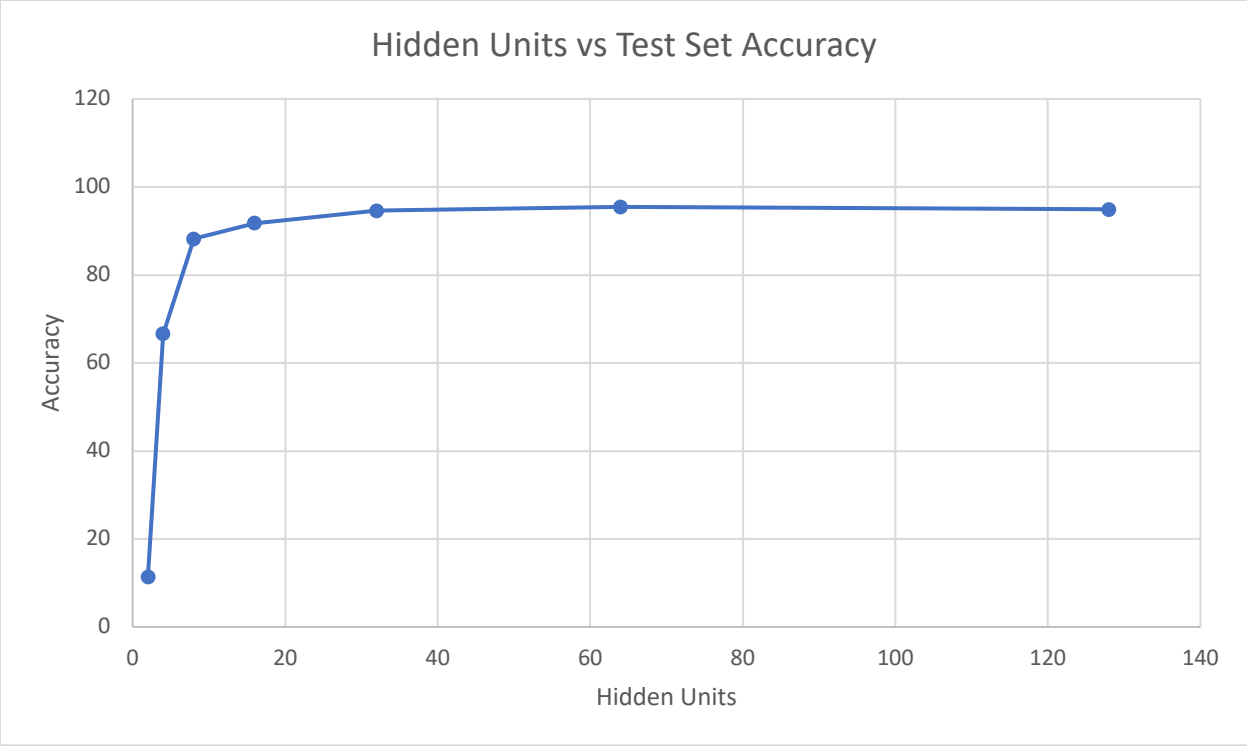
**CSE 574 Introduction to Machine Learning**
**Programming Assignment 2**
**Roshan Saundankar: 50419577**
**Sumeet Sahu: 50367891**
**Mrunmayee Rane: 50417094**

