University at Buffalo

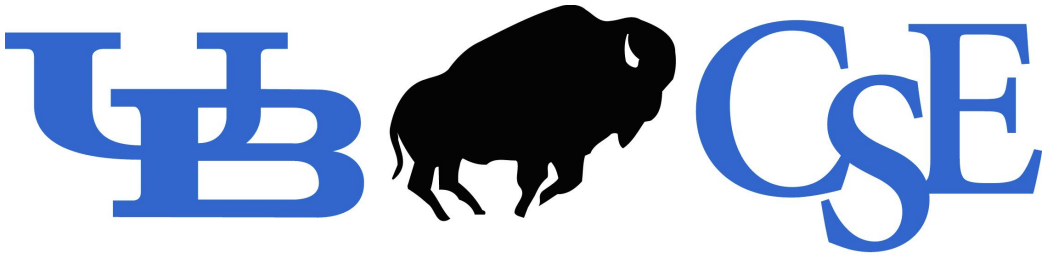
CSE 574 Introduction to Machine Learning, Spring 2017

Programming Assignment 1: Handwritten Digits Classification

Project Report

Submitted by
Anurag Devulapalli (5020 8153)
Sandeep Shenoy (5020 5705)
Vipin Kumar (5020 8397)

March 7, 2017



Department of Computer Science and Engineering
University at Buffalo *The State University of New York*

Contents

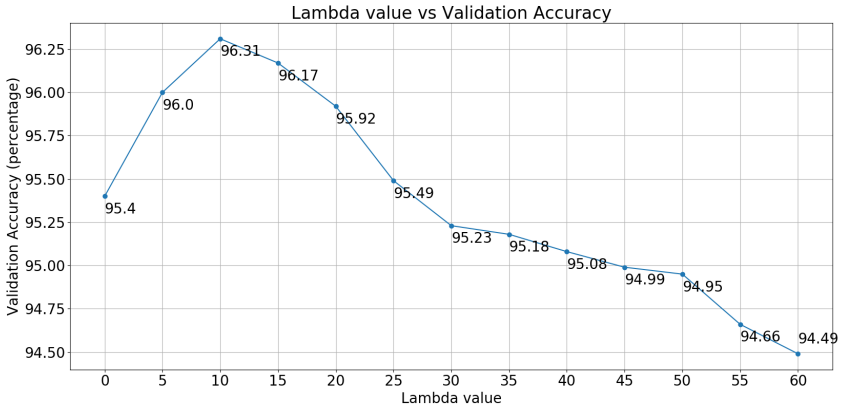
1 Choosing the hyper-parameters for Neural Network

Regularization coefficient (λ):

We start with $\lambda = 0$ and increment it by 5 units every run till 60. We note the validation data accuracy for each value.
Our observations are as follows:

As per the above observations, we note that the validation accuracy is highest when λ is set to 10.

Thus, we set the regularization coefficient as 10.



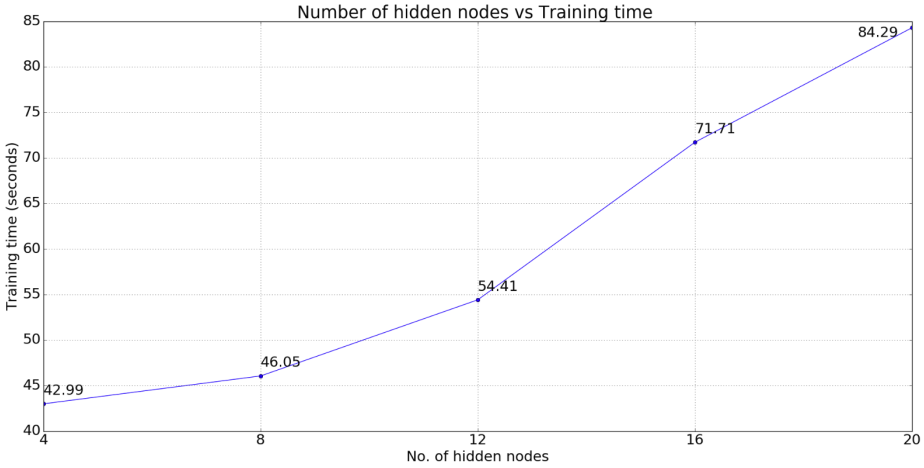
No of hidden units:

We set the λ value as 10 from above. We start with hidden units = 4 and increment it by 4 units every run till 20. We note the training time for each value of hidden unit.

Our observations are as follows:

No. of hidden nodes	Training Time (seconds)
4	42.99
8	46.05
12	54.41
16	71.71
20	84.29

Table 1: Comparison of hidden nodes and training time



As per the above observations we note that training time increases are we increase the number of hidden layers. We also know that accuracy increases as the number of hidden units are increased. Thus for best results we choose to keep the number of hidden units = 50.

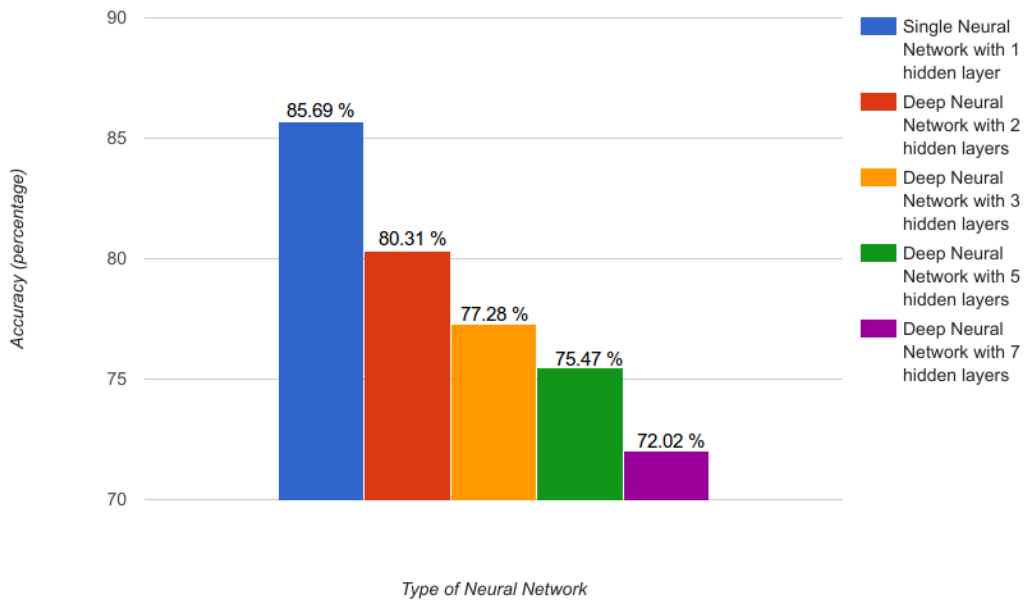
Maxiter:
We set this value to 100 for optimum results.
Final validation accuracy noted by our training: 96.31 %

2 Comparison of performance of single vs. deep neural network

No. of hidden layers / Type of neural network	Accuracy (percentage)	Training time (seconds)
1 / Single Neural Network	85.69 %	113.46
2 / Deep Neural Network	80.31 %	172.09
3 / Deep Neural Network	77.28 %	173.75
5 / Deep Neural Network	75.47 %	203.78
7 / Deep Neural Network	72.02 %	243.5

Table 2: Comparison of Neural Network and Deep Neural Network

2.1 Comparison of accuracy in single neural network and the deep neural network



2.2 Comparison of learning time in single neural network and the deep neural network