Department of Computer Science & Engineering

CSE 574 – MACHINE LEARNING

Handwritten Digits Classification

Programming Assignment – 2

Presented By:

Pratik **Agarwal** [50290570]
Sourav **Bihani** [50290969]
Tannu **Priya** [50290768]

Group - 23

Table of Contents:

1) Problem 1……3

2) Problem 2……8

3) Problem 3……10

Problem 1: Compute and tune Hyper - parameters for the Multilayer Neural Network:

FEEDFORWARD PROPAGATION

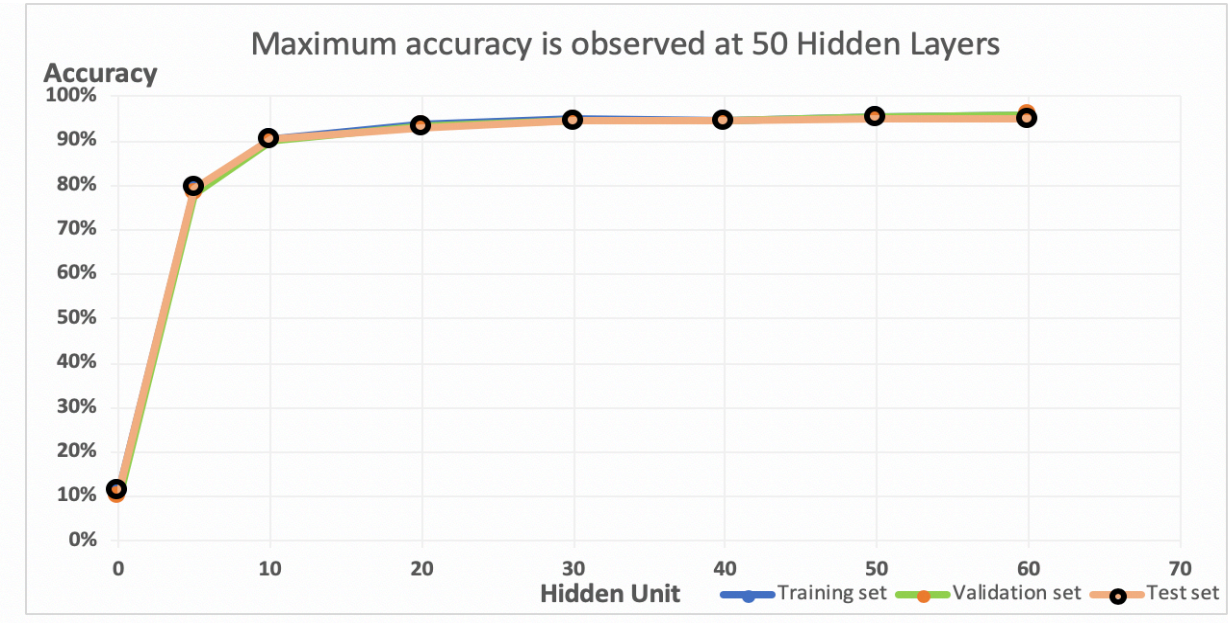
Objective:

- Using MNIST data, train the Multilayer perceptron neural network and evaluate its performance in Handwritten Digit Classification task.
- Compute the number of hidden nodes to achieve optimal Test Set Accuracy overcoming the problem of overfitting and under fitting.

Lambda value is set to default = 0. Following values are obtained:

