

CSE 474 / 574 Introduction to Machine Learning  
Programming Assignment 1  
Handwritten Digits Classification

Project Report  
Group Number 50  
March, 2017

*Team*

Muhammed Zaki Muhammed Husain Bakshi  
Nikhil Prakash  
Sunil Kunjappan Vasu

## Overview

In this assignment, our objective is to implement a Multilayer Perceptron Neural Network and evaluate its performance in classifying handwritten digits. We have also used Tensor Flow library to deploy deep neural networks and understand how having multiple hidden layers can improve the performance of the neural network. Multiple observation on accuracies of neural network are studied by changing the parameters.

## Hyper-Parameter for Neural Network

### Relation between $\lambda$ and Performance of Neural Network

We use regularization in Neural Network to avoid overfitting problem. We have analyzed the system performance by changing different value of  $\lambda$  to see its effect in predicting accuracy in validation set. We have varied  $\lambda$  from 0 (no regularization) to 60 in increments of 5 or 10 and the below results are obtained.

The observation is that the accuracy falls from  $\lambda$  value 0 to a value near 20. Again from there as the  $\lambda$  is increased further we can see that the accuracy is falling again. The below following reading are taken for lambda value ranging from 0 to 60 and their corresponding accuracy are also noted.

### Observations:

Lambda Value: 0

Training set Accuracy:95.116%

Validation set Accuracy:94.36%

Test set Accuracy:94.68%

Lambda Value: 10

Training set Accuracy:94.726%

Validation set Accuracy:94.07%

Test set Accuracy:94.45%

Lambda Value: 20

Training set Accuracy:95.212%

