

**Sardar Vallabhbhai National Institute of Technology  
Surat-395007**

**Department of Artificial Intelligence  
Deep Learning (AI302)**

**Lab Practical– 2**

Problem Statement: Handwritten Digit Recognition

**Objective:**

Your goal is to build and refine a Convolutional Neural Network (CNN) to classify handwritten digits from the MNIST dataset. While the base architecture is provided, the model currently performs poorly due to suboptimal choices in activation functions and optimization algorithms.

**The Base Architecture for CNN:**

- Input Layer: (28,28 )grayscale images.
- Conv2D Layer1: 32 filters,  $3 \times 3$  kernel, Activation Function: ReLU
- Conv2D Layer2: 64 filters,  $3 \times 3$  kernel, Activation Function: ReLU
- Max Pooling Layer:  $2 \times 2$  kernel
- Dropout:Dropout rate: **0.25**
- Dense Layer: 128 neurons, ReLU, (Fully connected Layer)
- Output Layer: 10 neurons (Softmax).

**The Base Architecture for MLP:**

- Flatten(784)
- Dense(256)
- BatchNormalization
- ReLU
- Dense(128)
- BatchNormalization
- ReLU
- Dense(10, Softmax)

You must experiment with different configurations

Model	FC layer	Optimizer	Epochs	Accuracy
CNN-1	128	Adam	10	?
MLP-1	512-256-128	SGD	20	?
MLP-2	256	Adam	15	?

## **Task 1: The Activation Function Challenge**

Compare the training loss and accuracy curves when using the following:

- Sigmoid: Observe if the model suffers from "vanishing gradients" or slow start.
- Tanh: Compare its speed to Sigmoid.
- ReLU: Document why this usually leads to faster convergence.

## **Task 2: The Optimizer Showdown**

Once you have selected the best activation function, keep it constant and switch between these optimizers:

- SGD (Stochastic Gradient Descent): Observe the stability of the loss.
- SGD with Momentum: Note how it handles "bumps" in the loss landscape.
- Adam: Observe how quickly it reaches high accuracy compared to basic SGD.

## **Task 3:**

**Run below specific scenarios to observe the contrast:**

- WITHOUT Batch normalization and drop out layer
- Without BN, drop out layer=0.1
- With BN, drop out layer=0.25

**You must submit a notebook containing:**

- Comparison Table: A table showing "Activation + Optimizer" combinations and their final Test Accuracy.
- Visualization: A plot showing the Loss Curves for at least three different experiments(training and testing accuracy plots).

**Example Table :**

Experiment	Activation	Optimizer	Epochs	Final Accuracy
1	Sigmoid	SGD	10	?
2	ReLU	SGD	10	?
3	ReLU	Adam	10	?