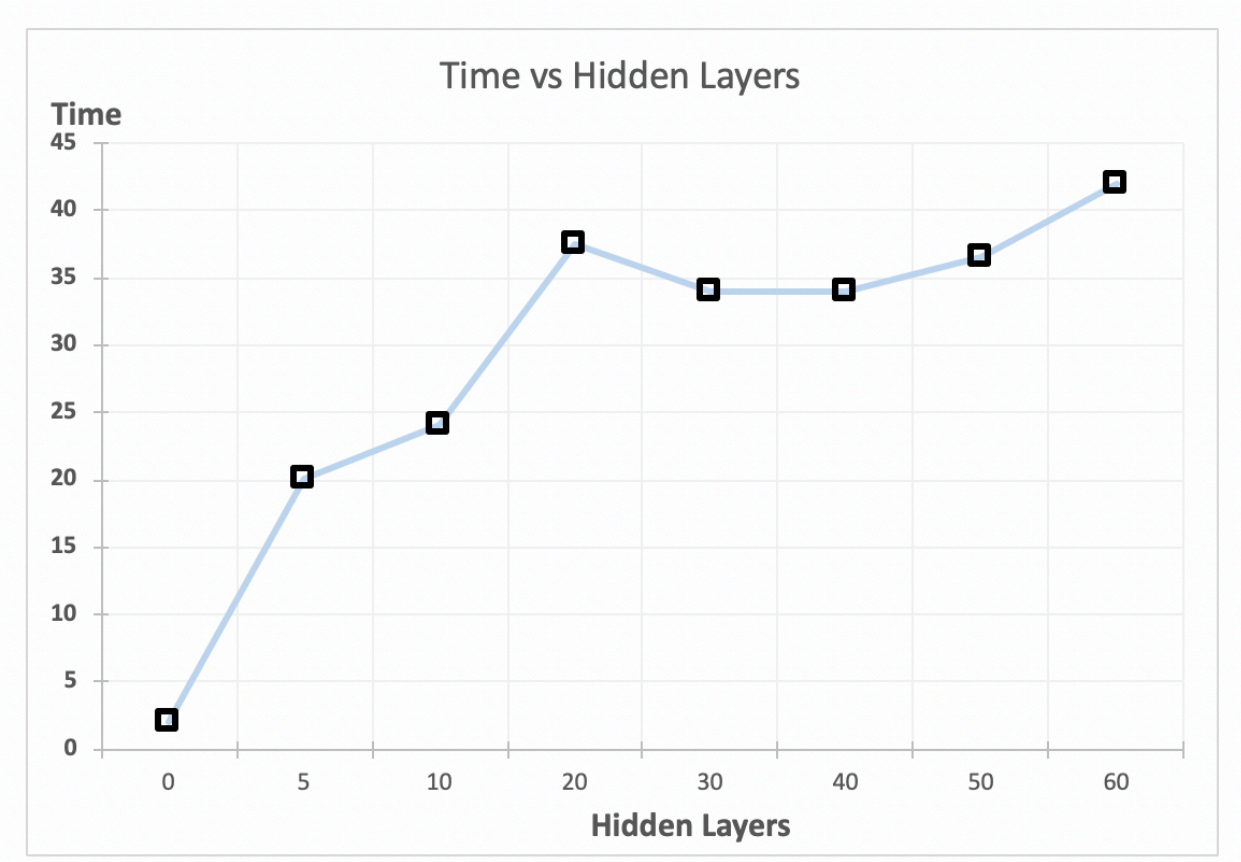
Hidden Unit	Training Set Accuracy	Validation Set Accuracy	Test Set Accuracy
0	11.48%	10.00%	11.35%
5	78.93%	78.11%	79.29%
10	90.48%	90.10%	90.32%
20	93.66%	93.38%	93.13%
30	94.80%	94.75%	94.50%
40	94.76%	94.40%	94.51%
50	95.50%	95.27%	95.13%
60	95.78%	95.88%	94.89%

Observations:



- With 50 hidden layers, we get the best Test Set Accuracy of 95.13%.
- With too many hidden layers we are facing the problem of overfitting and accuracy starts to drop when we increase the number to hidden layers to 60.
- With very few Hidden layers we can observe the problem of under fitting with a low accuracy of 79.29%.

Hidden Unit	Time(In sec)
0	2.0869
5	20.0605
10	24.131
20	37.552
30	34.044
40	34.068
50	36.632
60	41.973



Conclusion:

It is evident form graph that with the increase in number of hidden layers the accuracy of Neural Network increases and the computation time increase as well. But accuracy drops for a neural network with more than 50 nodes but the computation time continues to increase. Hence, we can conclude that with 50 hidden nodes our neural network works best on the MNIST data.

