

# Activation Function

| Activation Function | Equation  | Range   | Use Case   |
|---------------------|---|---------|--|
| Sigmoid             | $f(x) = \frac{1}{1 + e^{-x}}$                           | (0, 1)  | Binary classification, hidden layers in small networks.    |
| Tanh                | $f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$   | (-1, 1) | Hidden layers, encourages zero-centered outputs.           |
| ReLU                | $f(x) = \max(0, x)$                                     | [0, ∞)  | Default for most hidden layers, efficient computation.     |
| Leaky ReLU          | $f(x) = \max(0.01x, x)$                                 | (-∞, ∞) | Avoids dying ReLU problem, suitable for hidden layers.     |
| Softmax             | $f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$               | (0, 1)  | Output layers in multi-class classification problems.      |
| Swish               | $f(x) = x \cdot \text{sigmoid}(x)$                      | (-∞, ∞) | Recent innovation, smooth activation for better gradients. |
| GELU                | $f(x) = 0.5x(1 + \tanh(\sqrt{2/\pi}(x + 0.044715x^3)))$ | (-∞, ∞) | High-performing activation in transformers.                |

# Activation Function

Examples:

- Regression — Linear Activation
- Binary Classification— Sigmoid/Logistic
- Multiclass Classification—Softmax
- Multilabel Classification—Sigmoid

⋮

## Neural Network Activation Functions

