Deep Learning Lab Course: Assignment 01

Yufeng Xiong

November 10, 2016

1 Network Parameters

In my implementation of MLP, I used 1 Input layer, 4 FullyConnected layers and 1 SoftmaxOutput layer to classify mnist dataset. Here are some key parameters of the network architecture:

1.1 Hyperparameters

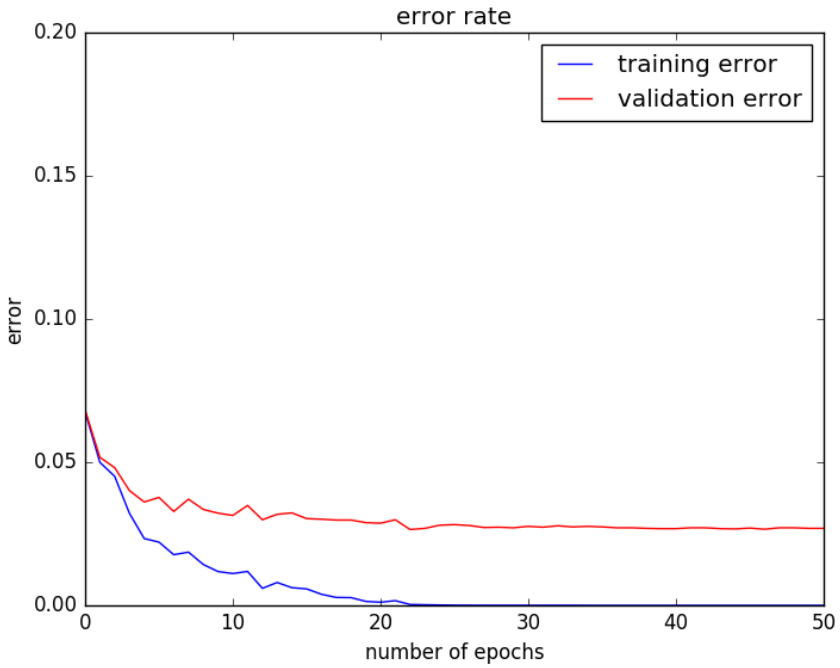
a. Max epochs: 50
b. Learning rate: 0.05
c. L2 regularization factor: 1e-5
d. Batch size: 64
e. Descent type: stochastic gradient descent

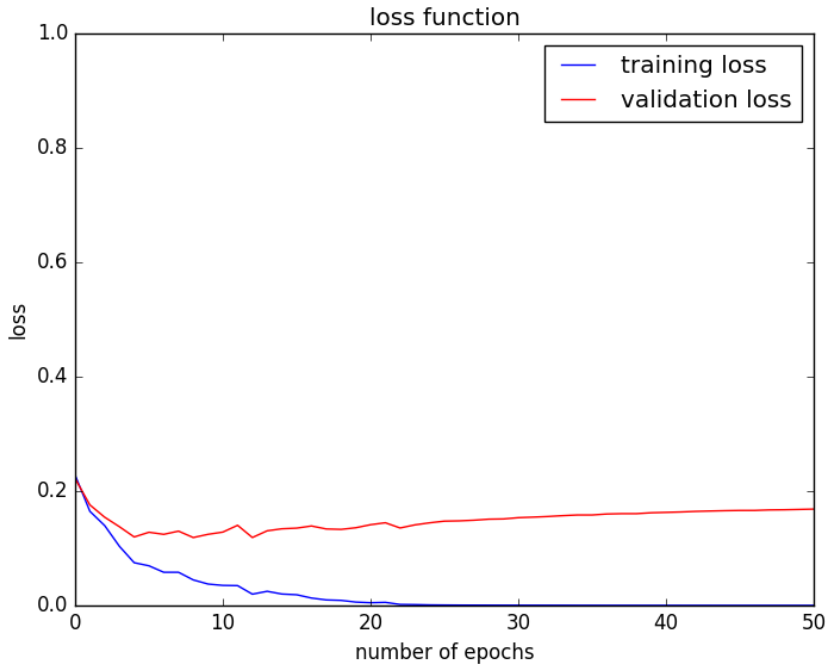
1.2 Network architecture

a. InputLayer(input_shape)
b. FullyConnectedLayer(layers[-1], num_units=256, init_stddev=1.0, activation_fun=Activation('relu'))
c. FullyConnectedLayer(layers[-1], num_units=128, init_stddev=0.1, activation_fun=Activation('relu'))
d. FullyConnectedLayer(layers[-1], num_units=64, init_stddev=0.1, activation_fun=Activation('relu'))
e. FullyConnectedLayer(layers[-1], num_units=10, init_stddev=0.1, activation_fun=Activation('relu'))
f. SoftmaxOutput(layers[-1])

2 Implementation results

After training the network on the complete data set, the validation error rate reaches around 2.76%, and the test error rate is about 2.79%. Here are the plotting figures of error rates and loss values:


error rate
training error
validation error
error
number of epochs
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50


loss function
training loss
validation loss
loss
number of epochs
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50