MLNS Assignment Part 1

Shreyas Kumar Sinha (2021113010, shrssnh on github)

The model

For part 1 (baseline), I used a CNN with 3 convolutional layers and 4 fully connected layers. The first convolutional layer uses a 28*28 kernel with a stride of 1, padding 14 and produces 10 convolutional layers, the second and third layers both use a 3*3 kernel with padding and stride equal to 1. Outputs from all three layers pass through a ReLU function and then through a Maxpool layer with kernel and stride equal to 2.
The output is then flattened and passed through four full connected layers that reduce the number of parameters to 256, 32, 4 and 1 respectively. Outputs from all fully connected layers pass through a ReLU function except the last one.

Training

The model is trained using PyTorch's Adam optimizer with a learning rate of 0.001 for 40 epochs. L1 Loss function is used as a measure of the loss. L1 loss is the Mean Absolute Error.

Results

Regression gives us Mean Absolute Error of **1.2624** on the full dataset. I used only regression and NOT classification so it does not make sense to provide an accuracy value.