

High Performance Computer Architectures

Practical Course

- Exercise 2 -

April 26, 2023

Neural Networks

The following exercises will ask you to implement, from scratch, a simple fully connected neural network in C++ and then apply it to the task of handwritten digit recognition (MNIST dataset). We provide you with the general structure of the neural network and leave only some of the math implementation blank. We will come back to this neural network again later in the course and improve its performance.

Problem 1

Follow the comments in the code to complete the neural network:

1. Complete the `feedForward()` function in `MLPNet.cpp`. You will need to write Affine transformation, Softmax and `MatMul2D1D` functions for this.
2. In addition to TanH, implement a ReLU activation. The ReLU activation function is defined as

$$ReLU(x) = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{otherwise} \end{cases} . \quad (1)$$

When the argument to ReLU is a vector, it is applied element-wise. Also write the back-propagation for ReLU using

$$ReLU'(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{otherwise} \end{cases} . \quad (2)$$

3. Complete the `backPropagation()` function. You will need to write a matrix transpose function for this. The `backPropagation()` function is a bit complex so please spend a little time just to understand it.
4. Download the MNIST data files from [this link](#). Add the paths to the downloaded data in the `loadMNISTEpochData()` function. The code should now compile and run.

5. Run the network for topology = $\{784, 10\}$, 10000 training examples, 10000 testing examples, batch size = 10, number of epochs = 10, activation function TanH. Write down the peak training and testing accuracies.

Problem 2

Play around with the neural network! Run the neural network with different hyperparameters for example with more hidden layers, less/more training examples, more epochs, different learning rate or a different activation function(s). Write down two observations you made about how the different hyperparameters affect the neural network performance. Please provide some data on which you base these observations.