DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY AT BUFFALO

CSE 574 Programming Assignment 2
Handwritten Digits Classification

Group 141

Shivam Sahu
(#50247673)

Parush Garg
(#50248921)

# PART 1 - How to choose the hyper-parameter for Neural Network:

Task 1 : By keeping the Lambda Constant, finding the optimal number of Hidden units

*Lambda - 0*
Hidden Units - 4,8,12,16,20,40,50,60

ACCURACY VARIANCE

Hidden Unit	Training Set	Validation Set	Test Set
4	65.012	63.42	64.56
8	89.72	88.75	89.57
12	91.308	91.17	91.21
16	92.054	91.71	91.67
20	92.828	92.44	92.18
40	94.906	94.64	94.69
50	95.238	94.99	94.8
60	95.022	94.90	94.7
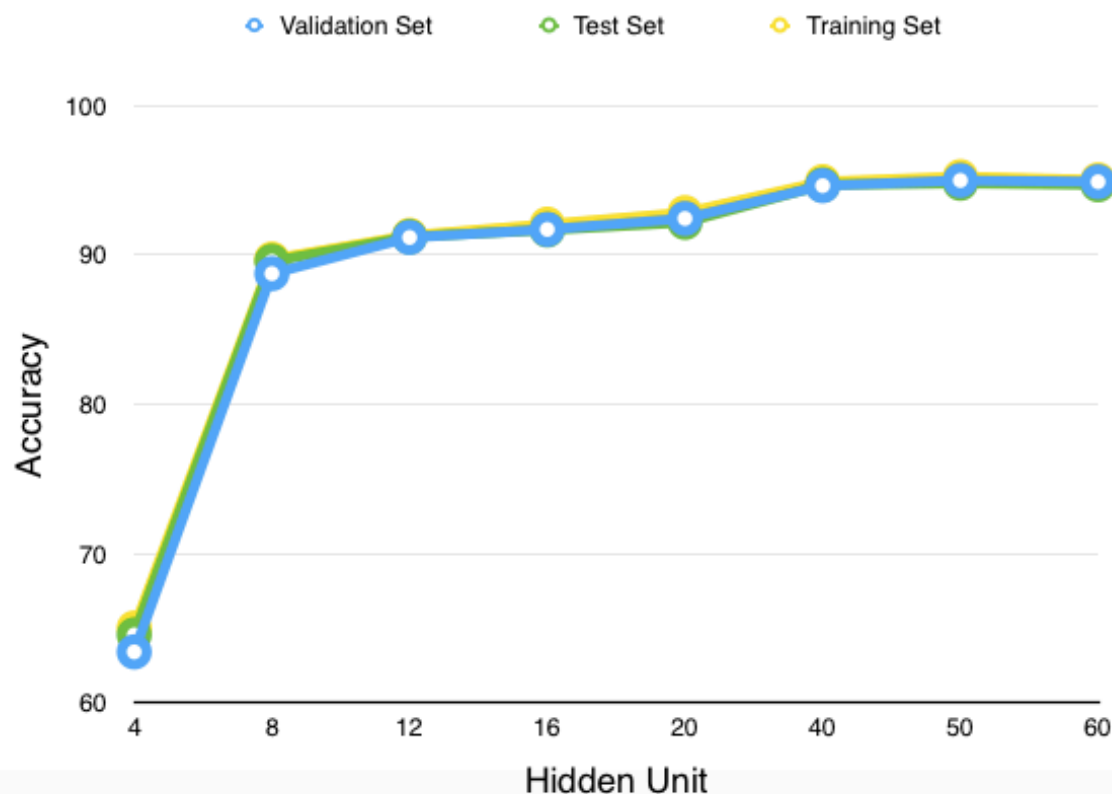
TIME VARIANCE

