Fall 2018 IE534/CS598: HW1

Name: Ziyu Zhou, **NetID**: ziyuz2

Test accuracy

98.28%

Usage

Type `python3 main.py` in terminal.

> Note that the default path for the dataset is `"data/MNISTdata.hdf5"`. If a different path is used, please change the input path to the `load_data` function in `main.py`.

Implementation

The implementation is separated into four files, namely:

- `main.py`: the main file to execute, which contains the high level pipeline of the overall implementation, including loading the dataset, initializing the model, training and testing.
- `model.py`: contains the architecture of the neural network with a single hidden layer. The model is implemented as a `NeuralNetwork` class which supports weight initialization, training and testing. There are mainly two public functions that can be called by the `NeuralNetwork` object:
  - `train`: train the neural network on the training dataset using SGD.
  - `test`: test the trained model on the testing dataset.

  The other functions, i.e., `_forward_step`, `_backward_step`, `_update_weights` and `_predict` are private functions which help with the training and testing process.
- `io_tools.py`: contains tools to load the MNIST dataset.
- `activate_functions.py`: implements activation functions for later use, including ReLU and softmax, as well as the gradient for ReLU.