

Training with sigmoid activation :-

Epoch 1/5 - loss: 2.0941

Epoch 2/5 - loss: 1.8724

Epoch 3/5 - loss: 1.7123

Epoch 4/5 - loss: 1.6588

Epoch 5/5 - loss: 1.6434.



Training with tanh activation :-

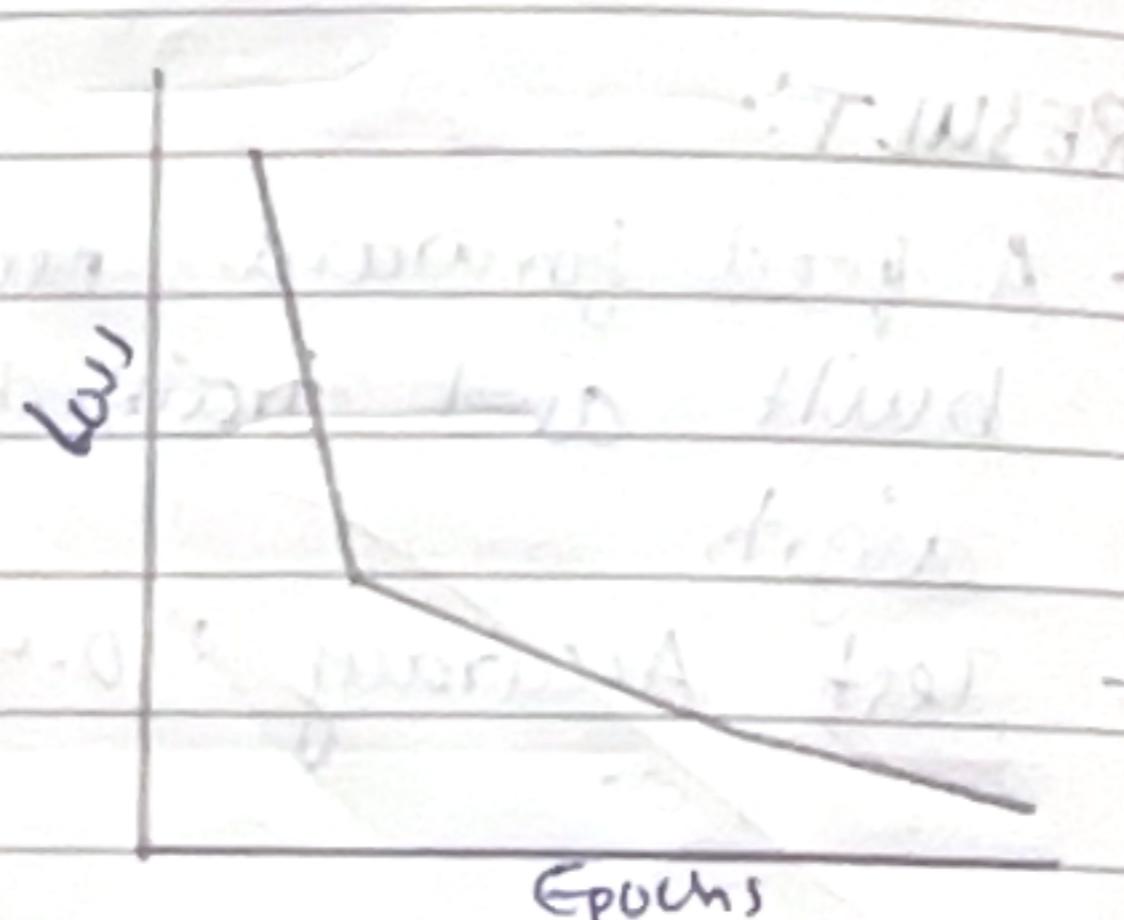
Epoch 1/5 - loss: 1.7028

Epoch 2/5 - loss: 1.5682

Epoch 3/5 - loss: 1.5511

Epoch 4/5 - loss: 1.5409

Epoch 5/5 - loss: 1.5326.