Hidden Unit	Time
4	59.0010333
8	62.1118622
12	58.9744882
16	65.56062552
20	75.41019427
40	98.82626572
50	94.43389053
60	90.02652682
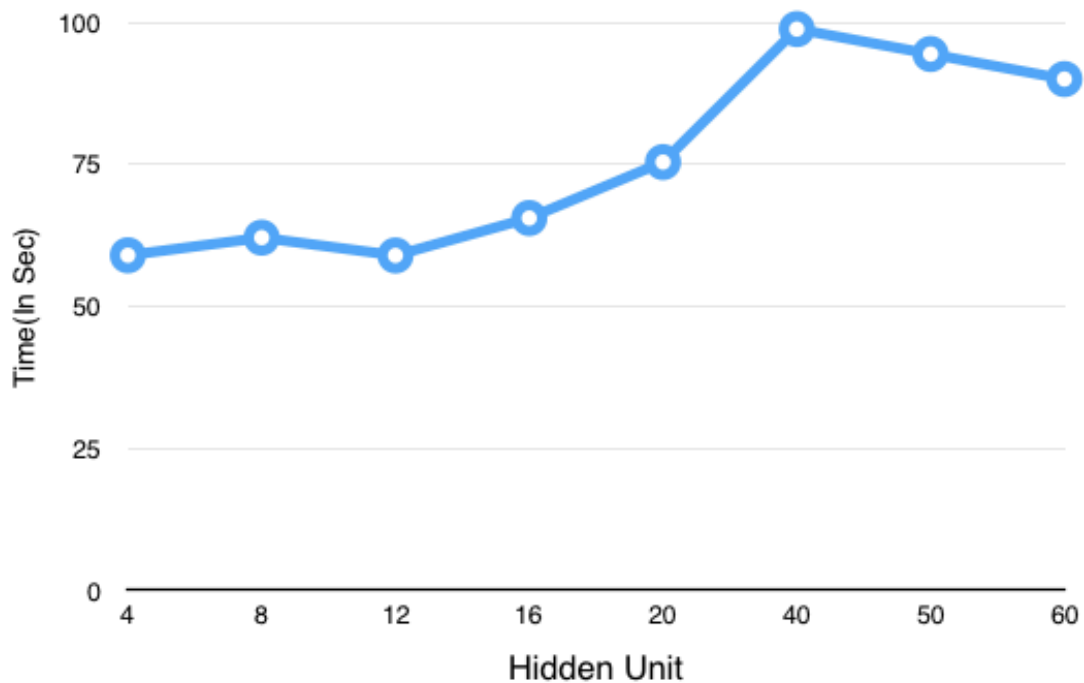
Conclusion-

By increasing the number of hidden units, Accuracy first increase and decrease start at value of 50. Time first increase up to certain hidden unit and then decrease. But we are more concern about accuracy so we are taking hidden unit value 50 as a optimal hidden unit.
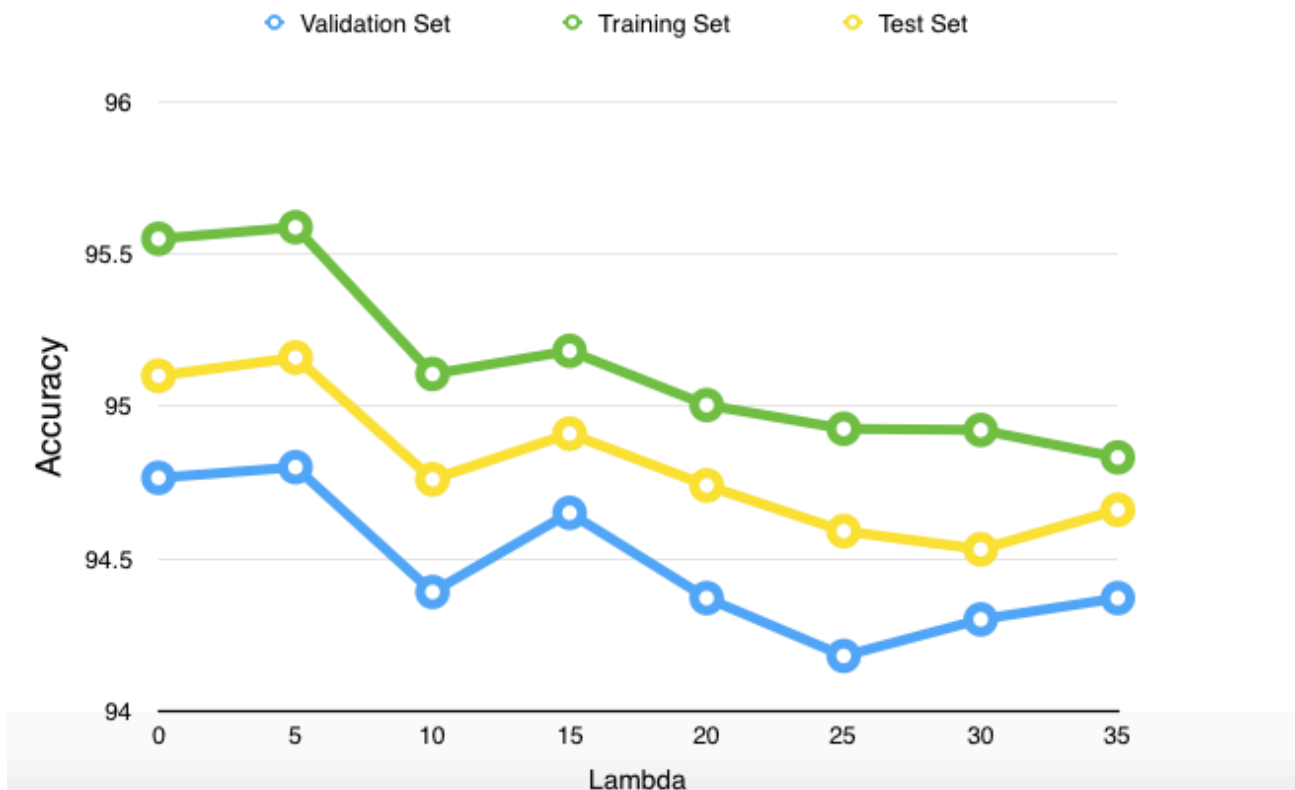
Task 2 : By keeping the Hidden units Constant, finding the optimal value of Lambda

