

# Fluxo de Operações: Rede Neural Multicamadas (XOR)

## Documentação de Aprendizado

Este documento descreve as operações matemáticas em um único ciclo da função `train()`, considerando uma rede com 2 entradas, 2 neurônios ocultos e 1 saída.

### 1 Passo 1: Feedforward (A "Ida")

O objetivo é transformar a entrada em uma predição.

#### 1.1 1.1 Camada Oculta

A entrada  $I = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  atravessa a primeira matriz de pesos  $W_{ih}$  e soma-se o Bias  $B_{ih}$ .

$$Z_h = \begin{bmatrix} 0.5 & 0.2 \\ 0.1 & -0.4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.2 \end{bmatrix}$$

Aplicando a função de ativação Sigmoide:

$$H = \sigma(Z_h) \approx \begin{bmatrix} 0.64 \\ 0.55 \end{bmatrix}$$

#### 1.2 1.2 Camada de Saída

O resultado de  $H$  torna-se a entrada para a camada final.

$$Z_o = [0.7 \quad -0.3] \cdot \begin{bmatrix} 0.64 \\ 0.55 \end{bmatrix} + [0.1] = [0.38]$$

\*\*Saída Final (Predição):\*\*

$$O = \sigma(0.38) \approx 0.59$$

### 2 Passo 2: Backpropagation - Saída (A "Volta")

A rede compara a saída com o alvo ( $T = 1$ ) para ajustar os pesos  $W_{ho}$ .

- **Erro da Saída:**  $E_o = T - O = 1 - 0.59 = 0.41$
- **Gradiente:**  $G_o = \sigma'(O) \cdot E_o \cdot LR = 0.24 \cdot 0.41 \cdot 0.1 \approx 0.0098$
- **Ajuste de Pesos ( $\Delta W_{ho}$ ):**  $G_o \cdot H^T = [0.0098] \cdot [0.64 \quad 0.55] = [0.006 \quad 0.005]$

**Nova Matriz  $W_{ho}$  (Soma e Gravação):**

$$W_{ho(novo)} = [0.7 \quad -0.3] + [0.006 \quad 0.005] = [0.706 \quad -0.295]$$

### 3 Passo 3: Backpropagation - Entrada (O Erro Invisível)

O erro é projetado de volta para a camada oculta.

1. **Erro da Oculta ( $E_h$ ):** Usa-se a Transposta dos pesos antigos.

$$E_h = W_{ho}^T \cdot E_o = \begin{bmatrix} 0.7 \\ -0.3 \end{bmatrix} \cdot [0.41] = \begin{bmatrix} 0.28 \\ -0.12 \end{bmatrix}$$

2. **Gradiente da Oculta ( $G_h$ ):** Aplicação do Produto de Hadamard ( $\odot$ ).

$$G_h = \sigma'(H) \odot E_h \cdot LR = \begin{bmatrix} 0.23 \\ 0.24 \end{bmatrix} \odot \begin{bmatrix} 0.28 \\ -0.12 \end{bmatrix} \cdot 0.1 = \begin{bmatrix} 0.006 \\ -0.003 \end{bmatrix}$$

3. **Ajuste Final ( $\Delta W_{ih}$ ):** Multiplica-se pela Transposta da entrada.

$$\Delta W_{ih} = G_h \cdot I^T = \begin{bmatrix} 0.006 \\ -0.003 \end{bmatrix} \cdot [1 \quad 0] = \begin{bmatrix} 0.006 & 0 \\ -0.003 & 0 \end{bmatrix}$$

## 4 Conclusão

Os novos pesos calculados ( $W_{ho}$  e  $W_{ih}$ ) substituem os antigos na memória da rede. No próximo ciclo de treinamento, a predição será mais próxima do alvo 1.