Training with ReLU activation :-

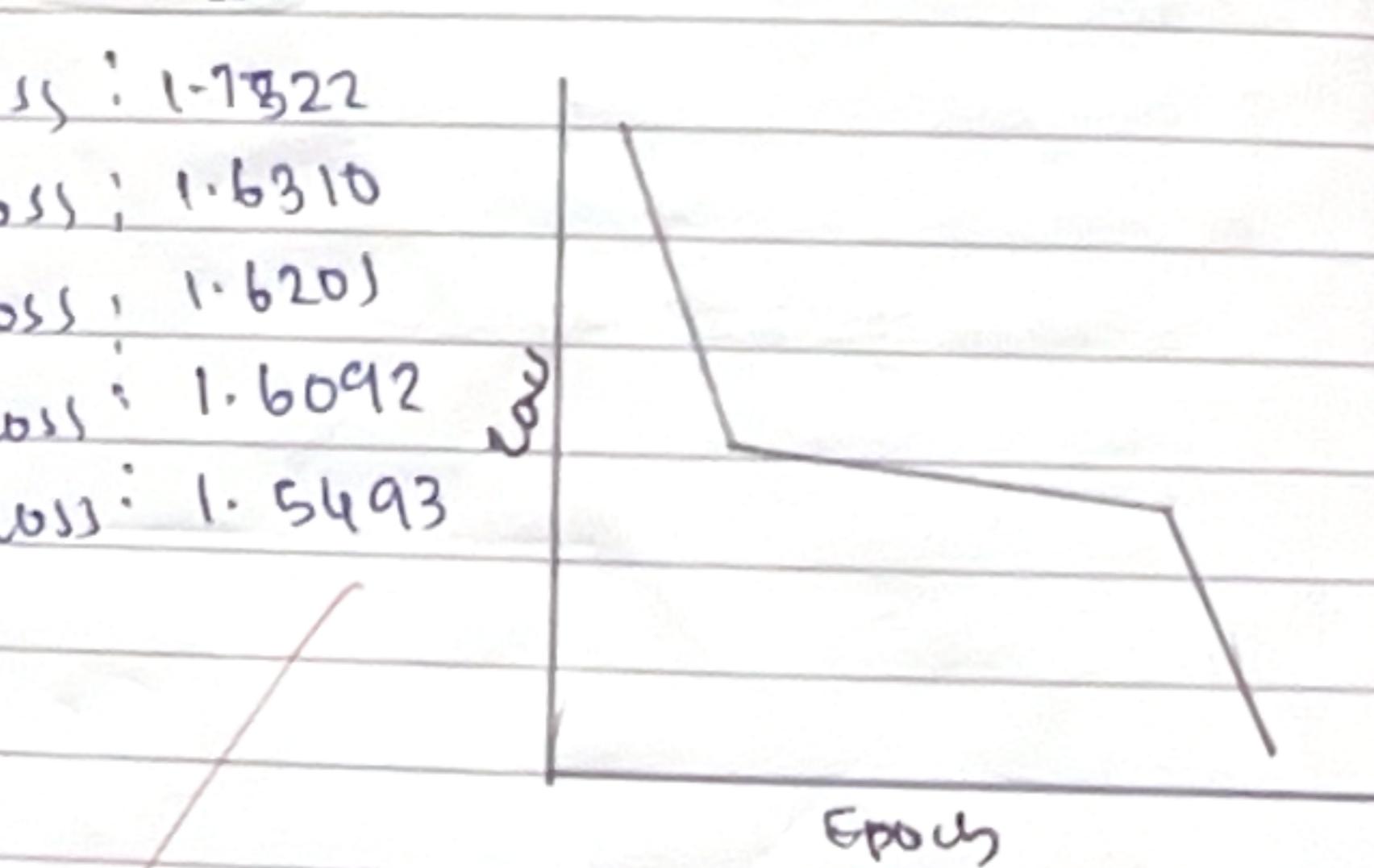
Epoch 1/5 - loss: 1.7322

Epoch 2/5 - loss: 1.6310

Epoch 3/5 - loss: 1.6201

Epoch 4/5 - loss: 1.6092

Epoch 5/5 - loss: 1.5493



01/09/25

LAB-5

Study Of Activation Functions

And Their Role.

AIM:-

To study the effect of different activation functions [Sigmoid, Tanh, ReLU, Leaky ReLU, Softmax] on a neural network trained for MNIST digit classification.

OBJECTIVES:-

- Understand the role of activation functions in hidden and output layers
- Implement a neural network for MNIST digit classification
- Train the network with different activation functions
- Compare loss and accuracy for each activation function.

PSEUDO CODE:-

1. Import required libraries
2. Load MNIST dataset with normalization
3. Define Data loader for training and testing
4. Create Neural Network class:
 - a. Input layer (784)
 - b. Hidden layer (128 neurons)
 - c. Output layer (10 neurons)
5. Define training functions.

Training with leaky-relu activation.

