**Artificial Neural Network and Deep Learning**

Project #1: Back-Propagation Neural Network Implementation and Application to Handwritten Digits Recognition

Project Deadline: October 27th, 2019

Evaluation Time: 1:10PM, October 28th, 2019

Location: Electrical Engineering and Computer Sciences Building, Room 1F04

**Subject:**

Implement the **back-propagation** neural network algorithm to classify/recognize the handwritten digits. A database of 40,000 examples of handwritten digits is given as a training set. The dataset contains 10 folders, namely 0, 1, 2, …, and 9. Each folder contains 4,000 sample images of the corresponding digit. All digits have been size-normalized and centered in a fixed-size image (28\*28 pixel, PNG format) as shown in the following examples.

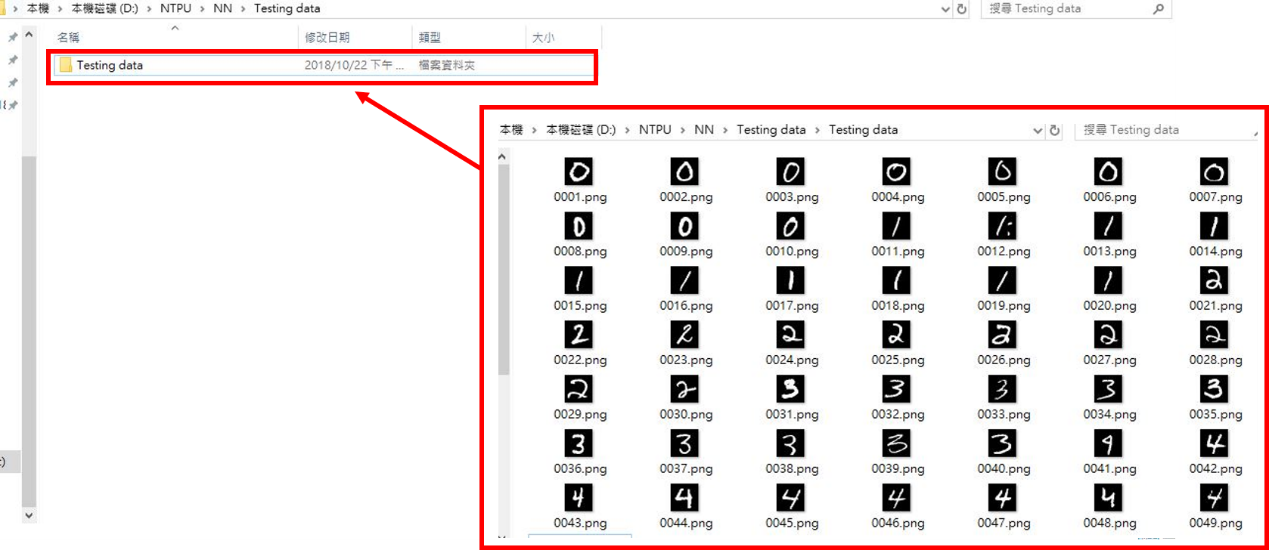
It is a good database for people who want to try learning techniques and pattern recognition methods on real-world data while spending minimal efforts on preprocessing and formatting. For network performance evaluation, you will need to meet TA at 1:10PM on October 28th and test on a set of 5,000 new examples.

**Note:**

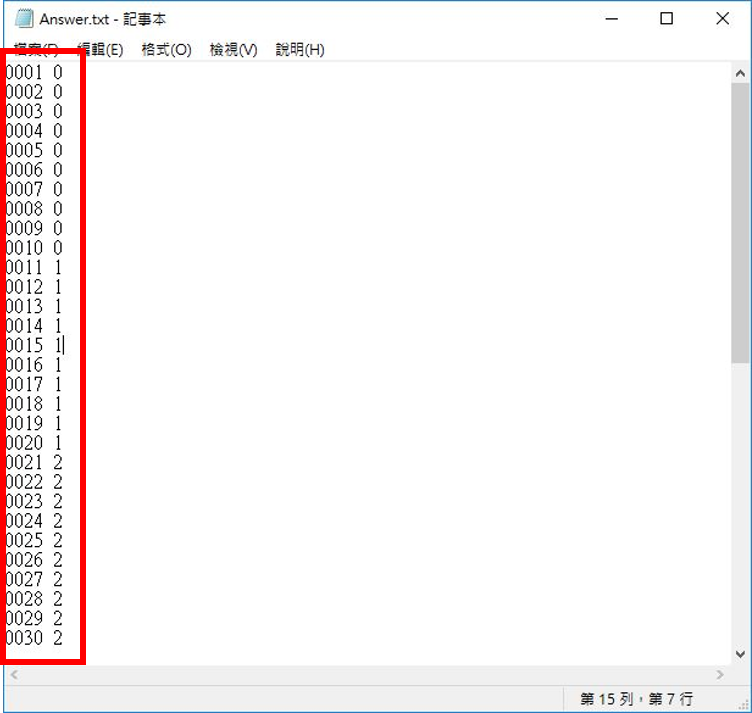
Design your own code using any programming languages. **DO NOT** apply any open source code of BP nor existing BP library/function for this project. To ensure the program is designed by yourself, you will be requested on-site on October 28th to re-train the BP model for only part of the digits with the given dataset and the evaluation will be also performed on-site for those particular digits.

**Evaluation Method:**

1. After you have done with the training as requested on-site, meet TA for evaluation. TA will prepare an USB memory stick containing a folder called “Testing data”. This folder includes 5,000 test images with file names 0001.png, 0002.png, …, and 5000.png as illustrated in the following:



1. Your program needs to read all test files in the folder, perform the recognition for each digit image, and then generate a text file output as shown in the following format. The output file must contain two columns: first column is the filename; the second column is the recognition result of the digit. For example, 0001 represents a file name of test image 0001.png and 0 represents your recognition result as digit 0.
2. Then TA will check your output file and let you know the accuracy of your recognition results.



Project Grading Policy:

1. (75%) The overall test accuracy
2. (25%) Presentation: sharing your findings or experiences.

Note: The presentation will be held together with Project #2.

Contact TA for any questions/comments:

Miss Wu.

Email: k3roneji@gmail.com