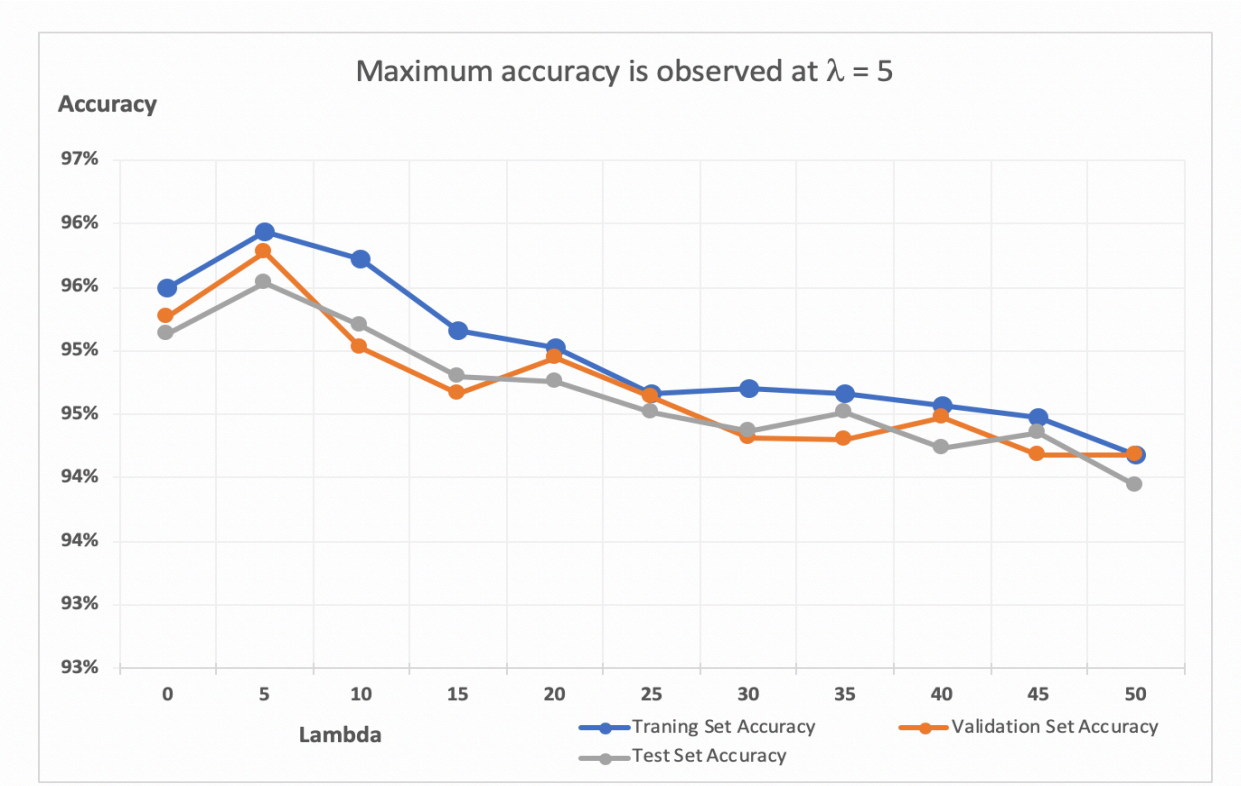
Note: In the above computations, the value of λ was set to 0.

ERROR FUNCTION AND BACKPROPAGATION:

Objective:

- Compute optimal Regularization Coefficient for Error function computation and Backpropagation.