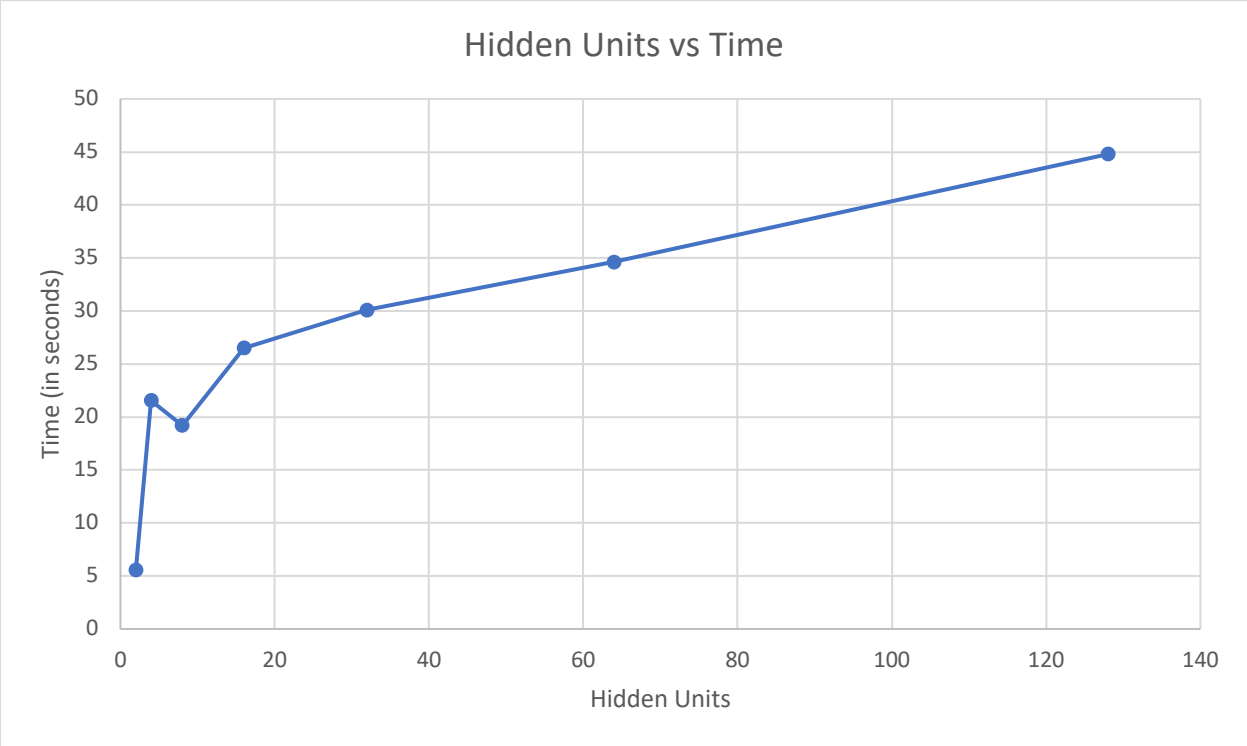
# HANDWRITTEN DIGITS CLASSIFICATION

## Part1: Selecting Hyper Parameters using nnScript:

In this part we will compare the accuracy and time taken for Neural Networks based on different values of hidden units and constant lambda.

| Hidden Units | Lambda | Training set Accuracy | Validation set Accuracy | Test Set Accuracy | Time taken in seconds |
|---|---|---|---|---|---|
| 2 | 0 | 11.272 | 11.06 | 11.35 | 5.59 |
| 4 | 0 | 66.884 | 66.64 | 66.58 | 21.56 |
| 8 | 0 | 88.354 | 87.42 | 88.19 | 19.2 |
| 16 | 0 | 92.068 | 90.84 | 91.72 | 26.5 |
| 32 | 0 | 95.056 | 94.1 | 94.57 | 30.09 |
| 64 | 0 | 95.762 | 94.57 | 95.46 | 34.62 |
| 128 | 0 | 95.148 | 94.740 | 94.89 | 44.79 |


Hidden Units vs Test Set Accuracy
Accuracy
Hidden Units
0
20
40
60
80
100
120
140


Hidden Units vs Time
Time (in seconds)
Hidden Units
0
5
10
15
20
25
30
35
40
45
50
0
20
40
60
80
100
120
140

**Observations:**

- As the number of hidden units increases the accuracy increases. The objective function also converges fastly as the number of hidden units increases.
- The best Train, Validation, Test accuracy can be seen at 64 units.
- The time also starts increasing as we add more hidden units.