Hidden Unit - 50
Lambda - 0,5,10,20,25,30,35

ACCURACY VARIANCE

Lambda	Training Set	Validation Set	Test Set
0	95.55	94.765	95.1
5	95.588	94.8	95.16
10	95.106	94.39	94.76
15	95.182	94.65	94.91
20	95.004	94.37	94.74
25	94.926	94.18	94.59
30	94.922	94.3	94.53
35	94.832	94.37	94.66

TIME VARIANCE

Lambda	Time
0	90.75906087
5	90.57872089
10	86.22784495
15	86.33816746
20	85.70223242
25	83.63559294
30	86.93623542
35	85.96905843
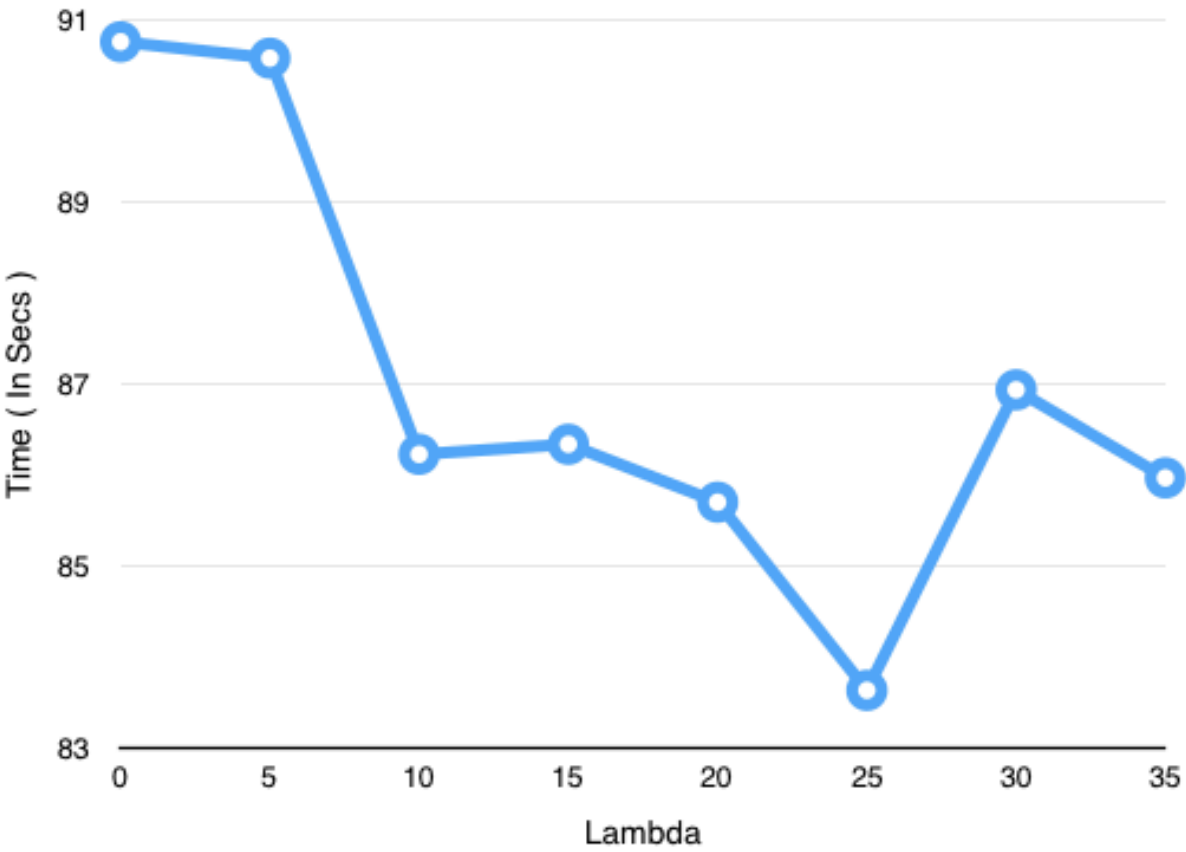
Conclusion-