Epoch 1/5 - loss: 1.7431

Epoch 2/5 - loss: 1.6241

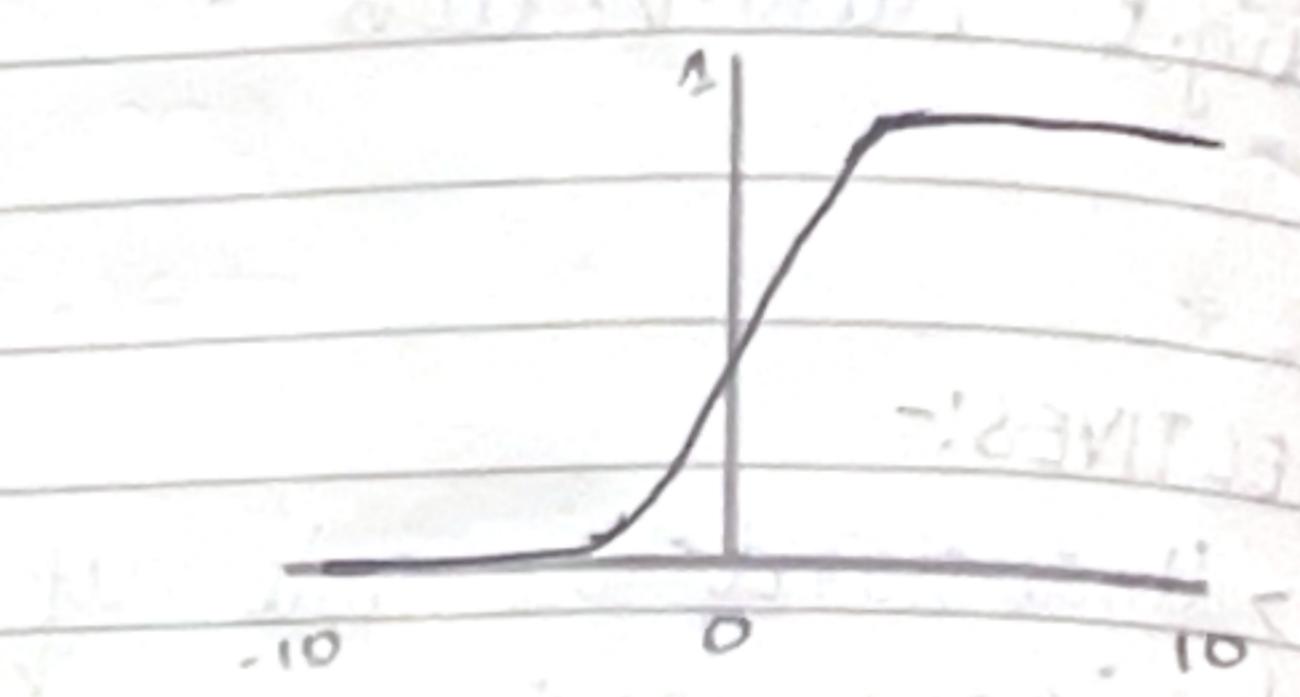
Epoch 3/5 - loss: 1.5623

Epoch 4/5 - loss: 1.5477

