

NATIONAL UNIVERSITY OF MODERN LANGUAGES
ISLAMABAD



Artificial Neural Network (LAB)

Assignment: 03

Submitted to
Ms. Hina Ali

Submitted By
Junaid Asif
(BSAI-144)

Submission Date: November 28th , 2024

Question: You are provided with code to classify images from the MNIST dataset using PyTorch. Your task is to make the following modifications to the model and record the results of each update separately:

1. **Experiment with fewer epochs:** Reduce the number of training epochs to 5 and record the results.
 2. **Modify the convolutional layers:**
 - Decrease the number of filters in the first and second convolutional layers separately and observe the impact.
 - Reduce the size of filters in the first and second convolutional layers separately and record the changes.
 3. **Experiment with different activation functions:** Replace the current activation function and try at least three different activation functions. Record the results for each.
 4. Analyze how these changes affect the overall accuracy of the model.
- Create a table to summarize the results for each parameter change separately.

Answer:

1. Experiment with fewer epochs (reduce epochs to 5)

With 10 Epochs:

- That means this model is more likely to learn its weights to minimize its training loss.
- The loss will converge more assuming that the learning rate was appropriate and that the model wasn't overfitting.

With 5 Epochs:

- Training loss may be higher since the model has less passes through the dataset.
- In very complex datasets or large models, the model might have not converged at all.

2. Modify the convolutional layers (reduced filters and filter sizes)

Configuration	Conv1 Filter size	Conv2 Filter size	Training loss	Test accuracy	Comments
Original	3x3	3x3	Low	High	Baseline setup
Reduced Conv1	2x2	3x3	Moderate	Moderate	Struggles to capture early features.
Reduced Conv2	3x3	2x2	Moderate	Moderate	Struggles to capture complex features

3. Experiment with different activation functions (ReLU, Sigmoid, Softmax, TanH)

Activation Functions	Accuracy	Analysis
ReLU	98.87%	Performs well, avoid vanishing gradients.
Sigmoid	98.44%	Lead to vanishing gradients. Training might me slower.
Softmax	61.04%	It is better suited for output layers.
Tanh	98.65%	Like sigmoid.

4. Result Summary table

Experiments	Original value	New value	Accuracy	Observations
Epochs	10	5	98.82%	Slightly lower accuracy due to reduced training time.
Conv. Layer 1 (no. of filters)	32	16	98.43%	Accuracy slightly drops by reducing filters.
Conv. Layer 2 (no. of filters)	64	32	98.43%	Accuracy slightly drops by reducing filters.
Conv. Layer 1 (filter size)	3x3	1. 3x3 2. 2x2	Moderate 98.44%	Struggles to capture early features.
Conv. Layer 2 (filter size)	3x3	1. 2x2 2. 3x3	Moderate 98.44%	Struggles to capture complex features