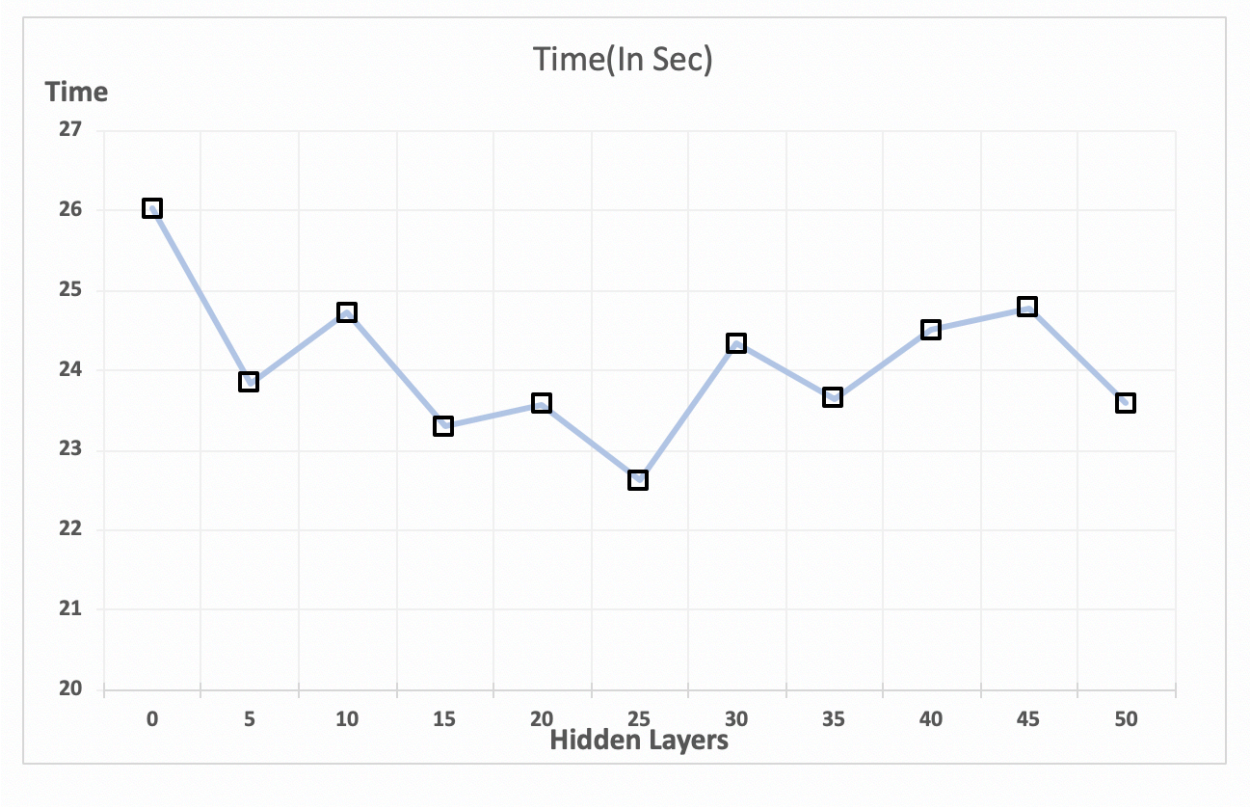
Lambda	Hidden Unit	Training Set Accuracy	Validation Set Accuracy	Test Set Accuracy
0	50	95.50%	95.27%	95.13%
5	50	95.94%	95.77%	95.53%
10	50	95.72%	95.02%	95.20%
15	50	95.16%	94.66%	94.80%
20	50	95.02%	94.94%	94.76%
25	50	94.66%	94.64%	94.52%
30	50	94.71%	94.31%	94.37%
35	50	94.67%	94.30%	94.51%
40	50	94.57%	94.48%	94.24%
45	50	94.48%	94.18%	94.36%
50	50	94.18%	94.18%	93.94%



Observations:

- The above graph shows the different Training, Validation and Test Set accuracy for different regularization parameters.
- We can observe that with increase in regularization parameter the accuracy increases and gradually drops from λ = 10.
- Non-Regularized (λ = 0) performs least efficiently as the algorithm was not able to converge at local minima, consequently making it harder for the system to learn weight.

Lambda	Hidden Unit	Time(In Sec)
0	50	26.02205676
5	50	23.83797178
10	50	24.71857586
15	50	23.29902044
20	50	23.57502859
25	50	22.62298982
30	50	24.33622745
35	50	23.64994618
40	50	24.50004585
45	50	24.7822738
50	50	23.58422237



Conclusion:

With increase in regularization parameter, accuracy increases. But on further increase in value i.e. for $\lambda > 10$, accuracy starts to drop and error does not converge. Hence, we can conclude for backpropagation on multilayered neural network with 50 hidden nodes, works best with the regularization parameter of 5. It can also be observed that there is a removal of *124 features*.

