Mathematics of Artificial Neural Networks

1. Neuron Function (Perceptron)

Each neuron computes a weighted sum of inputs and applies an activation function:

$$z = \sum(w_i * x_i) + b$$

$$a = \phi(z)$$

Where:

- x_i: input features
- w_i: weights
- b: bias
- $\phi(z)$: activation function

2. Layer Operations (Matrix Form)

For a layer of neurons:

$$Z = W \cdot X + B$$

$$A = \phi(Z)$$

Where:

- X: input matrix
- W: weights
- B: bias
- A: output

3. Loss Function

Measures prediction error:

- Regression (MSE): $L = (1/n) * \sum((y_i - \hat{y}_i)^2)$
- Classification (Cross Entropy): $L = -\sum(y_i * \log(\hat{y}_i))$

4. Backpropagation (Gradient Descent)

To minimize loss, update weights via:

Mathematics of Artificial Neural Networks

w := w - eta * dL/dw

Use the chain rule to compute gradients.

5. Activation Functions

- Sigmoid: 1 / (1 + exp(-z))
- ReLU: max(0, z)
- Tanh: (exp(z) - exp(-z)) / (exp(z) + exp(-z))

6. Visual Diagram