HW2-ANN (10 points for Final Exam)

In the context of this assignment, you will build an ANN structure by using MNIST train data to predict the MNIST handwritten digits. You will use MNIST test data to evaluate your trained ANN structure in terms of evaluation metrics that are accuracy and F1 score. You will codify your code with any programming languages (Python is a good choice). You can take advantage of some libraries such as scikit-learn, etc.

To make comparison in your report, you will train your ANN structure with activation functions that are sigmoid, ReLU, and tanh separately. You will use the same activation functions in hidden layer and output layer for your ANN structure. You will add the convergence curve of your ANN structure along iterations to the report. You will explain your ANN structure in terms of the included number of neurons in hidden layer and output layer. You will explain which value of learning rate is used for your ANN structure. You will explain which activation function provides the highest or the lowest classification performance. You will upload your report with your source code file as a zip file. The name of zip file should be saved in form of schoolnumber_name_surname.

Information about MNIST dataset

The original dataset is in a format that is difficult for beginners to use. This dataset uses the work of Joseph Redmon to provide the MNIST dataset in a CSV format.

The dataset consists of two files:

1. mnist_train.csv
2. mnist_test.csv

The mnist_train.csv file contains the 60,000 training examples and labels. The mnist_test.csv contains 10,000 test examples and labels. Each row consists of 785 values: the first value is the label (a number from 0 to 9) and the remaining 784 values are the pixel values (a number from 0 to 255).