Problem 2: Comparison between Neural Network and Deep Neural Network

Objective:

- Compute the accuracy of CelebA data using facennScript using the regularization parameter obtained from nnScript. (*Neural Network*)

Lambda and Hidden Unit values are set to default. Following values are obtained:

Lambda	Hidden Unit	Training set Accuracy	Validation Set Accuracy	Test Set Accuracy
10	256	85.02%	84.20%	85.39%

Lambda and Hidden Unit values are set to the values obtained from the previous observation. Following values are obtained:

Lambda	Hidden Unit	Training set Accuracy	Validation Set Accuracy	Test Set Accuracy
5	50	85.11%	84.84%	85.47%

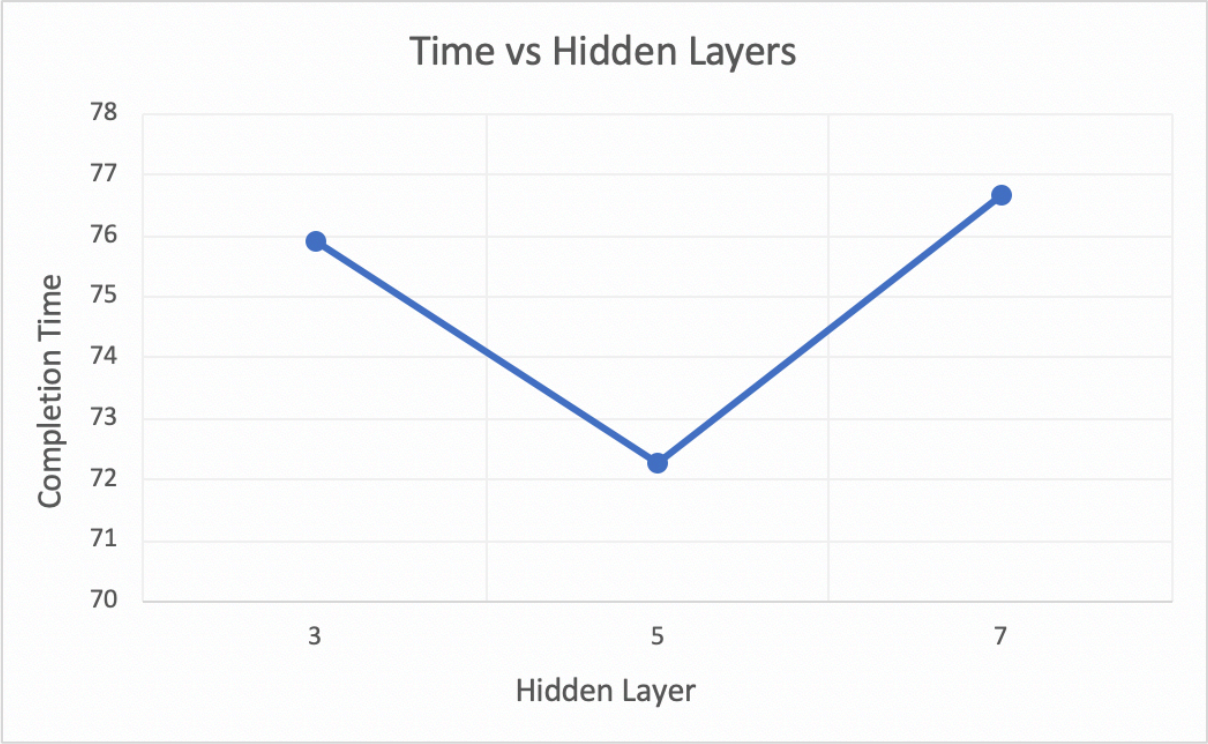
Objective:

