

CSE 574 - Programming Assignment 2

Handwritten Digits Classification

Submitted by

1. Sajid Khan - 50248743
2. Vinay Vardhaman - 50248852.

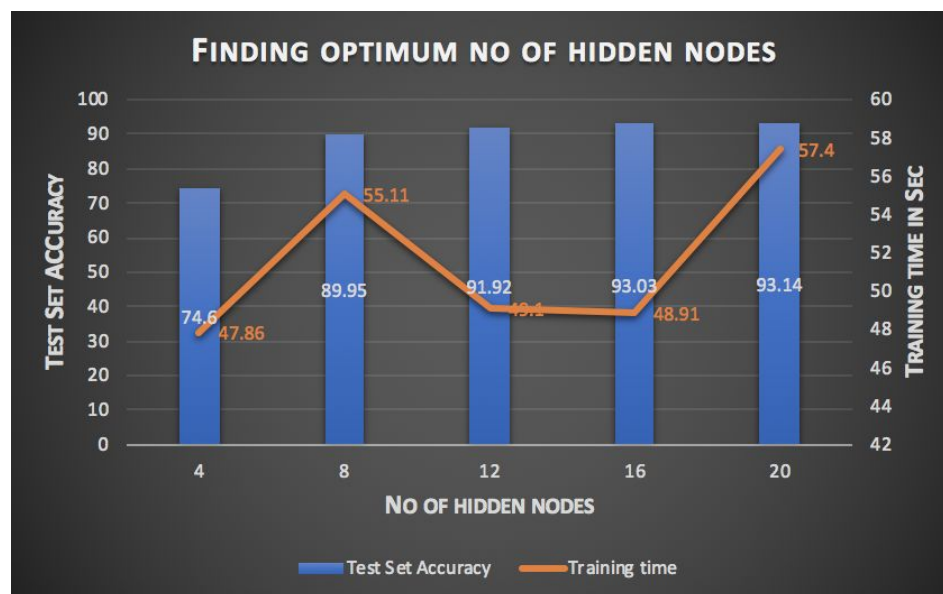
1. Finding optimal hyper parameters for Neural network

Optimum number of hidden units

The following results were obtained with lambda value set to 20

Hidden Units	Test Accuracy(in percent)	Training time(in sec)
4	74.6	47.86
8	89.95	55.11
12	91.92	49.1
16	93.03	48.91
20	93.14	57.4

As shown in the table, it can be observed that for a fixed lambda value, the test accuracy increases almost linearly with the number of hidden nodes. Training time also increases with increased complexity with more number of hidden units.



The plot above shows the test accuracy and training time plotted against the number of hidden nodes.

Observation:

It can be observed that test accuracy and training time almost increase linearly with number of hidden units. Highest accuracy of 93.14 is obtained for 20 hidden units.

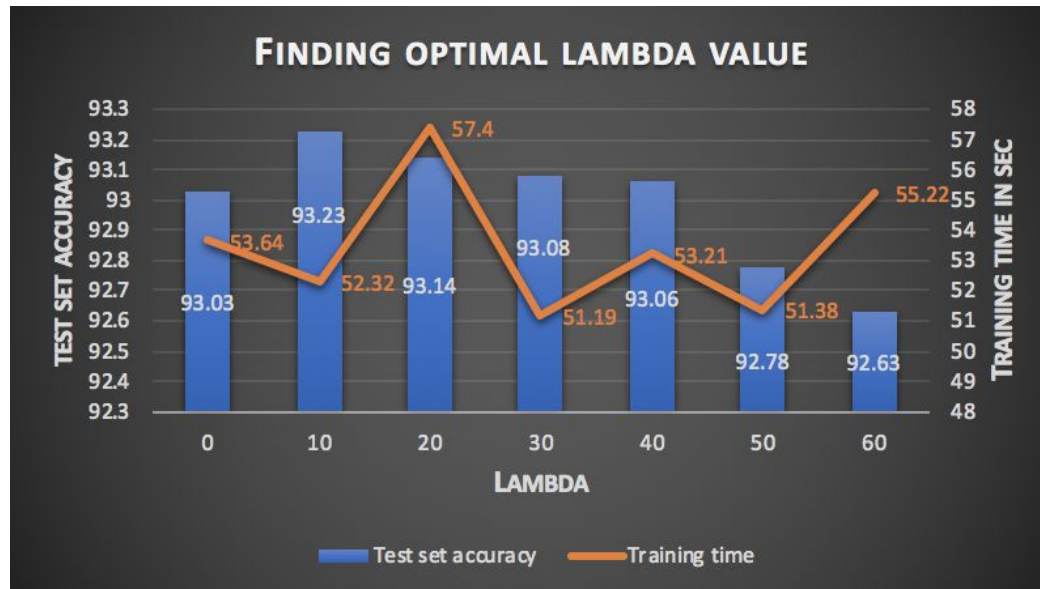
So the optimum value of hidden units is 20.

Optimum Regularization Parameter - lambda

Now we will change the lambda value with fixed number of hidden units(20) and observe the test set results.