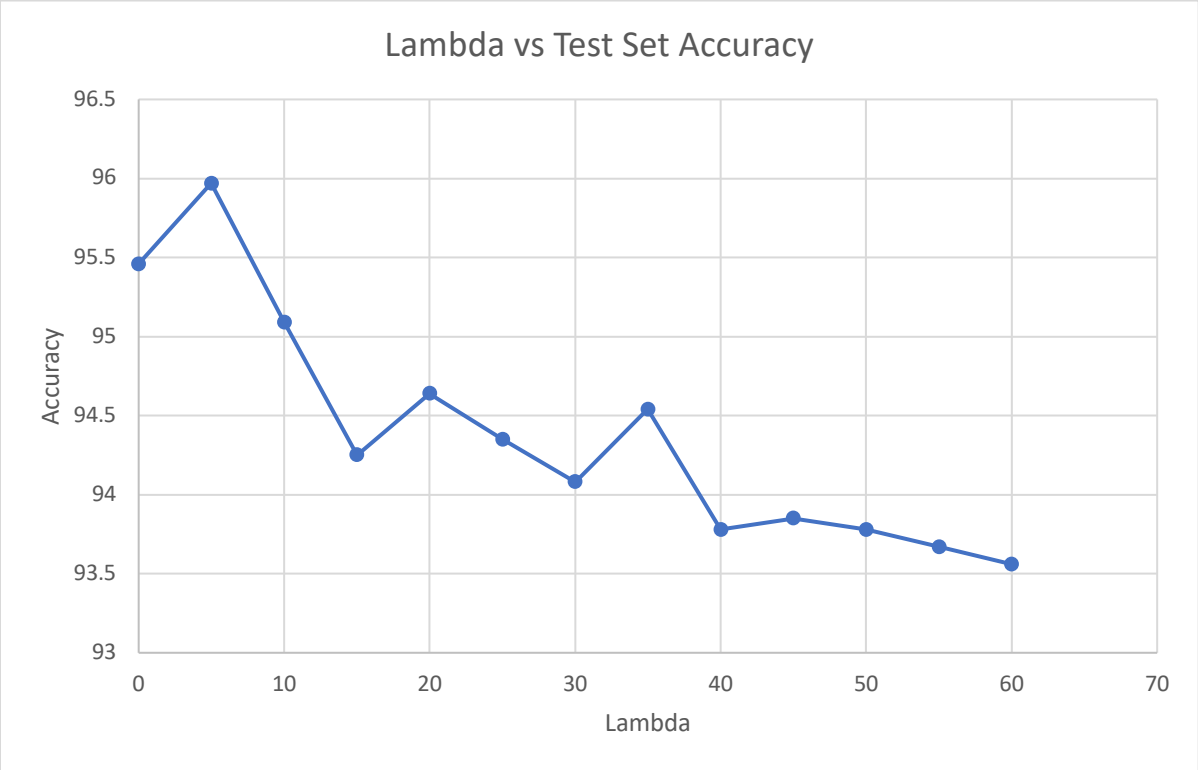
**Conclusions:**

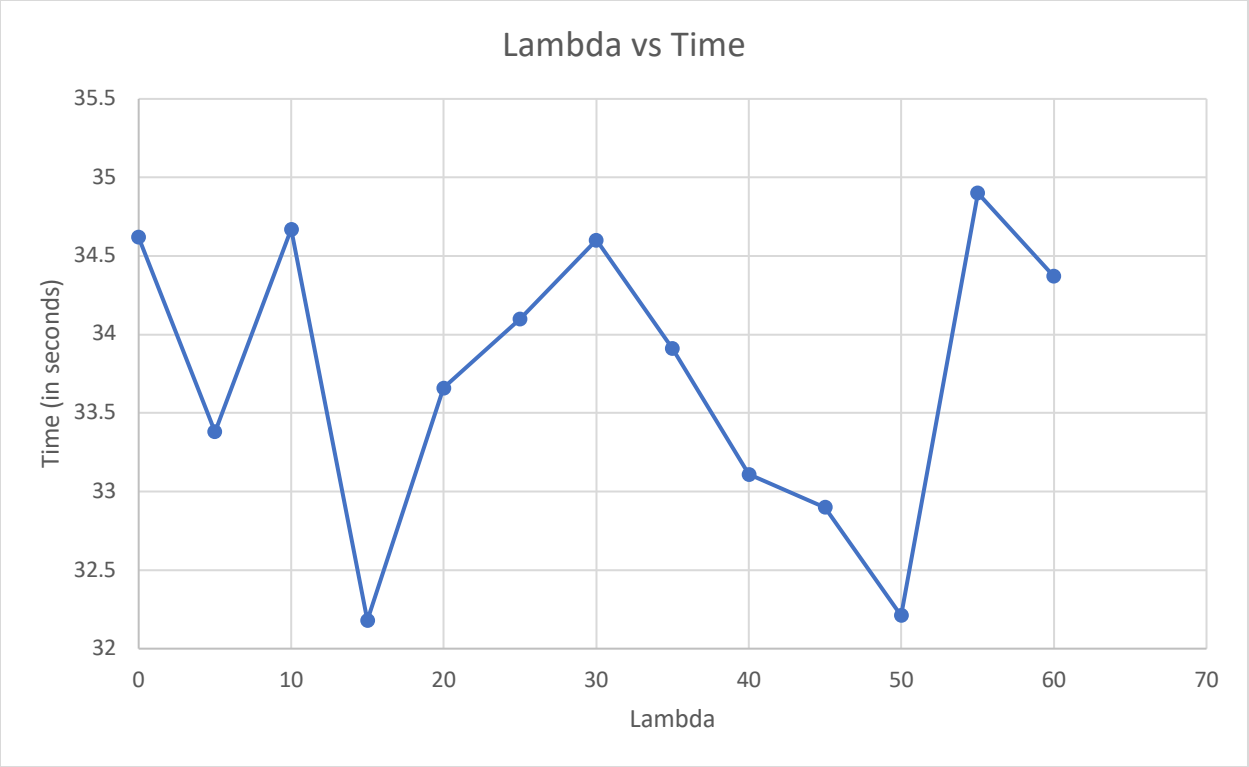
- It can be concluded that the neural network performs its best when there are 64 hidden units.

Now we compare the accuracy and time taken for Neural Networks based on different values of lambda and constant best value of hidden units we got previously.

| **Lambda** | **Hidden Units** | **Training set Accuracy** | **Validation Set Accuracy** | **Test Set Accuracy** | **Time taken in seconds** |
|---|---|---|---|---|---|
| 0 | 64 | 95.762 | 94.57 | 95.46 | 34.62 |
| 5 | 64 | 95.282 | 94.38 | 95.97 | 33.38 |
| 10 | 64 | 95.304 | 94.72 | 95.09 | 34.67 |
| 15 | 64 | 94.428 | 93.86 | 94.25 | 32.18 |
| 20 | 64 | 94.64 | 94.320 | 94.64 | 33.66 |
| 25 | 64 | 94.53 | 94.14 | 94.35 | 34.1 |
| 30 | 64 | 94.186 | 93.65 | 94.08 | 34.6 |
| 35 | 64 | 94.582 | 94.07 | 94.54 | 33.91 |
| 40 | 64 | 93.746 | 93.71 | 93.78 | 33.11 |

| | | | | | |
|---|---|---|---|---|---|
| 45 | 64 | 94.114 | 93.67 | 93.85 | 32.9 |
| 50 | 64 | 93.69 | 93.25 | 93.78 | 32.21 |
| 55 | 64 | 93.804 | 93.17 | 93.67 | 34.9 |
| 60 | 64 | 93.53 | 92.97 | 93.56 | 34.37 |






**Conclusions:**

- The optimal value obtained of lambda obtained is 5 according to our observation.
- The optimal value for hidden units is 64 and lambda is 5.
- If the lambda value is too large underfitting will happen.

**Part 2: Comparison between Deep Neural Network and Neural Network.**

In this part the accuracy of a single layer hidden neural network on CelebA using facenn script.

**Observations:**

- For default values, Lambda 10 and hidden units 256.

