MNIST Classification with Softmax

Homework 1 for Introduction to Deep Learning, Fall 2020

Deadline: 2020.10.11 23:59:59

1 Introduction

MNIST digits dataset is a widely used dataset for image classification in machine learning field. It contains 60,000 training examples and 10,000 testing examples. The digits have been size-normalized and centered in a fixed-size image. Each example is a 784×1 matrix, which is transformed from an original 28×28 grayscale image. Digits in MNIST range from 0 to 9. Some examples are shown below. **Note**: During training, information about testing examples should never be used in any form.

In this homework, you are required to use **Softmax Classifier** to perform MNIST classification.

2 Softmax for MNIST Classification

2.1 Files Description

There are several files included in the **./homework1-softmax/** folder:

- **homework1-softmax.ipynb** is an IPython Notebook file which describes the main contents of this homework. Data loading, hyerparameters setting, training and testing are included in this file. **Please read this file carefully**.
- **mnist_data_loader.py** is used to load MNIST dataset. Not required reading. **To use this file please install TensorFlow**. (conda install -n [your_env_name] tensorflow).
- **softmax_classifier.py** describes the softmax classifier. **You are required to complete** the function $softmax\_classifier(W, input, label, lamda)$. Part of this function is provided and you need to write down your code at '# TODO' to calculate the loss, gradient and prediction.

2.2 Requirements

You are required to complete the '# TODO' parts in above files. If implemented correctly, just by running the IPython notebook step by step, you can obtain lines of output and reach a relatively high test accuracy. **You need to submit all codes and a short report** with the following requirements:

- Record the training and testing accuracy, plot the training loss curve and training accuracy curve in the report.
- The given hyerparameters maybe performed not very well. You can modify the hyerparameters by your own, and observe how does these hyerparameters affect the classification performance. Write down your observation and record these new results in the report.

3 Attention

- You need to submit all codes and a report (at least two pages **in PDF format**). Delete the MNIST dataset before submit.
- Pay attention to the efficiency of your implementation. Try to replace the use of **for-loops** with matrix multiplication.
- Do not paste a lot of codes in your report (only some essential lines could be included). Any extra modifications of above homework files or adding extra Python files should be explained and documented.
- Any open source neural network toolkits, such as TensorFlow, Caffe, PyTorch, are **NOT** permitted in finishing homework-1.
- **Plagiarism is not permitted**.