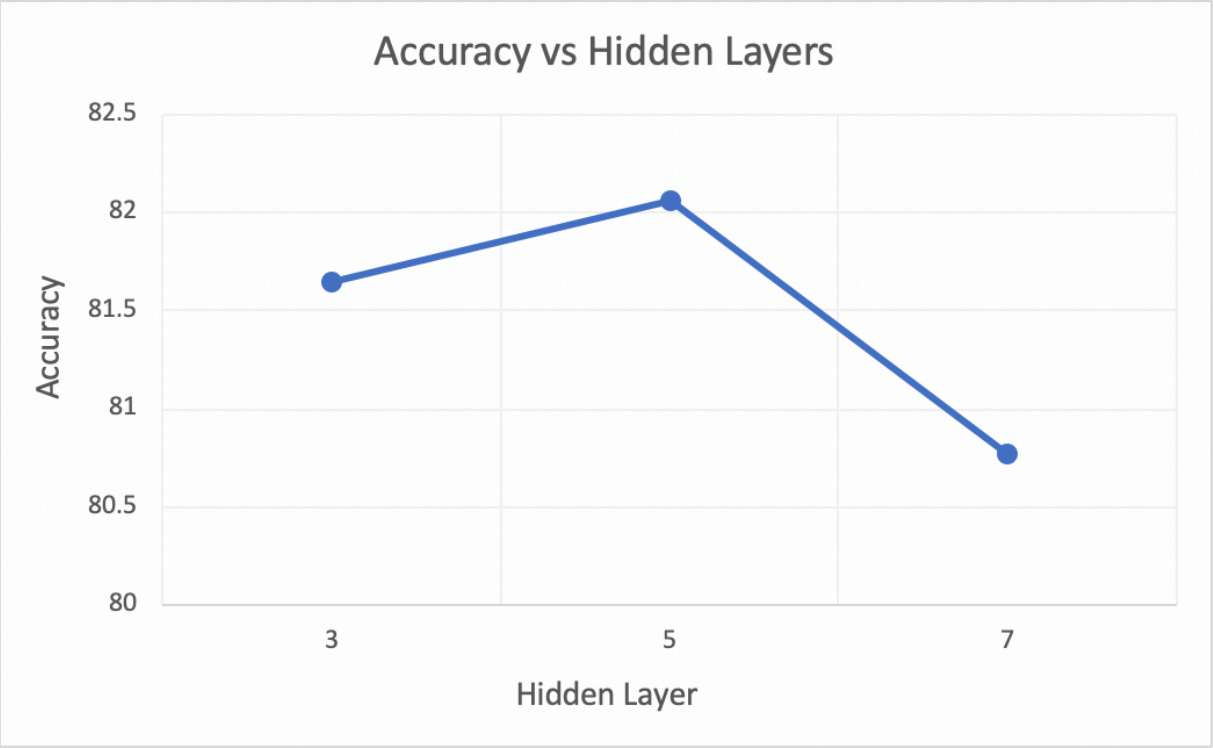
- Compute the accuracy of CelebA data using the deepnnScript – Deep Neural Network

Hidden Layers	Accuracy	Completion Time (in seconds)
3	81.642693	75.91169906
5	82.059044	72.29148412
7	80.772144	76.68332076

Conclusion:

- We conclude from the above observation that our Neural Network does a better job than the deep neural network as our task is not that complicated and adding more hidden layers merely leads to overfitting.


Time vs Hidden Layers
78
77
76
75
74
73
72
71
70
Completion Time
3
5
7
Hidden Layer


Accuracy vs Hidden Layers
82.5
82
81.5
81
80.5
80
Accuracy
3
5
7
Hidden Layer

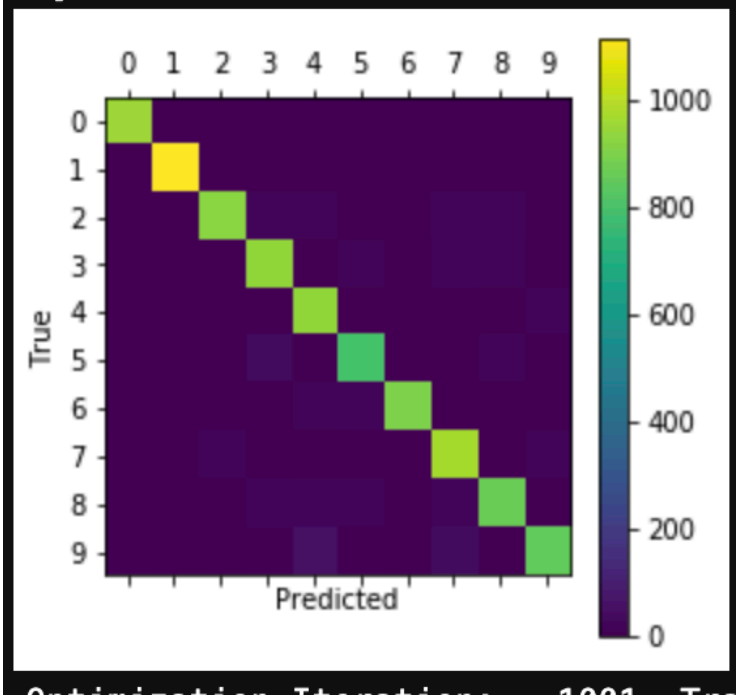
Problem 3: Convolutional Neural Network

Objective:

- Results from Convolutional Neural Network in terms of accuracy and time.