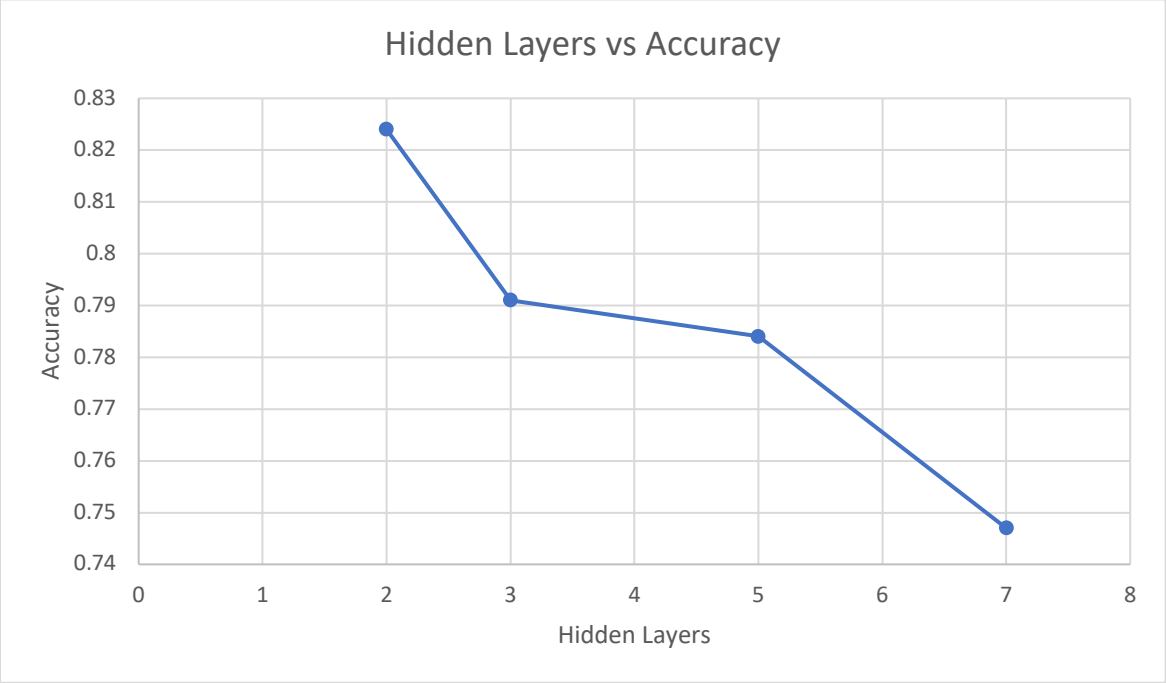
| Lambda | Hidden Units | Training set Accuracy | Validation set Accuracy | Test Set Accuracy | Time Taken in Seconds |
|---|---|---|---|---|---|
| 10 | 256 | 85.12 | 84.16 | 85.69 | 190.23 |

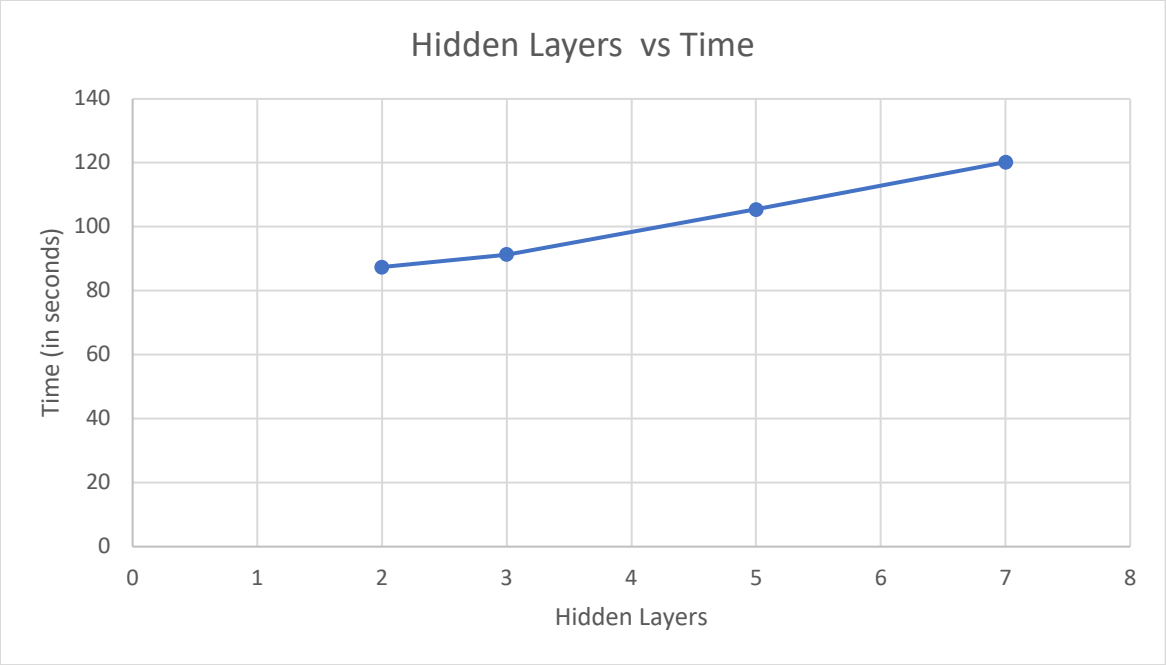
- For optimal values obtained previously:

