proved to be better compared to a deep neural network. This proves that although the runtime of a deep neural network using the Tensorflow library might be better but the accuracy of the test data was much higher in the single layer neural network. Also, we can say that deep neural networks might be beneficial in some cases or test data but not in every data set.