Validation set Accuracy:93.97%

Test set Accuracy:94.88%

Lambda Value: 30

Training set Accuracy:94.884%

Validation set Accuracy:94.41%

Test set Accuracy:94.71%

Lambda Value: 40

Training set Accuracy:94.498%

Validation set Accuracy:93.97%

Test set Accuracy:94.17%

Lambda Value: 50

Training set Accuracy:93.946%

Validation set Accuracy:93.09%

Test set Accuracy:93.84%

Lambda Value: 60

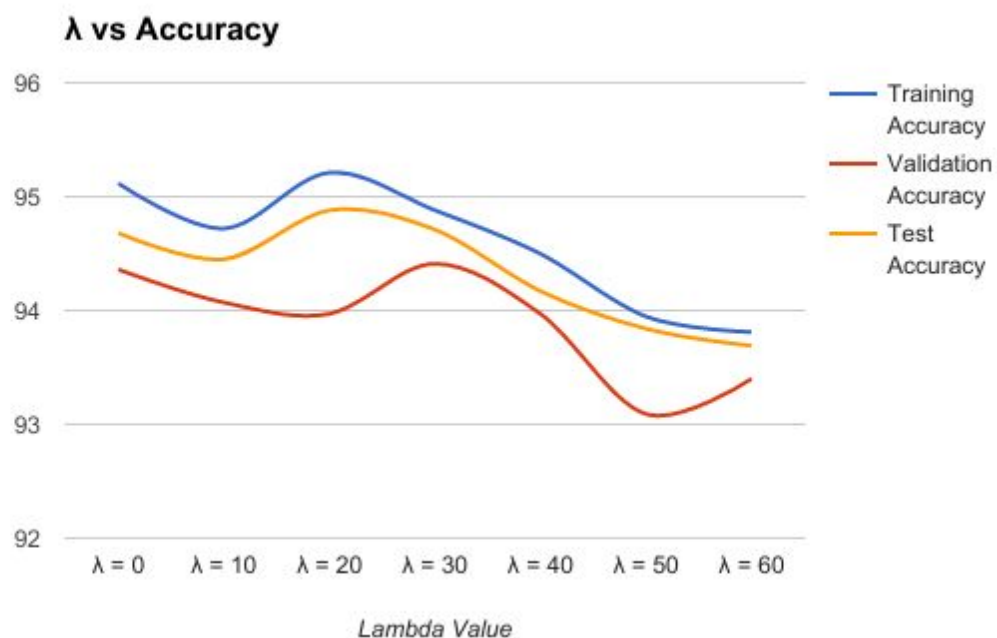
Training set Accuracy:93.81%

Validation set Accuracy:93.4%

Test set Accuracy:93.69%

The overfitting and underfitting of the system can result in lower performance of the system and this can be solved with regularization. The lower accuracy around lambda value 0 is due to the fact that the lambda value does not control the magnitude of parameters during calculation of error. And the lower accuracy at larger lambda value is due to too higher contribution to error function thus negatively impacting calculation. From our observation we conclude that the lambda value need to be chosen in between both of these extreme cases.

The final plot based on all the above reading is plotted below. We could conclude that lambda needs to be chosen as a value from the middle of the graph, i.e the value that corresponds to the maximum accuracy would be located near the center of the graph.



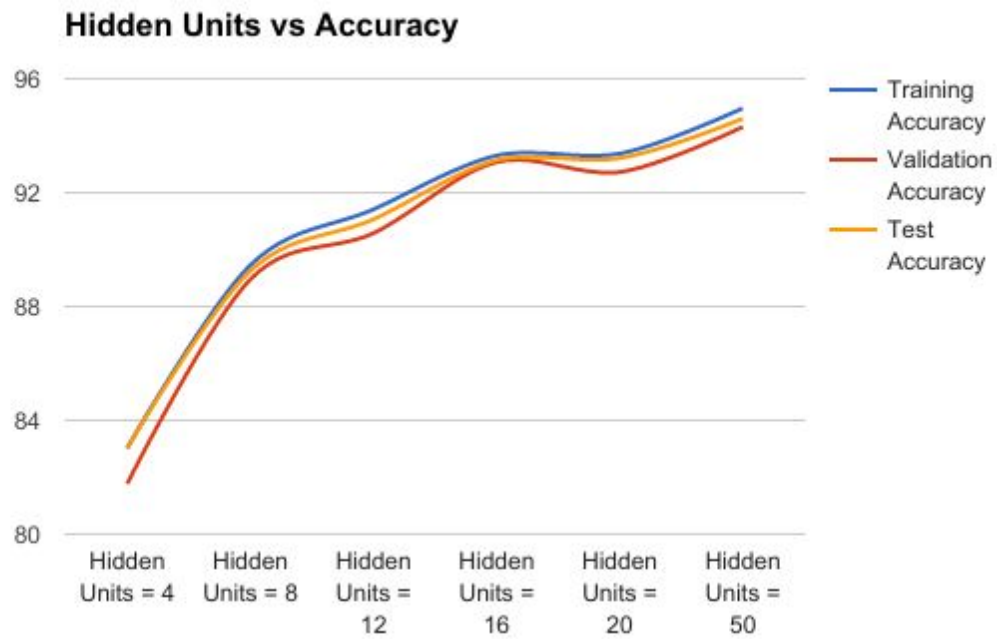
## Effect of Hidden units on Performance of Neural Network

We have varied the number of hidden units from small number of units and gradually increase the size. We have tried for hidden units ranging from 4, 8, 12, 16, 20, 50. We have fixed the lambda value as 10 for these observation. The following observations were obtained for this range of hidden units.