Epoch 5/5 - loss: 1.5369

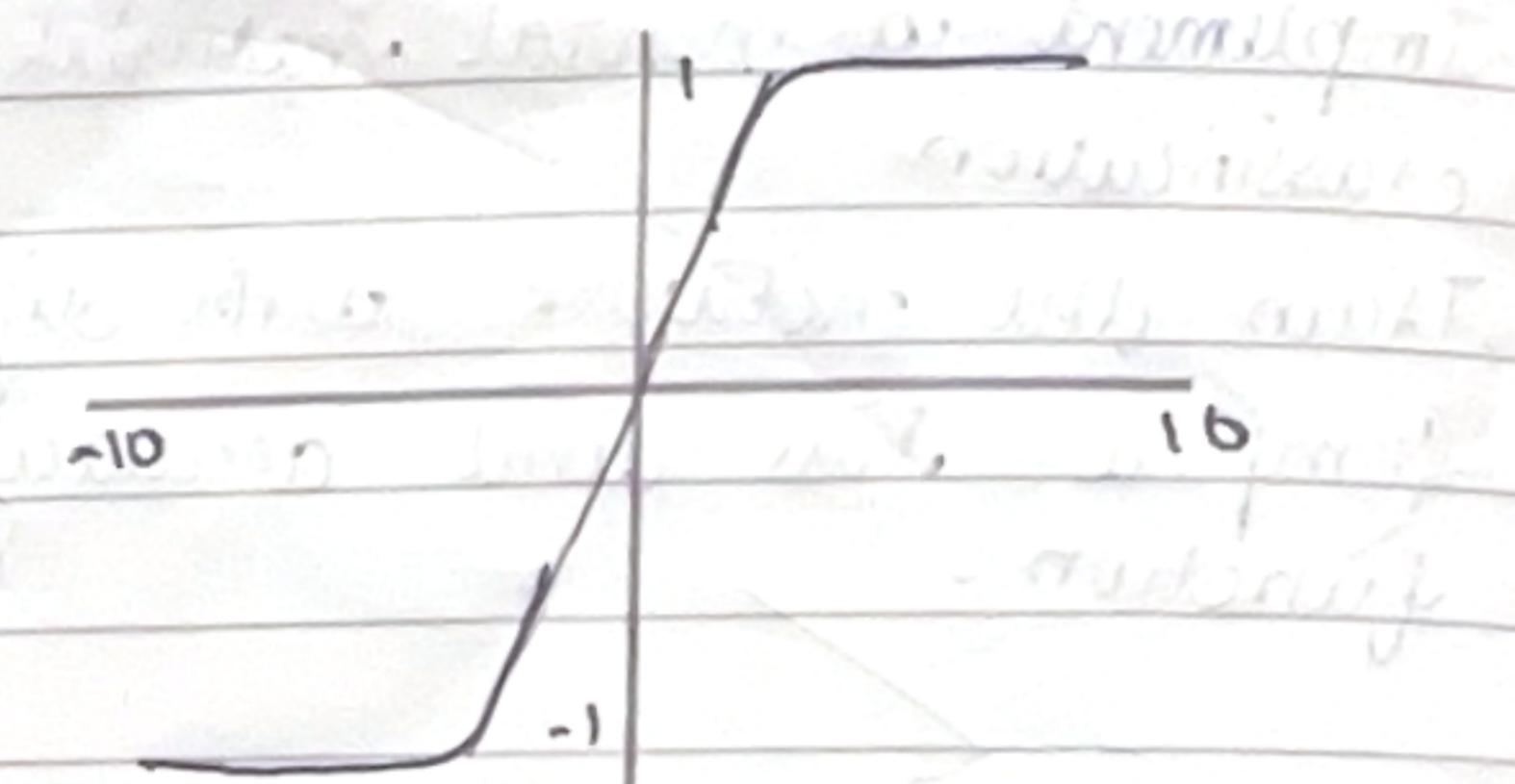
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