| Lambda | Hidden Units | Training Set Accuracy | Validation Set Accuracy | Test Set Accuracy | Time Taken in Seconds |
|---|---|---|---|---|---|
| 5 | 64 | 95.282 | 94.38 | 94.99 | 33.38 |

Now we will evaluate the accuracy of deep nn on CelebA dataset:

| Hidden Layers | Accuracy | Time in Seconds |
|---|---|---|
| 2 | 0.824 | 87.43 |
| 3 | 0.791 | 91.34 |
| 5 | 0.784 | 105.42 |
| 7 | 0.747 | 120.23 |






**Observations:**

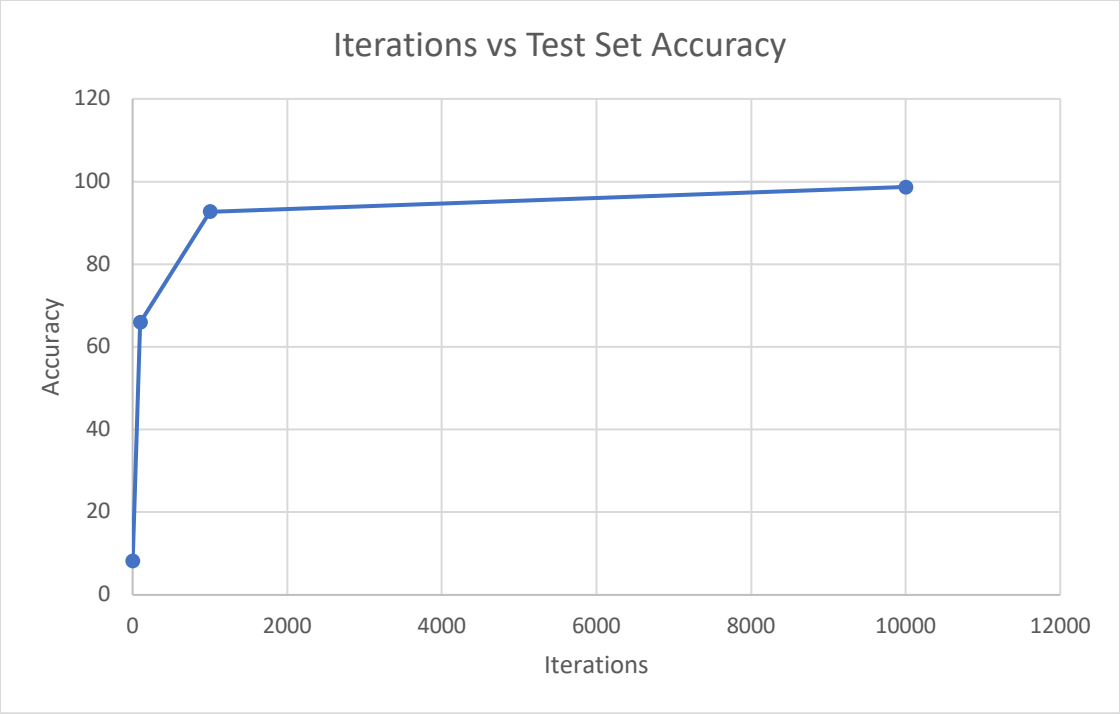
- It can be seen that as we increase the number of layers the accuracy decreases with it.
- The Time taken also increases with increase in the number of layers.