By increasing the value of Lambda, Accuracy first increase and then decrease start at value 5. Time is comparable for all values of Lambda. But we are more concern about accuracy so we are taking Lambda value 5 as a optimal Lambda.

| Hyper -Parameters	Optimal Value
Hidden Unit	50
Lambda	5
Accuracy	Training Set - 95.588% Validation Set - 94.8% Test Set - 95.16%
Time (In Secs)	90.578

PART 2: Compare the results of neural network with one hidden layer and deep neural network on the CelebA data set
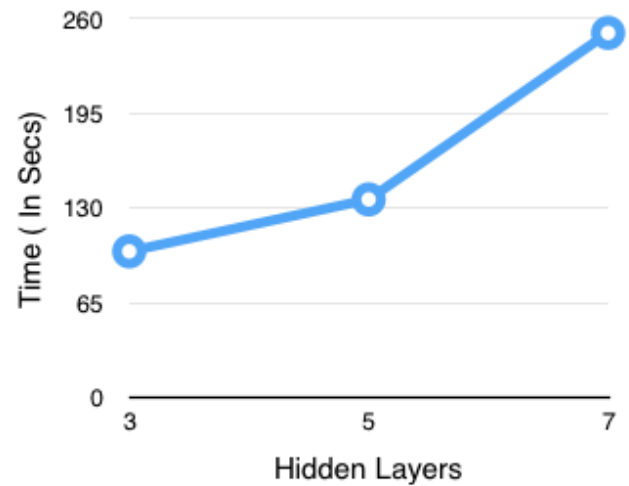
facennScript.py:

Hyper -Parameters	Optimal Value	Default Value
Hidden Unit	50	256
Lambda	5	10
Accuracy	Training Set - 84.9573459716% Validation Set - 83.489681050% Test Set - 85.0492051476%	Training Set - 84.2654028436% Validation Set - 82.288930581% Test Set - 84.3300529902%
Time (In Secs)	114.62280667107552	183.65132400603034

deepnnScript.py:

Layers	Accuracy
3	82.1342
5	82.3429
7	80.5231
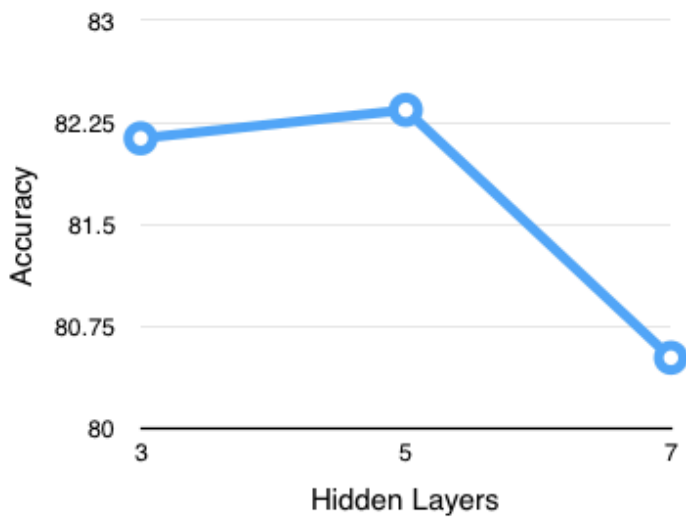
| Layers	Time
3	100.52
5	136.13
7	250.32

Conclusion -

We can see that facennScript.py that is neural network with one hidden layer have accuracy of 84.33 on default value of hidden unit and lambda and have accuracy 85.04 on optimal value of hidden unit and lambda. On the other hand, deepnnScript.py gave 82.13 for 3 layer, 82.34 for 5 layer and 80.5231 for 7 layer. On comparing date, we conclude that the neural network with one hidden layer perform better than multiple layers.

