Activation Functions

- Introduce non-linearity to neural networks.
- Allow networks to learn complex patterns in data.

Sigmoid Activation Function

- Takes any real number as input and outputs a number between 0 and 1.
- Often used in output layers for classification problems.
- Formula: σ(x) = 1 / (1 + e^(-x))

ReLU Activation Function

- Takes any real number as input and outputs the maximum of 0 and the input value.
- Often used in hidden layers.
- Formula: ReLU(x) = max(0, x)

ReLU vs Sigmoid

- ReLU is more popular than sigmoid because it avoids the vanishing gradient problem.
- Sigmoid is still useful for classification problems because it outputs values between 0 and 1 (probabilities).