MLNS Deep Learning Assignment - Part 1

Yashas SB

January 2025

1 Convolutional Neural Network (CNN) Architecture

The model consists of the following components:

1.1 Convolutional Layers

The network has two convolutional layers, each followed by ReLU activation and max-pooling. The first layer extracts features using 32 filters, while the second layer uses 64 filters. These layers help capture spatial patterns in the image.

1.2 Fully Connected Layers

After the convolutional layers, the output is flattened and passed through a fully connected layer with 128 neurons. Dropout is applied to prevent overfitting. The final layer outputs a single value representing the predicted sum of the digits.

1.3 Summary of the Architecture

$$\text{Input} \rightarrow \text{Conv2D} \rightarrow \text{ReLU} \rightarrow \text{MaxPool} \rightarrow \text{Flatten} \rightarrow \text{FC} \rightarrow \text{Output}$$

2 Model Training

The model is trained using the following process:

2.1 Data Preprocessing

The input images are first transformed using a series of operations. These include converting the images to tensors and normalizing them with a mean of 0.5 and a standard deviation of 0.5. This helps in stabilizing the training process.

2.2 Dataset Split

The dataset is divided into training and validation sets, with 80% of the data used for training and 20% for validation. The data is loaded in batches using a DataLoader.

2.3 Training Loop

During each epoch, the model is trained on the training set using the following steps:

- The model's predictions are compared to the true labels using Mean Squared Error (MSE) loss.
- The optimizer (Adam) updates the model's weights based on the computed gradients.
- The training loss is computed and tracked for each batch.

2.4 Validation

After each epoch, the model is evaluated on the validation set. The loss is computed for the validation data, and several accuracy metrics are calculated, including:

- **Floor Accuracy:** The proportion of predictions rounded down to the nearest integer that match the true value. I observed a floor accuracy of 10.6% on my validation set.
- **Ceiling Accuracy:** The proportion of predictions rounded up to the nearest integer that match the true value. I observed a ceiling accuracy of 12.03% on my validation set.
- **Round Accuracy:** The proportion of predictions rounded to the nearest integer that match the true value. I observed a rounded accuracy of 11.72% on my validation set.
- **Mean Absolute Error (MAE):** The average absolute difference between the predicted and true values. I observed a mean absolute error of 2.8462 on my validation set

2.5 Model Saving

The best model (with the lowest validation loss) is saved after training. This model can be used later for inference on new data.

2.6 Training Summary

The model is trained for 100 epochs, and the training process is monitored by tracking both the training and validation losses. The training and validation accuracy metrics are plotted to visualize the model's performance over time.

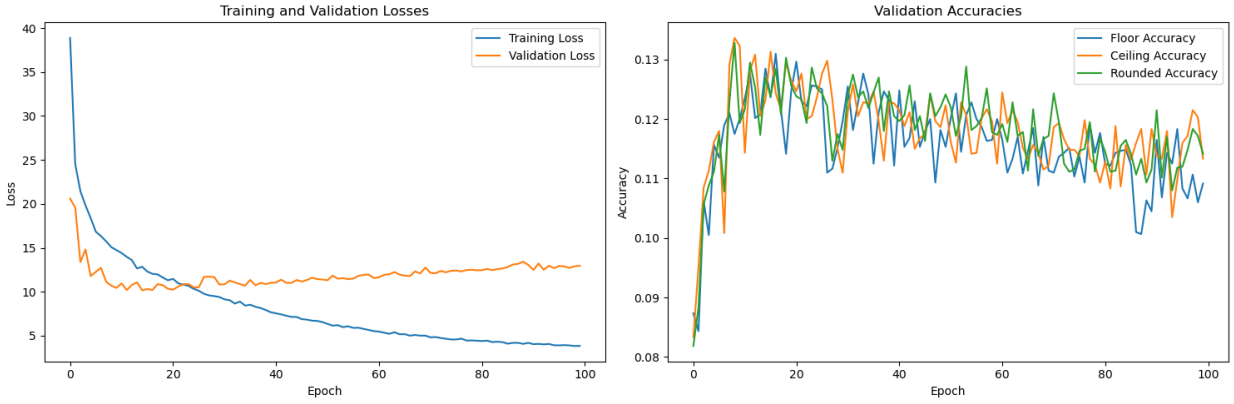

Figure 1: Training and validation history

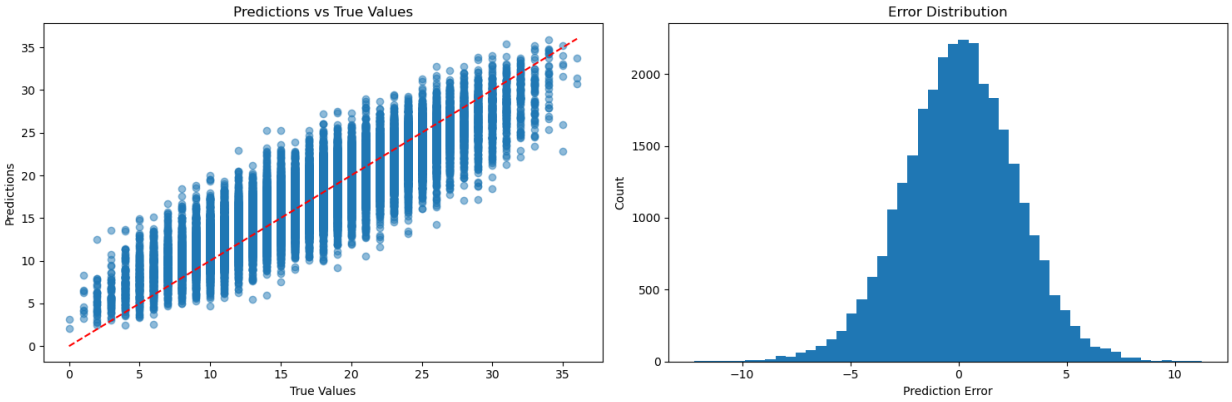

Figure 2: Inference results