Machine Learning

ECE 4332 / ECE 5370

Project 4

1. Use your implementation of logistic regression to classify the following two datasets.
    a. MNIST dataset of handwritten digits.
    b. C. elegans dataset: Two classes: 1-worm (t=1), and 2- no worm (t=0).
2. Generate a table for each dataset that includes:
    a. Training information such as
        i. Visual verification of input data
        ii. Training, validation, and testing data splits
        iii. Input image size used for training/testing
        iv. Image preprocessing, if any
        v. Parameters of logistic regression
        vi. Optimizer type and corresponding parameters
    b. Testing information presented as a confusion matrix
    c. Training and testing execution times
3. Your code must conform to the following format so that it can be tested on an independent dataset.
    a. Input:
        i. Name of the directory path containing test images (*.tif images for MNIST and *.png images for C. elegans)
        ii. Trained model
    b. Output written to an Excel file (be sure to expose the filename):
        i. Two-column list with the first column indicating the image filenames and the second column indicating the corresponding labels
        ii. Total number of images for each label
4. This is a group project. No more than three students are to form a group and make a single submission.

Archive

- **.m or .py file(s)**
- **report in pdf format**
- **trained model**

Name your file as Lastname1_Lastname2_Lastname3_Project4.zip and upload to Blackboard prior to the deadline.