Lambda	Test Set Accuracy(in percent)	Training time(in sec)
0	93.03	53.64
10	93.23	52.32
20	93.14	57.4
30	93.08	51.19
40	93.06	53.21
50	92.78	51.38
60	92.63	55.22

The table above shows Test set accuracy and Training time plotted with increasing lambda values. It can be observed that highest accuracy is obtained for lambda value 10. Increasing lambda value more doesn't help in regularization.



The plot above shows test set accuracy and training time plotted against different values of lambda. The highest accuracy can be observed at lambda value 10.

Observation:

From the table and plot, it can be observed highest test accuracy of 93.23 is observed for lambda value 10.

So the optimum value of lambda is 10.

2. Accuracy of classification method on handwritten digits test data

From the results above, we can conclude that highest accuracies are observed with more number of hidden units and optimal value of lambda i.e. 10

The highest accuracy observed on handwritten digits data is 93.23

3. Accuracy of classification method on CelebA data set

The optimum lambda and number of hidden units obtained in the previous exercise are used to classify the celebA data set.

For lambda value 10 and number of hidden nodes 20

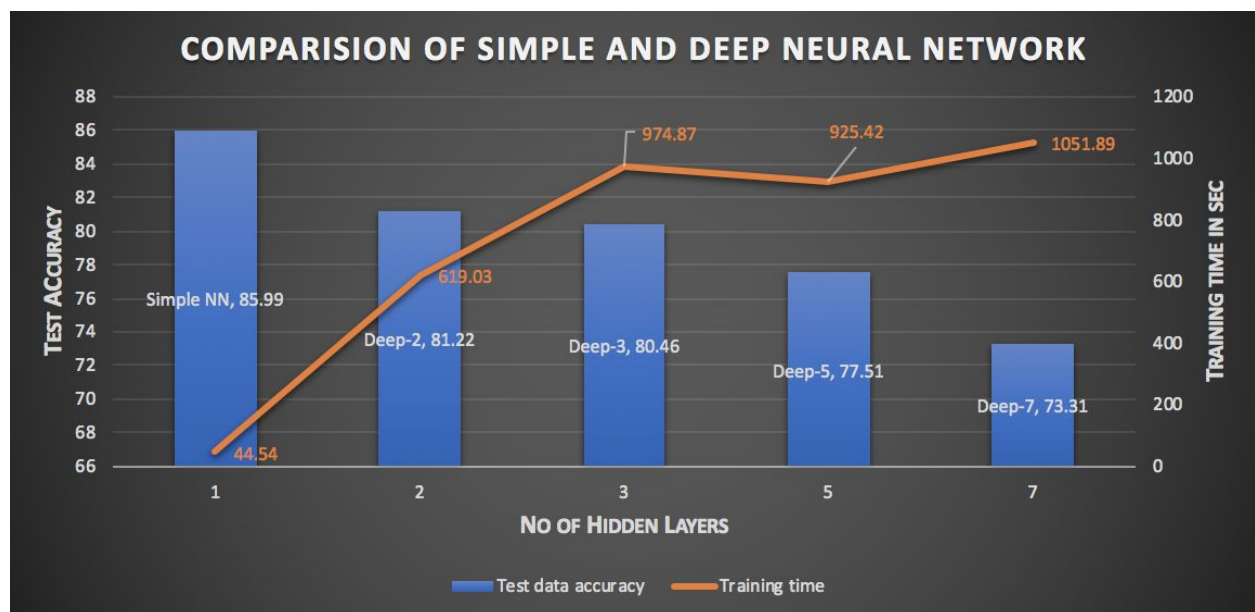
Testing Data Accuracy is 85.99% and Training time is 44.54 sec.

4. Comparison of Simple Neural Network with Deep Neural Network

Now, we will try to compare results from Simple Neural Network with Deep Neural Network.

Type	No of Hidden layers	Test Data Accuracy	Training Time (in sec)
Simple NN	1	85.99	44.54
Deep NN	2	81.22	619.03
Deep NN	3	80.46	974.87
Deep NN	5	77.51	925.42
Deep NN	7	73.31	1051.89

The table above shows Test data accuracy and Training time plotted for simple neural network and deep network with different number of hidden layers.



The plot above shows the Test accuracy and Training time for different number of hidden layers. It can be observed that **highest test accuracy is obtained for simple neural network with just one layer**. As the number of layers increase, accuracy decreases due to over-fitting.

Also, **training time is the lowest with Simple neural network**. With deep networks, as the number of layers increase, training time increases linearly due to increased cost and computation.

