Homework 1

Implement a deep learning model with Python Tensorflow for classification of digit (0-9) from computer fonts with variations.

The training set and validation set are in the directory. Keep in mind that we'll use a testing set other than both the datasets we give to you for the performance evaluation of your model. Thus, you'd better not use the validation set for training in order to avoid over-fitting.

Name your top-level python file (module) as "machine_learning_hw1.py"

Provide two functions in that module:

1. *train(data_dir) -> None*

   data_dir: <string> the directory containing all training images

2. *test(data_dir) -> predict, ground_truth*

   data_dir: <string> the directory containing all testing images

   Return:

   predict: <list of int> predicted class labels

   ground_truth: <list of int> answer class labels in corresponding order to *predict*

****

The order of the labels depends on how you load images from the directory, so you are responsible to return the ground truth along with the prediction. However, we'll view your code, so POINTS WOULD BE TAKEN IF FAKE GROUND TRUTH ARE GIVEN.

****

See autograder.py in the directory for how we use the functions in detail.

We may use code not exactly the same as in autograder.py to evaluate your model performance.

****

The score outputted by the provided autograder is not your final score but an indicator for your model performance.

****

The homework is graded by:

- Performance: classification accuracy
- Code cleanness: e.g. comments, meaningful variable names, …
- Explanation: write explanation for your code (e.g. model structure, methods, …) in a PDF file.

Good luck!