### Observations:

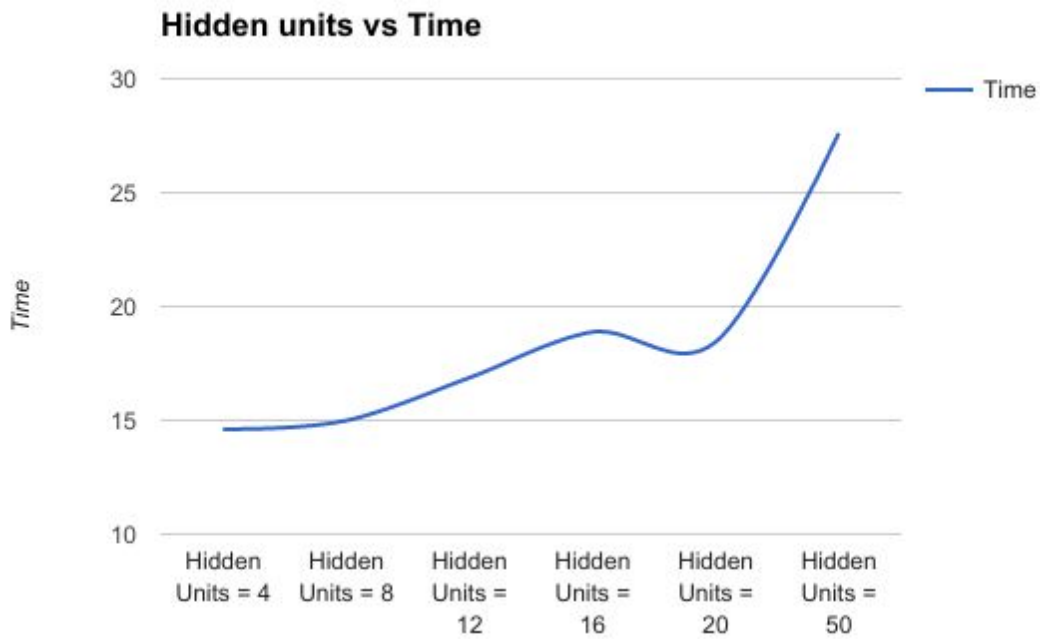
Hidden Units: 4	Hidden Units: 8
Training set Accuracy:83.022%	Training set Accuracy:89.456%
Validation set Accuracy:81.77%	Validation set Accuracy:88.95%
Test set Accuracy:83.02%	Test set Accuracy:89.25%
Rum time: 14.610327959060669	Rum time: 14.999963760375977
Hidden Units: 12	Hidden Units: 16
Training set Accuracy:91.41%	Training set Accuracy:93.304%
Validation set Accuracy:90.57%	Validation set Accuracy:93.07%
Test set Accuracy:91.06%	Test set Accuracy:93.14%
Rum time: 16.877825021743774	Rum time: 18.889121770858765
Hidden Units: 20	Hidden Units: 50
Training set Accuracy:93.368%	Training set Accuracy:94.964%
Validation set Accuracy:92.72%	Validation set Accuracy:94.31%
Test set Accuracy:93.21%	Test set Accuracy:94.6%
Rum time: 18.449976921081543	Rum time: 27.619040727615356

We have plotted the number of hidden units vs accuracy and hidden units vs run time and we have obtained the below results.



We could see that as the number of hidden units increase the accuracy of the model increases. This can be due to the reason that as there are more units the system is able to find more minute error and adjust the error hence resulting in more accuracy.

When the Hidden Units are increased we could see that the run time increases. The run time is proportional to the number of hidden units. This is due to the high computing and complexity involved in the hidden units. Hence we can conclude that the number of hidden units needs to be chosen in such a way that there is balance between runtime and accuracy.



## Accuracy of classification method on the CelebA data set

We have run the tests for lambda value of 20, and hidden Units of 256 units and the following accuracies were obtained. We have taken lambda value of 20 as this related to the best accuracy in our test for Handwritten digits dataset. The test ran for a runtime of 7 minutes. The large run time is due to the total number of hidden units of 256. But the large number of hidden units has ensured the high accuracy even though the run time has been compromised to a limited extend. The below observation are obtained in our test.

Training set Accuracy:84.1279620853%

Validation set Accuracy:83.1144465291%

Test set Accuracy:84.7085541257%

## Comparison of your Neural Network with a Deep Neural Network

The Accuracy of the observation in Deep Neural Network increases initially but we could see that after a specific threshold of the number of hidden layers the accuracy drops.

### Observations for Tensor Flow Test:

Hidden Layer: 2

```
Optimization Finished!  
Accuracy: 0.813777
```

Hidden Layer: 3

```
Optimization Finished!  
Accuracy: 0.787661
```

Hidden Layer: 5

```
Optimization Finished!  
Number of hidden layers: 5  
Accuracy: 0.761544
```

The accuracies for the above observations are plotted as follows.