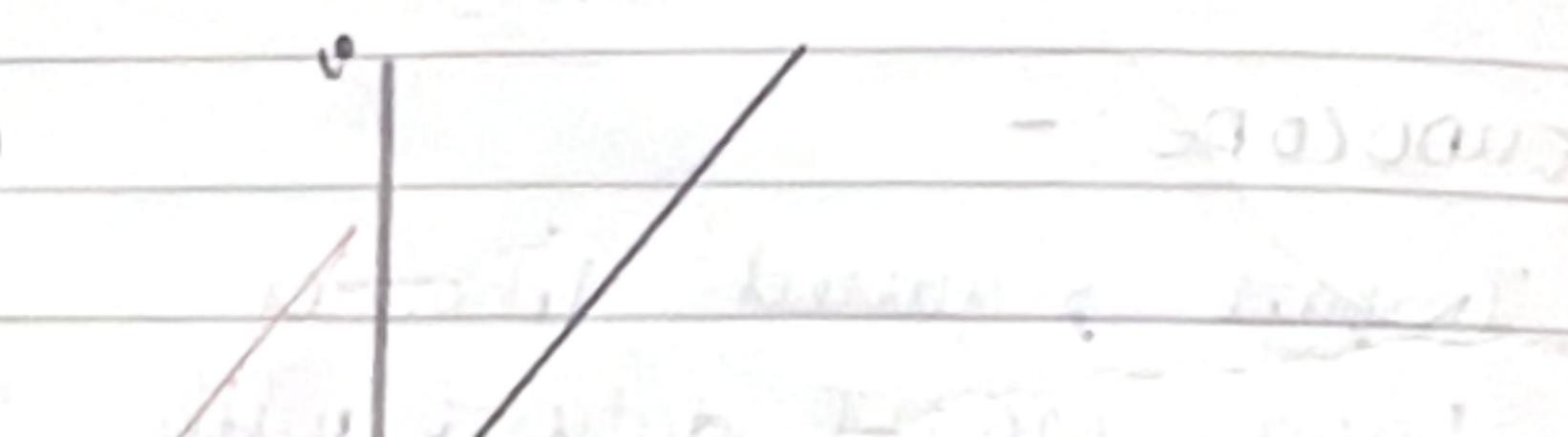
tanh

$$\tanh(x)$$



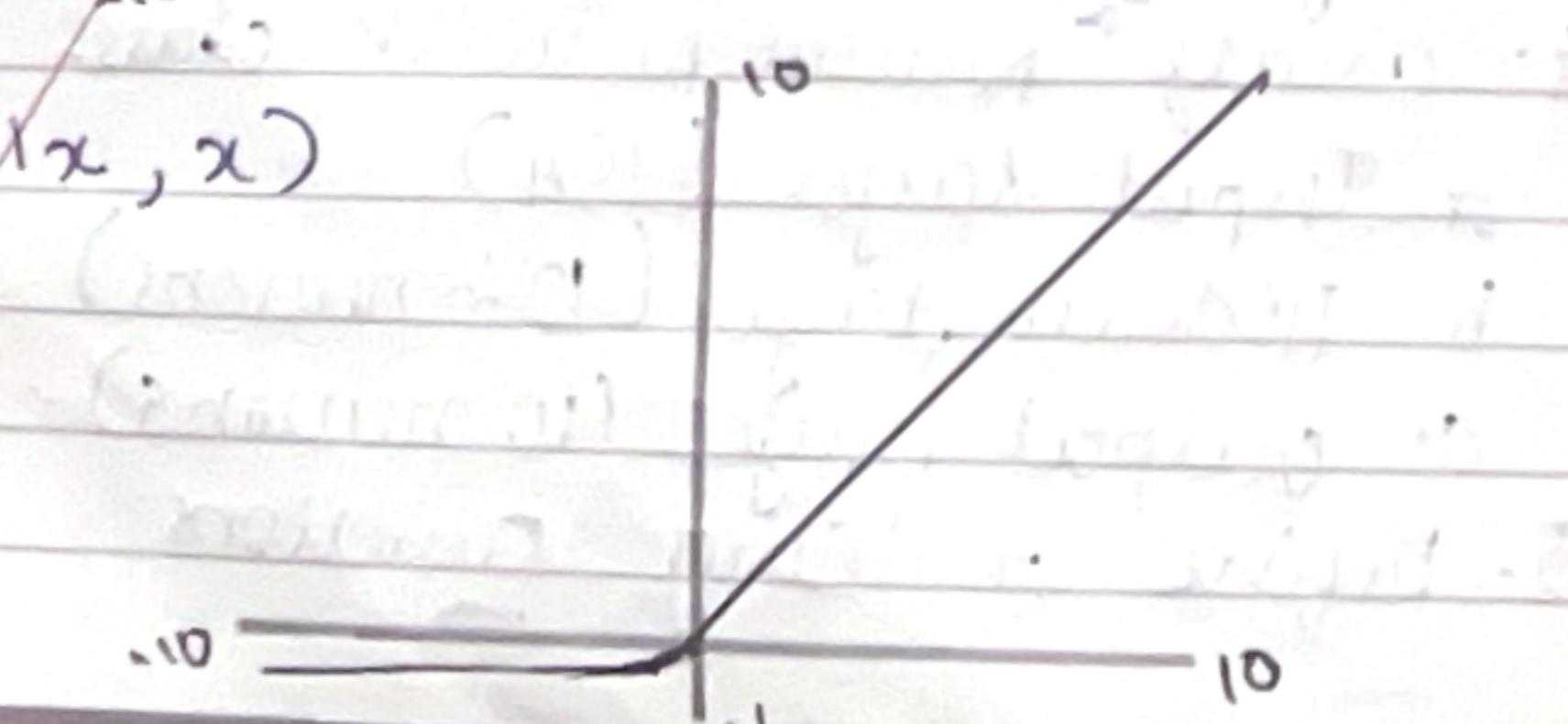
ReLU

$$\max(0, x)$$



Leak ReLU

$$\max(0, \alpha x, x)$$



a. Use crossEntropyLoss

b. Initialize optimizer (SGD)

c. Train for N epochs

d. Track loss and accuracy

e. Loop through Activations:

[sigmoid, tanh, relu, leaky-relu, softmax]

f. Train network

g. Plot training loss

h. Compute test accuracy

i. Compare activations

RESULT:-

- ReLU and Leaky ReLU are best due to fast convergence
- Choosing proper activation function is crucial for accuracy and training efficiency.