# Deep Neural Network

Simulation

# Problem Description

- Handwritten image recognition is a challenging tasks for machines to detect. Hence, in this experiment deep neural networks are leveraged for classifying handwritten numbers.
- The dataset leveraged for this task is MNIST (Modified National Institute of Standards and Technology)

# Activation Functions Used

- ReLU
- Swish (Recent)
- tanh
- Sigmoid

# ReLU

- The ReLU function works by giving the value x itself if its positive or zero when it is applied pointwise.

$$x = max(0, x)$$

# Swish

- This is a relatively recent activation function founded in 2017.
- It resembles the ReLU activation function but smooth for values below zero. But offers smoothness
- It has shown promise for various deep learning related tasks.

$$\text{Swish}(x) = x \cdot \sigma(\beta x)$$

# tanh

- The tanh or hyperbolic tangent activation function is another activation function that can be used to model non linearity to the function composition used in a neural network
- It varies between -1 to 1

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

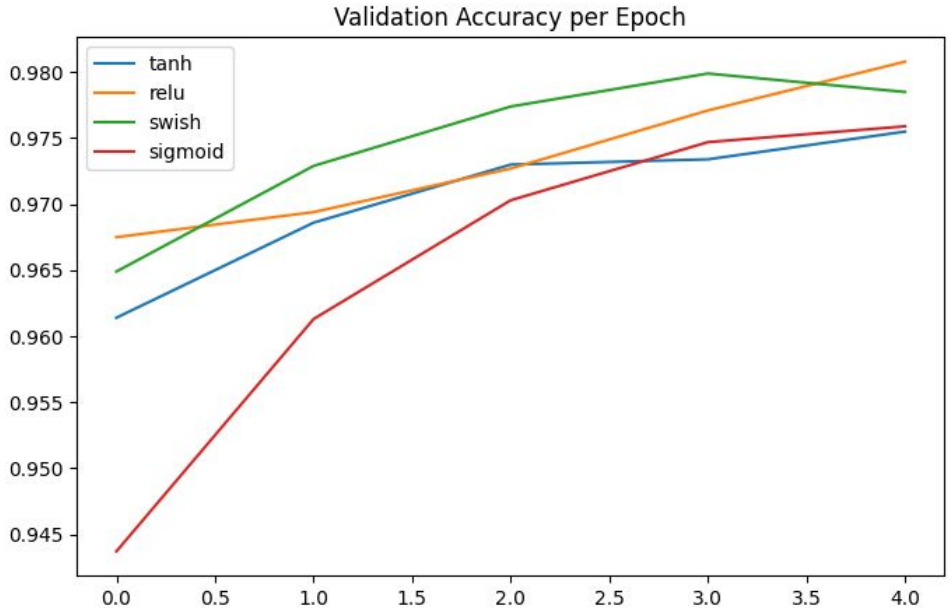
# sigmoid

- The sigmoid activation function has some similarities with the tanh activation function.
- But the sigmoid function has outputs between 0 to 1

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

# Comparison Results

- The following activation functions were compared as an activation function. With 3-ary neural networks.
- The Output layer was softmax

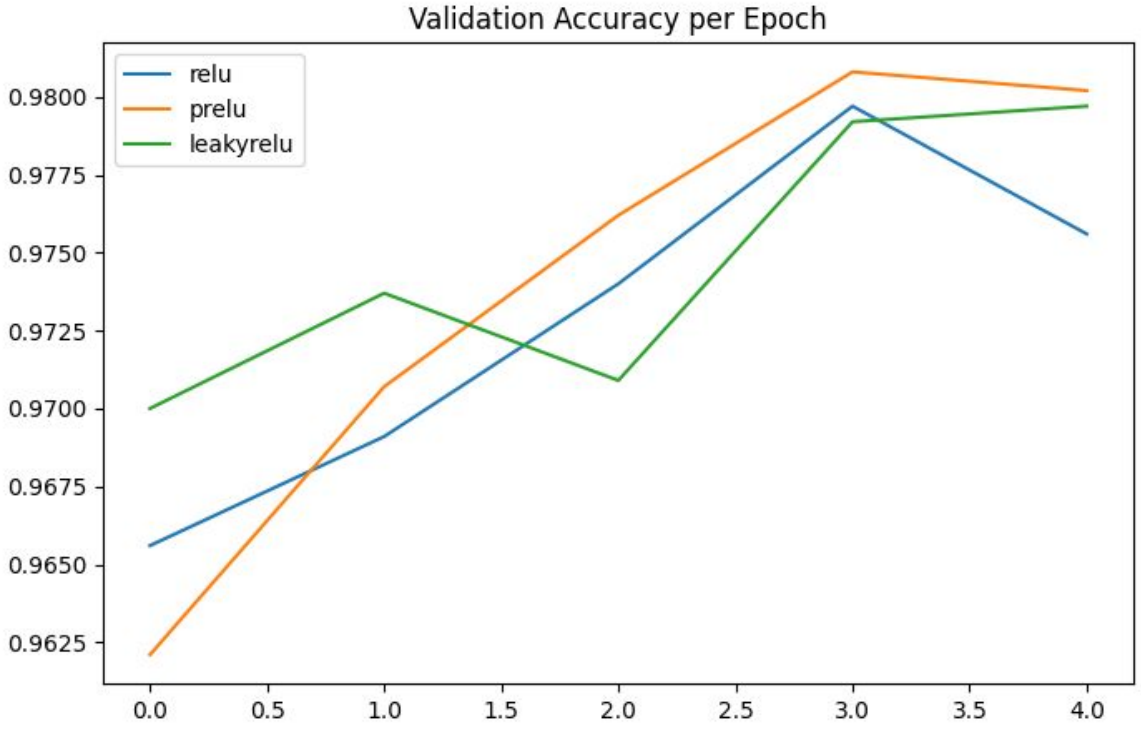

Validation Accuracy per Epoch
tanh
relu
swish
sigmoid
0.980
0.975
0.970
0.965
0.960
0.955
0.950
0.945
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

# Leaky ReLU

- The leaky ReLU problem is designed to address the dying relu problem.
- This is done by assigning low values multiplied by $\alpha$

# PReLU

- The parametric ReLU changes the alpha as a learnable parameter.
- It changes as the other parameters due to backprop.


Validation Accuracy per Epoch
relu
prelu
leakyrelu
0.9800
0.9775
0.9750
0.9725
0.9700
0.9675
0.9650
0.9625
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

# Depth Evaluations