Observations:

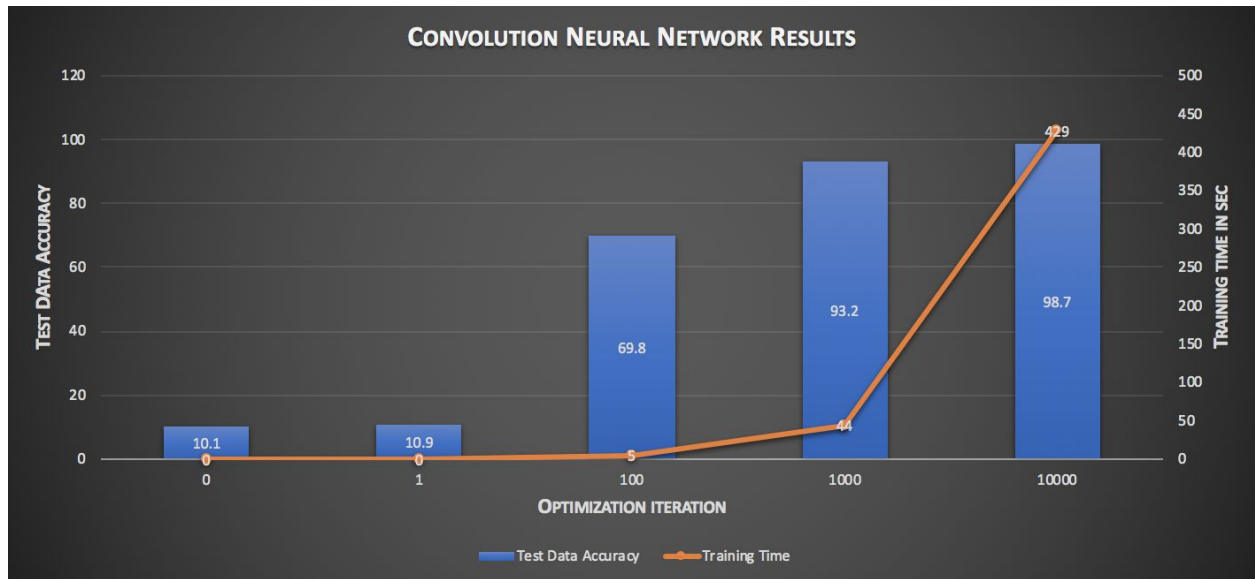
1. By increasing the no of hidden layers, training time increases due to increase in computation and complexity.
2. Adding more number of hidden layers causes overfitting, thus resulting in lower accuracies.
3. Simple neural networks does the work with better accuracy and less training time than Deep neural networks with multiple layers.

5. Convolutional Neural Network Results

Optimization Iteration	Test Set Accuracy (in percent)	Training time (in sec)
0	10.1	0
1	10.9	0
100	69.8	5
1000	93.2	44
10000	98.7	429

The table above shows results of Test set accuracy and Training time for respective number of iterations. As the number of iterations increase, the accuracy of the predictions increases. So does the training time.

The graph in the next page plots the Test data accuracy(in percent) and Training time(in sec) against number of optimization iterations.



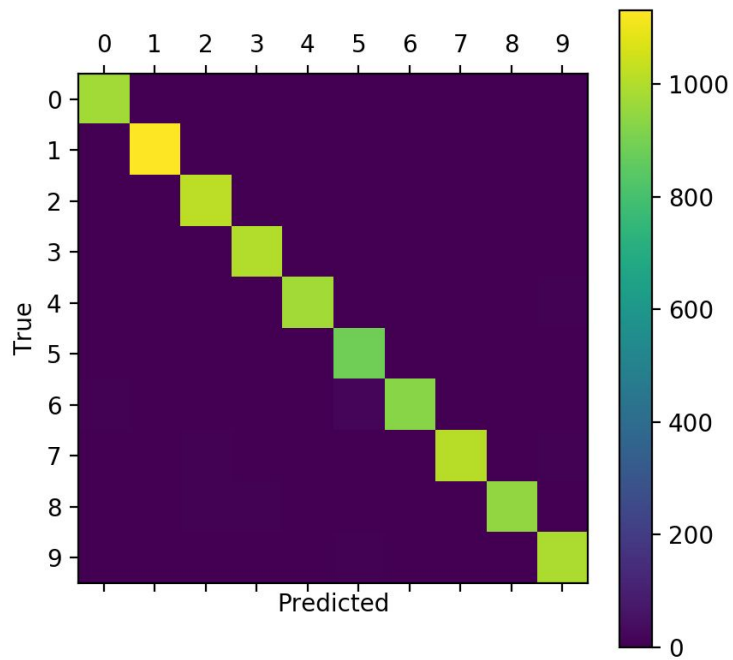
As we can observe, the Test data accuracy and Training time increase exponentially with the number of iterations.

Confusion matrix:

Confusion matrix obtained after 10,000 iterations

```
[[ 975  0  1  0  0  1  0  0  2  1]
 [ 0 1132  1  0  0  0  0  2  0  0]
 [  1  4 1019  2  1  0  0  2  3  0]
 [  1  0  1 1002  0  3  0  1  2  0]
 [  0  0  1  0 974  0  0  0  1  6]
 [  2  0  0  4  0 885  1  0  0  0]
 [  8  3  0  1  2 15 929  0  0  0]
 [  0  3  6  2  0  0  0 1011  1  5]
 [  4  0  6  6  1  2  0  2 949  4]
 [  1  3  0  2  3  6  0  2  0 992]]
```

Matrix plot:



The figure above shows matlab representation of Confusion matrix

Observation:

Highest accuracy of 98.7% is obtained for over 10000 iterations with training time of 429 sec.