

Spring 2022 Introduction to Deep Learning

Homework Assignment 3

Due date: April 8 2022

Problem (Build a FC layer-based neural network to recognize hand-written digits). In this problem, you are asked to train and test a neural network for *entire* MNIST hand-written digit dataset. Some information of the network is as follows:

- Its structure is **784-200-50-10**. Here 784 means the input layer has 784 input neurons. This is because each image in MNIST dataset is 28x28 and you need to stretch them to a length-784 vector. 200 and 50 are the number of neurons in hidden layers. 10 is the number of neurons in output layer since there are 10 types of digits.
 - The two hidden layers are followed by **ReLU layers**.
 - The output layer is a **softmax** layer.
- (a) **(Mandatory)** Use deep learning framework to train and test this network. You are allowed to use the corresponding autograd or nn module to train the network.
- (b) **(Optional)** Use only Numpy to train and test this network. You are NOT allowed to use deep learning framework (e.g. Pytorch, Tensorflow etc.) and the corresponding autograd or nn module to train the network.

Performance Requirement and Submission:

- The test accuracy should achieve above 95%
- Submission should include your source codes and screen snapshot of your train and test accuracy, plus the training time

Suggestion for hyperparameter setting (not necessary to follow):

- Learning rate can be set as 0.01
- If you choose to use mini-batch SGD, the batch size can be set as 128
- Number of epochs can be set as 10.

About dataset loading:

- For MNIST dataset loading using Numpy, you can refer to the solutions in HW1. Notice that in HW1 we only use part of the dataset.
- For dataset loading in Pytorch, you can refer to the following example code
<https://github.com/pytorch/examples/blob/master/mnist/main.py>