



CS 539: Machine Learning

Assignment 4 Part 2

Jiaming Nie

November 17, 2018

Content

1	Part II Intro	1
2	Parameter: Epoch = 30, batch_size = 10, $\eta = 3.0$	1
3	Hyper-parameter and Network Topology Result	1
3.1	Parameter I, Epoch = 30, $\eta = 0.1$	1
3.2	Parameter II, Epoch = 30, Hidden Units = 100, $\eta = 3.0$	2
3.3	Parameter III, Epoch = 70, Batch Size = 16, $\eta = 3$	3
4	Summary on the Parameter of Neural Network	3

1 Part II Intro

This part is to summarize the results running on the MNIST dataset using the code provided in the assignment.

The result will include related hyper-parameters of the training progress and network architecture.

2 Parameter: Epoch = 30, batch_size = 10, $\eta = 3.0$

Using the parameter defined in the assignment, run 3 times and record these parameters.

The network structure remains the same with the 30 hidden units.

Table 1: Result Table for Running Three Times

Trial Number	Predict Number	Prediction Rate
1	9484	0.9484
2	9463	0.9463
3	9472	0.9472
Median	9472	0.9472

According to the table 1, the median result on test dataset is 9472/10000.

3 Hyper-parameter and Network Topology Result

3.1 Parameter I, Epoch = 30, $\eta = 0.1$

In this section, the parameter is set to 30 and the learning rate η set to 0.1.

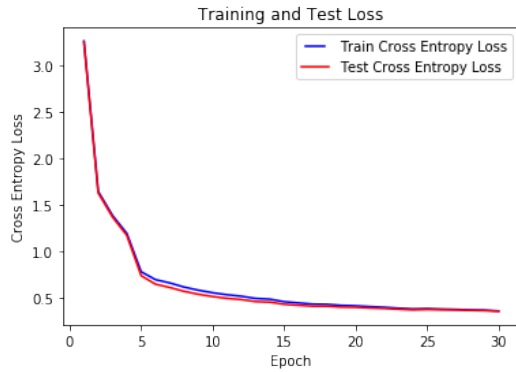


Figure 1: Cross Entropy Plot

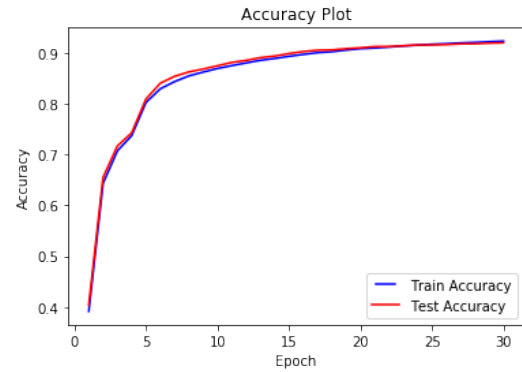


Figure 2: Accuracy Plot

The accuracy on validation data is 0.9225 (9225/10000). The confusion matrix for the classification is illustrated in the figure 3.

In the training progress of the preliminary stage, the convergence is a bit slower and the cross-entropy loss is higher than 1.

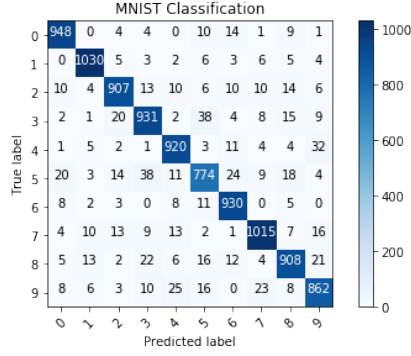


Figure 3: Classification Confusion Matrix on Validation Data

This stage only changed the learning rate η to 0.1, which illustrated a slower convergence speed.

3.2 Parameter II, Epoch = 30, Hidden Units = 100, $\eta = 3.0$

Set the number of neurons to 100 at the hidden layer, with the learning rate $\eta = 3$ as the default value.



Figure 4: Cross Entropy Plot

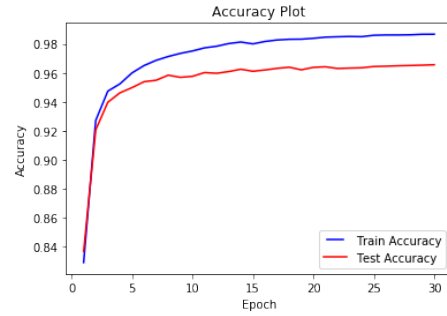


Figure 5: Accuracy Plot

The accuracy on validation data is 0.9662 (9662/10000). The confusion matrix for the classification is illustrated in the figure 6.

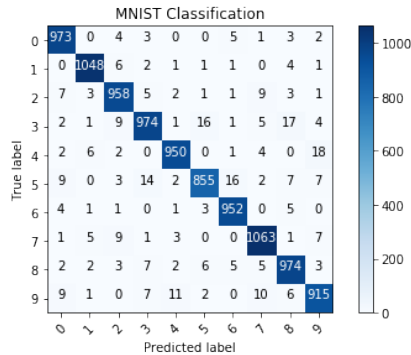


Figure 6: Classification Confusion Matrix on Validation Data

3.3 Parameter III, Epoch = 70, Batch Size = 16, $\eta = 3$

Based on the parameters obtained in the second subsection, set the Epoch to 70, the batch size to 16, the hidden units to 100.

The modification of the neural network is to verify the tuning parameters based on the good parameters can achieve better training results.

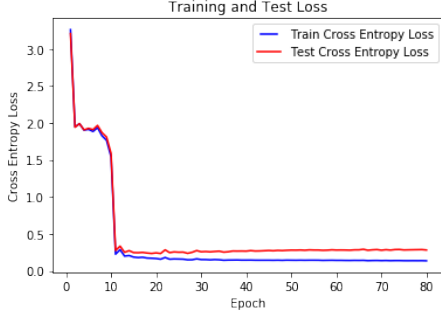


Figure 7: Cross Entropy Plot

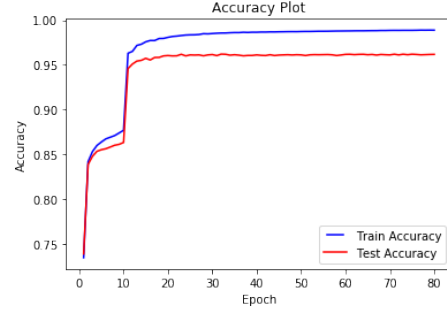


Figure 8: Accuracy Plot

The accuracy on validation data is 0.9644 (9644/10000). The confusion matrix for the classification is illustrated in the figure 6.

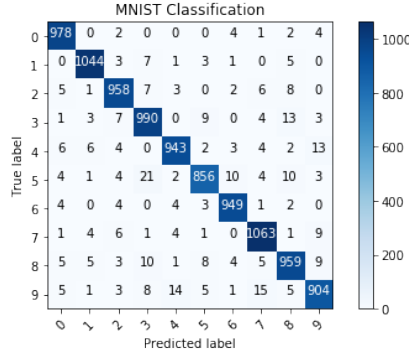


Figure 9: Classification Confusion Matrix on Validation Data

4 Summary on the Parameter of Neural Network

The summary of the parameters tuning on the neural network is on the following:

- With the same settings on other parameters, more epochs will contribute to higher accuracy.
- Under certain circumstances, more hidden neurons will contribute to higher convergence speed and higher accuracy.
- The change of the learning rate will lead to different convergence condition, even divergence.