Homework1

Instructions

All solutions must be written in your own words, and you must program the algorithms yourself. Your programs are suggested to be written in Python. The relevant code to the problem should be in the PDF you turn in. If a problem involves programming, then the code should be shown as part of the solution to that problem. If you solve any problems by hand just digitize that page and submit it (make sure the problem is labeled).

In this homework, you will be implementing neural network to recognize handwritten digits using stochastic gradient descent and the MNIST data. MNIST dataset contains images of handwritten digits, together with their labels. The dataset can be downloaded here using Pytorch.
The images are 28 by 28 greyscale images. It contains 60,000 training data and 10,000 test data. Break the training data into training (70%) and validating (15%) and testing (15%) dataset.

1) Describe the training process of neural network in your own words. You may need to write the update equation for a_j^l and z_j^l in the forward step, δ_j^l in the backward step and the update equation for parameters $w_{j,k}^l$.
2) Implement backpropagation to train a feedforward neural network (show in Fig.1) with specified hyperparameters (number of layers, number of neurons in each layer). You need to convert the 28 by 28 pixel values into a long vector with 784 elements as your input. Add a bias neuron for input and hidden layer. Use MSE error as the loss function and sigmoid function as the activation function. You can use the validating data to find a fairly good learning rate for each architecture. You need to code from scratch for this problem.
3) Reimplement the same network in 2) with Pytorch. Compare the result with 2) to see if you could get same results in terms of loss function and classification accuracy.
4) Try to use the training tips you learned from class to improve the training process in terms of accuracy on training and test dataset, learning speed. Compare the learning curves when you applied these training tips.

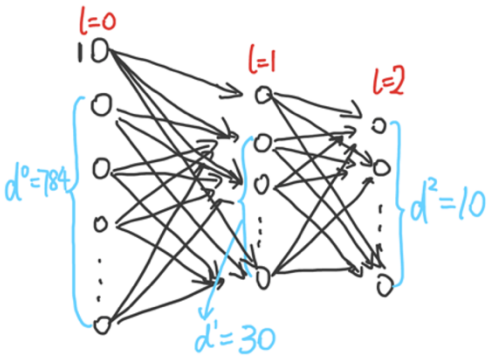

Figure1 A 2-layer neural network with 30 neurons in the hidden layer (l=1).