CECS 551
Assignment 3
Total: 50 Points

General Instruction

- Submit uncompressed file(s) in the Dropbox folder via Canvas (Not email).
- Use `Python 3`, any other programming language is not acceptable.
- You can import modules in the following list of libraries (please check the full list *here*). If you want to use any other library, please consult with the instructor.

---

1. Tune hyper-parameters of a neural network.
   (a) Read and Understand: Familiarize yourself with the article *PyTorch Lightning classifier for MNIST*.
   (b) (5 points) Data Split: Divide `train` dataset into `sub_train` and `valid`.
      - MNIST train dataset contains 60,000 samples.
      - Allocate 50,000 samples to `sub_train` and remaining 10,000 sample to `valid`.
   (c) (5 points) Build a Neural Network: Construct a neural network according to the following specifications
      - **Exclude** convolution ('Conv') layer.
      - Use the cross entropy loss function.
      - You may design your own network architecture by adjusting the width, depth, batch size, and activation functions.
   (d) (40 points) Tune the hyperparameters for each of the following optimizers: Stochastic Gradient Descent with Momentum (SGD), AdaGrad, RMSProp, and Adam.
      - Tune the following hyper-parameters.
        - SGD (*reference*): γ, μ, τ
        - AdaGrad (*reference*): γ, τ, η
        - RMSProp (*reference*): α, γ, μ
        - Adam (*reference*): γ, β_1, β_2
      - Train the model with `sub_train` and validate it using `valid`.
      - Identify the best combination of hyperparameters for each optimizer based on validation accuracy.
      - For each optimizer, using its optimal settings, plot the curves for training and validation loss as depicted in Figure 1. Additionally, report the test accuracy (not the validation accuracy).
   (e) Submit your `ipynb` file.

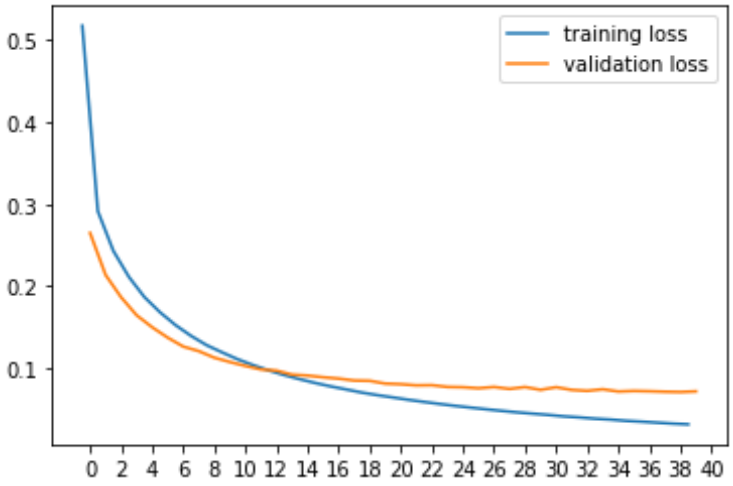

Figure 1: Training and validation loss vs. epoch