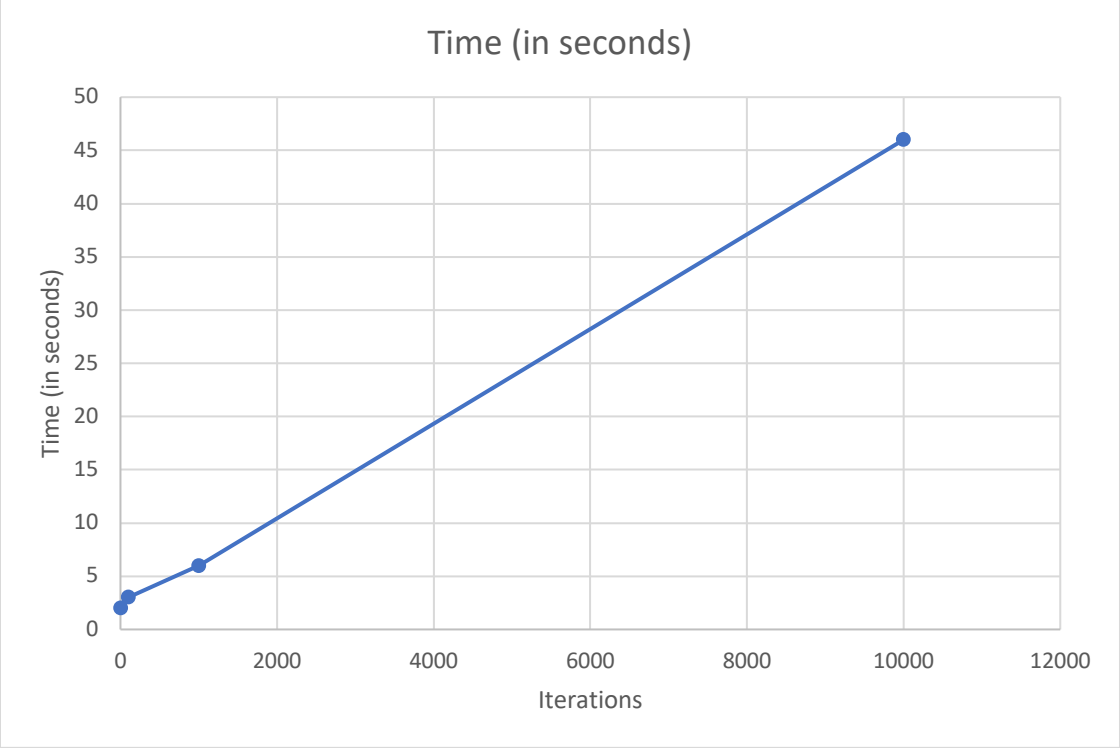
**Comparing the performance of single layer vs deep neural networks in terms of accuracy for test set**

- **Accuracy:** It can be concluded from above that the neural network works much better with one hidden layer with accuracy of 94.99 on celeb dataset than with multiple layers with accuracy 82.4 for two hidden layers.
- **Time:** Time required is also less in single layer than deep neural network.

**Part 3: Convolution Neural Network:**

| **Iterations** | **Test Set Accuracy** | **Time (in seconds)** |
|---|---|---|
| 1 | 8.2 | 2 |
| 100 | 65.9 | 3 |
| 1000 | 92.7 | 6 |
| 10000 | 98.7 | 46 |




**Observations:**

- It can be seen from the above table that as we increase the number of Iterations of the convolutional neural network the accuracy increases with it.

**Output:**

Confusion Matrix obtained after 10000 iterations.

```
[[ 974    0    1    0    0    0    1    0    3    1]
 [   0 1129    3    0    0    0    1    1    1    0]
 [   1    0 1020    1    0    0    0    2    8    0]
 [   0    0    2 1000    0    5    0    1    2    0]
 [   0    0    2    0  977    0    1    1    0    1]
 [   1    0    1    7    0  880    1    1    1    0]
 [   4    2    0    1    1    6  944    0    0    0]
 [   2    2    7    4    0    0    0 1004    2    7]
 [   3    0    3    2    0    1    0    2  962    1]
 [   4    4    1    3    9    3    0    3    2  980]]
```