

DESCRIPTION:

This experiment is about the implementing the perceptron learning algorithm. In this algorithm, training and testing is done on MNIST dataset. MNIST dataset contains 60,000 training data and 10,000 test data of handwritten digits. Perceptron will learn to recognize handwritten digit. There are a 10 perceptron, as group it will classify the hand written digit. Each perceptron has 785 inputs and one output. Each perceptron's target is one of the 10 digit (0-9). Input is scale down between 0 and 1 by dividing the input value with 255. Perceptron's weights value is random value between -0.05 and 0.05. Perceptron will train the model on three different learning rate. The training period is for 50 epoch. After each epoch the accuracy of training and testing is compared and plot the graph for epoch vs accuracy. It is shown below.

Results:

The graphs of epoch vs accuracy are shown below for different learning rate.

1. Learning rate: 0.001

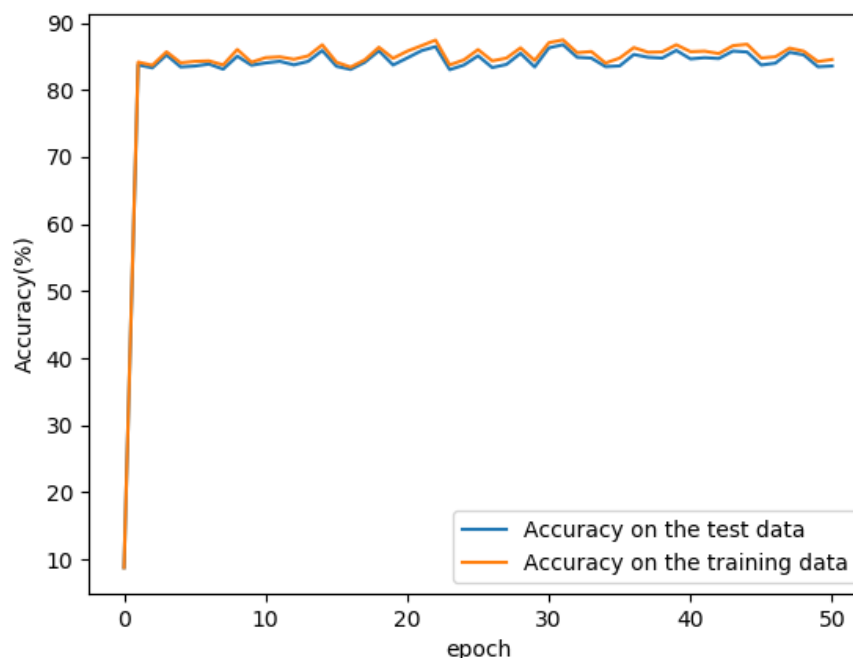


Figure 1: graphs of epoch vs accuracy for learning rate 0.001

The initial accuracy (before epoch 0) is 7.8 % and after training for one epoch the accuracy is 85.8% for the training data and 86.2% for the test data. As we can clearly see the oscillation after one epoch to next 50 epoch.

Confusion matrix:

		Predicted digit									
Digit	0	1	2	3	4	5	6	7	8	9	

```

[[ 924    0    5    4    1    8   17    4   11    6]
 [    0 1118    1    2    0    1    4    0    9    0]
 [    1   48  763  121   13    7   20   11   44    4]
 [    2    5   16  937    2   10    4   10   12   12]
 [    1    3    4    3  811    0   17    1   11  131]
 [    5    5    3  107   20  619   37    4   55   37]
 [    5    5    1    2    5   21  912    0    7    0]
 [    1   15   17   14    8    1    1  863    8  100]
 [    4   40    6  106   24   10   21    7  706   50]
 [    4    9    0   21   15    4    2   12    6  936]]

```

As we can see that, digit 4 is more confused with digit 9 and more predicted it. Digit 2 is more predicted as digit 3. Digit 1 and digit 3 is more accurately predicted.

2. Learning rate: 0.01

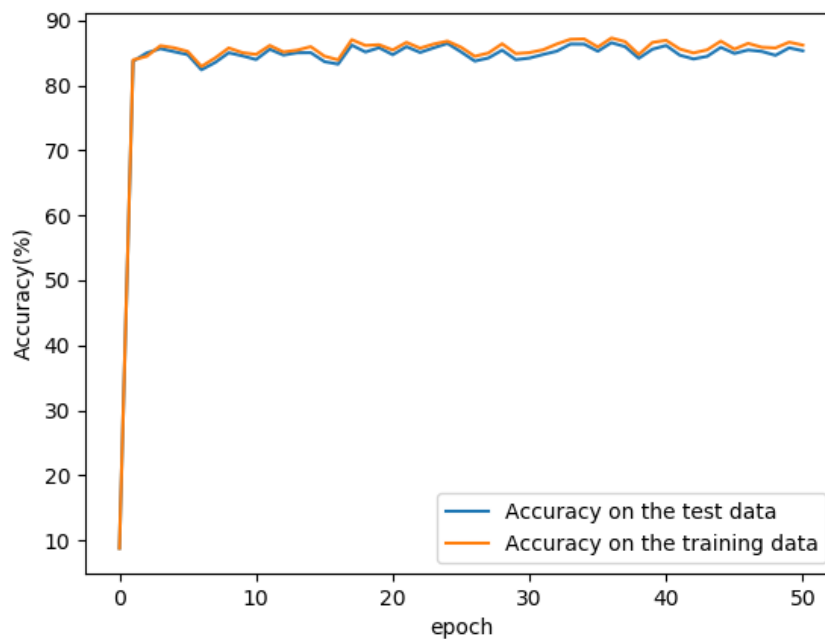


Figure 2: graphs of epoch vs accuracy for learning rate 0.01

After training for one epoch the accuracy is 83.7% for the training data and 84.4% for the test data. As we can clearly see the oscillation after one epoch to next 50 epoch.

