

Exercise 1 Report

Deep Learning Lab Course

Lab Date: 16/10/2018

Due Date: 30/10/2018

Ahmed Abdelhadi

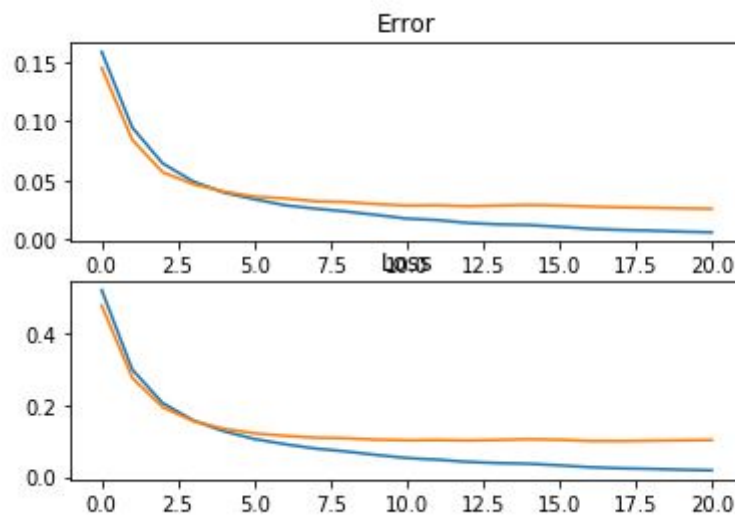
Introduction

In this exercise, a neural network was implemented from scratch. The neural network aimed to solve multiclass classification problems, namely the MNIST dataset. Both stochastic gradient descent and gradient descent were implemented as optimizing techniques. A cross-entropy loss function was implemented as well.

After implementing the neural network, its performance was tested against the MNIST dataset with the objective to optimize the hyperparameters to get to the best possible validation errors. Hyperparameters included: the learning rate, no. of layers, no. of units per layer, the activation function of each layer, max. epochs and batch size.

Predefined Network

The predefined network yielded a validation error of approximately 0.0256 after 20 epochs. The following is a graph for the train/validation error and loss.



Hyperparameter Optimization

As mentioned earlier, there are many hyperparameters. I chose to experiment with the activation functions and learning rate while keeping the other hyperparameters fixed.

Following is a table of the different results:

No. of Layers	No. Of Units	Activation Function	Max Epoch	Batch Size	Learning Rate	Validation Error
4	256-128-64-10	relu - tanh - relu	20	64	0.25	0.0247
					0.5	0.0226
		relu - tanh - tanh			0.25	0.0227
					0.5	0.0227