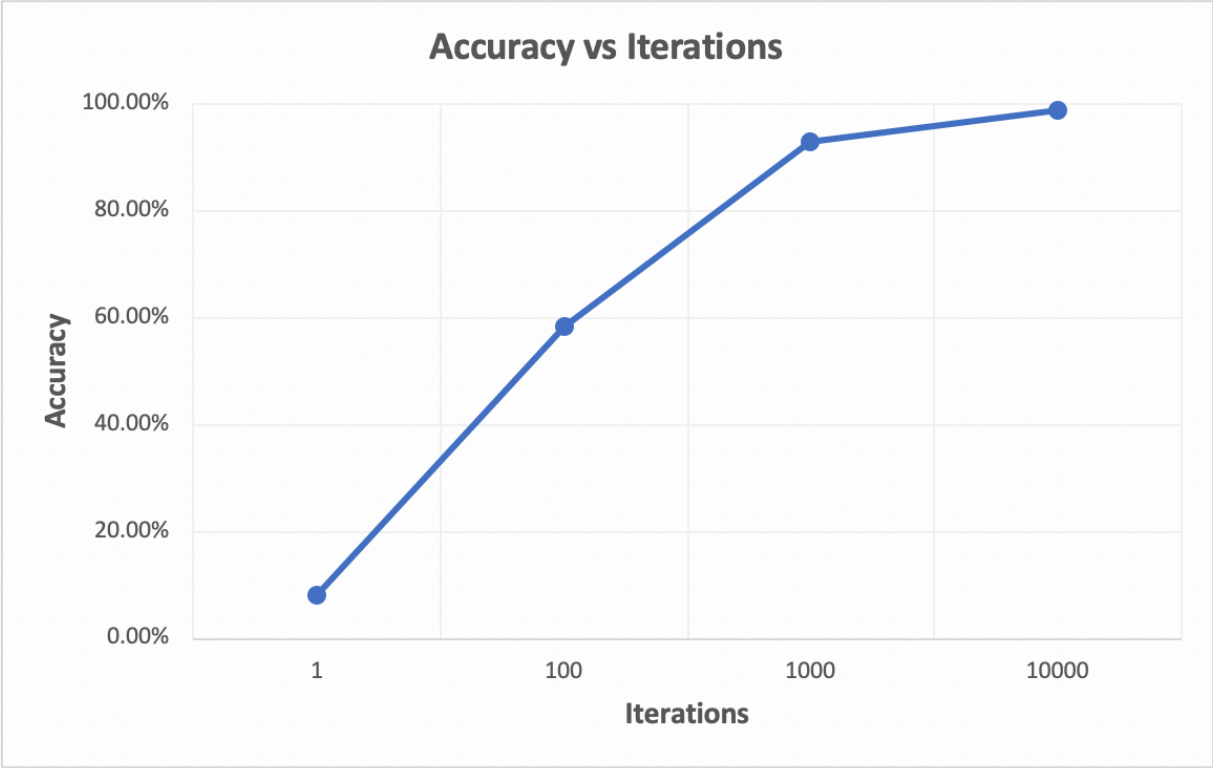
Results:

```
Time usage: 0:00:28
Accuracy on Test-Set: 92.9% (9293 / 10000)
Confusion Matrix:
[[ 953    0    0    1    0    4   10    4    8    0]
 [   0 1119    3    2    0    1    4    0    6    0]
 [  10    2  931   16   17    1    9   20   26    0]
 [   1    1   11  942    0   14    0   21   17    3]
 [   1    2    2    0  942    0    9    2    3   21]
 [   6    2    3   36    8  804   13    2   16    2]
 [   9    5    2    0   16   18  903    2    3    0]
 [   0    9   18    6    6    1    0  970    2   16]
 [   7    6    5   22   14   17    6   17  871    9]
 [   9    5    4   12   57   11    0   48    5  858]]
```



Observations:

Iterations	Test Accuracy
1	8.30%
100	58.50%
1000	93.10%
10000	98.70%



Iterations	Time (in secs)
1	0
100	3
1000	28
10000	285