PART 3: Convolution Neural Network

Iteration	Time (In Secs)	Test Accuracy(%)
1	0	10.3
100	8	55.4
1000	78	93.2
10000	740	98.8
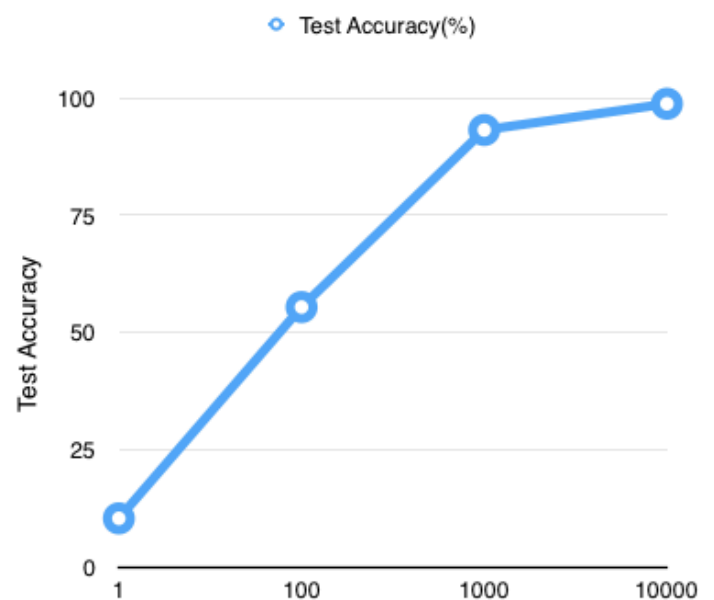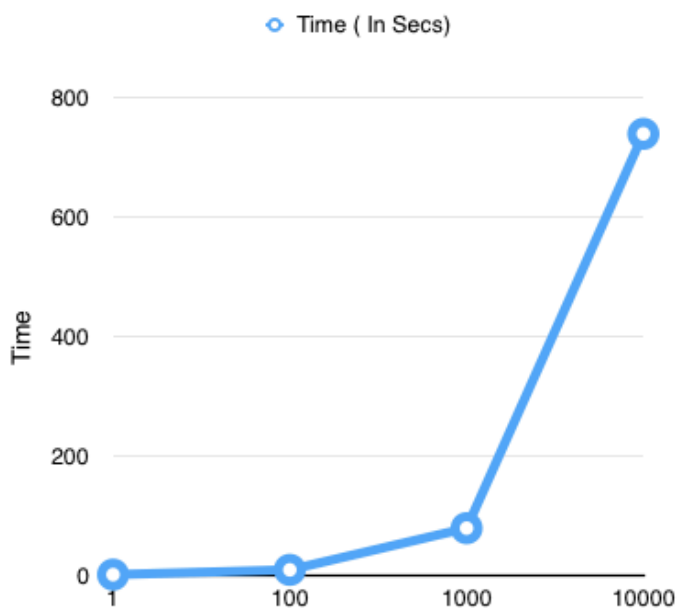
Confusion Matrix after 10000 iteration :

```
Confusion Matrix:
[[ 977    0    0    0    0    0    0    1    2    0]
 [   0 1133    1    0    0    0    0    1    0    0]
 [   2    3 1017    0    1    0    0    4    5    0]
 [   1    0    1 1005    0    1    0    0    2    0]
 [   0    1    0    0  974    0    0    1    2    4]
 [   2    0    0    6    0  880    1    0    1    2]
 [   9    3    1    0    6    8  931    0    0    0]
 [   0    2    3    2    0    0    0 1019    0    2]
 [   6    2    3    4    2    1    0    2  949    5]
 [   3    5    0    3    3    0    0    3    0  992]]
```