**Homework 2**

Step 1: Data

1. How many data samples are included in the dataset?

Ans. The MNIST dataset contains 70,000 data samples.

1. Which problem will this dataset try to address?

Ans. The problem this dataset addresses is handwritten digit classification, where you need to classify images of handwritten digits (0-9) into their respective classes.

1. What is the minimum value and the maximum value in the dataset?

Ans. The minimum value in the dataset is 0 (black pixel) and the maximum value is 255 (white pixel) since it's a grayscale image.

1. What is the dimension of each data sample?

Ans. Each data sample is a 28x28 pixel image.

1. Does the dataset have any missing information? E.g., missing features.

Ans. No

1. What is the label of this dataset?

Ans. The labels of this dataset correspond to the digits 0 through 9.

1. How many percent of data will you use for training, validation and testing?

Ans. I am planning to split the data into 70% for training, 15% for validation, and 15% for testing.

8)What kind of data pre-processing will you use for your training dataset?

Ans. For pre-processing, I converted the datatype to float32, normalize pixel values to [0, 1], resized images to 64\*64 and converted RGB channels by duplicating the single channel.

Step 2: Model

|  |  |
| --- | --- |
| Model | Accuracy (best model) |
| DNN | 0.9795 |
| ConvNet | 0.9887 |
| ResNet | 0.9933 |

Step 3: Objective

I’ve used cross-entropy as the loss function to train my models.

Step 4: Optimization

I’ve used Adam optimizer to optimize my model.

Step 5: Model Selection:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | 0.1 | 0.01 | 0.001 | 0.0001 |
| DNN | 0.7154 | 0.9559 | 0.9795 | 0.9725 |
| ConvNet | 0.1135 | 0.9698 | 0.9882 | 0.9887 |
| ResNet | 0.4527 | 0.9159 | 0.9933 | 0.9928 |

Step 6: Model Performance

All the requirements are there in the screenshots folder which must be submitted with this docx file.