

Deep Learning Project I

Blaine Swieder

21 April 2025

1 Project I

1.1 Objective

Utilize your preferred deep learning software to implement a **fully connected neural network** to solve the classification problem for the modifications of the MNIST dataset listed below. The original MNIST dataset contains 60,000 training samples and 10,000 testing images (some may divide the 60,000 training datasets as 50,000 training and 10,000 validation set).

1.2 Task

For each of the modified datasets:

- Train your neural network and present your best results in a table that includes the following:
 - Mini-Batch Size
 - Number of Hidden Layers
 - Listing of the number of hidden units
 - Total number of parameters
 - Optimization Method Used
 - Learning Rate
 - Initializer
 - Regularization (if used)
 - Number of training epochs
 - Final Training Accuracy
 - Final Testing Accuracy
 - Ratio of the test accuracy to the 10^{th} root of the total number of parameters.
- This ratio will be used to score my results.

- Plot the convergence curve of the loss function, the training accuracy, and the testing accuracy against the training epochs.
- Discuss your results.

1.3 Dataset

1.3.1 Dataset I

- **Training Set:** The first 50,000 of the original 60,000 training images.
- **Testing Set:** Same as the original testing set.

1.3.2 Dataset II

- **Training Set:** The first 1,000 of the original training images.
- **Testing Set:** Same as the original testing set.

1.3.3 Dataset III

Use the full MNIST training and testing set but add 765 to all pixel values of the images and then divide by 4 (the resulting data may be thought of as brightened images).

1.3.4 Dataset IV

- **Training Set:** Take the first 40,000, the next 40,000 starting from the 10,001st, and the last 40,000 of the original training images, line them up in three rows, and then concatenate three images in each column to make 40,000 three-digit images.
- **Testing Set:** Concatenate in the same way the first 8,000, the next 8,000 starting from the 1,001st, and the last 8,000 of the original testing images. The output will be ordered three-digit that may be labeled as a number between 0 and 999.

1.4 Notes

Here some things to keep in mind:

- Try not to use a convolutional layer for this project.
- Refer to section 11.4 of *Deep Learning* by Goodfellow et al for some practical advice.
- Practice compiling into a pdf.