Confusion matrix:

		Predicted digit									
Digit	0	1	2	3	4	5	6	7	8	9	

```

[[ 925    0    3    2    1    4   29    3    9    4]
 [    0 1104    1    5    0    0    3    1   20    1]
 [    5   31  720  153    7    3   39   13   54    7]
 [    4    1    9  947    2    3    5    8   14   17]
 [    1    3    3    1  769    1   32    5   17  150]
 [    8    5    2  144   11  532   43   11   82   54]
 [    5    3    1    1    3   12  923    1    9    0]
 [    2    8   19   19    6    0    1  896    7   70]
 [    5   17    6  104   13    7   17    7  754   44]
 [    5    4    1   17   10    2    2   14   10  944]]

```

As we can see that, digit 2 is more confused with digit 3 and more predicted it. Digit 4 is more predicted as digit 9. Digit 1 and 3 is more accurately predicted.

3. Learning rate: 0.1

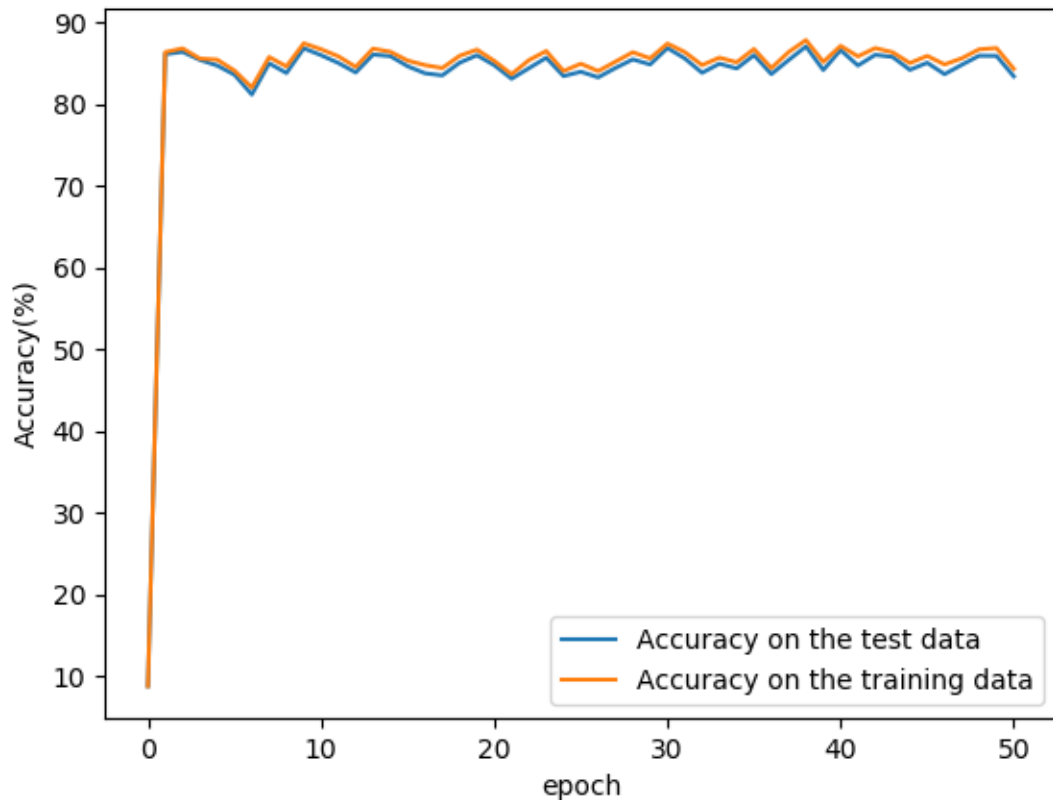


Figure 2: graphs of epoch vs accuracy for learning rate 0.01

After training for one epoch the accuracy is 86.7% for the training data and 86.4% for the test data. As we can clearly see the oscillation after one epoch to next 50 epoch. The oscillation increased for this learning rate compare to other two learning rate.

Confusion matrix:

Digit	Predicted digit									
	0	1	2	3	4	5	6	7	8	9
0	927	0	5	3	0	3	22	3	12	5]
1	0	1111	2	2	2	0	3	1	13	1]
2	4	42	799	92	6	4	16	13	52	4]
3	4	2	13	932	2	7	3	10	17	20]
4	1	2	9	2	815	0	19	1	15	118]
5	8	2	5	123	18	558	36	11	76	55]
6	4	5	5	1	2	18	914	1	8	0]
7	1	11	26	8	15	2	1	870	12	82]
8	6	34	6	65	31	4	18	7	741	62]
9	5	5	3	15	21	1	1	10	8	940]]

As we can see that, digit 6 is more confused with digit 3 and more predicted it. Digit 4 is more predicted as digit 9. Digit 1 and digit 9 is more accurately predicted.

Conclusion:

We can observe that as learning rate increase, the oscillation is increased, too. For above results, it is not clearly visible that there is any overfitting but we can observe from the graphs that if we decrease the learning rate then there might be overfitting. As learning rate increase, the accuracy fluctuate more over epoch.