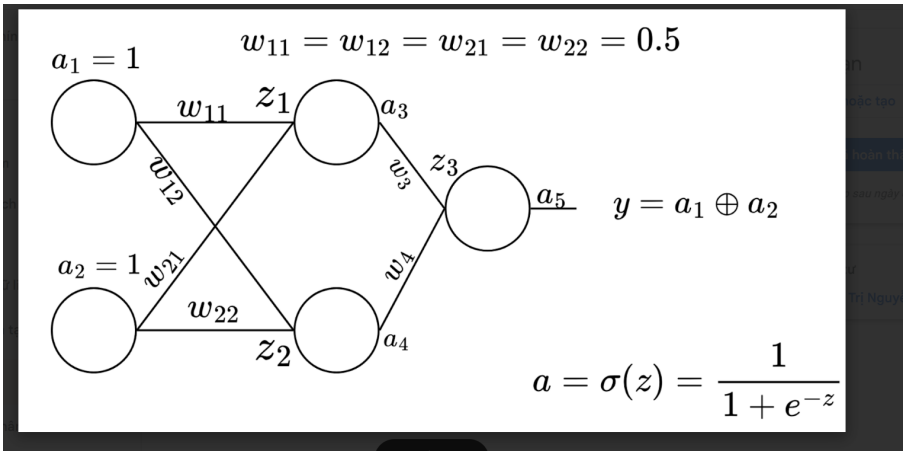
# Deep Learning

Phạm Phước Bảo Tín

March 2025

## Tính feedforward và backpropagation cho neural network như hình dưới




Hình 1: Cấu trúc neral network

Ta có:

- $w_{11} = w_{12} = w_{21} = w_{22} = w_3 = w_4 = 0.5$
- $a_1 = a_2 = 1$
- $a = \sigma(z) = \frac{1}{1+e^{-z}}$
- $y = a_1 \oplus a_2 = 0$
- $\eta = 1.0$
- $L = -[y \cdot log(a) + (1-y) \cdot log(1-a)]$

### Feedforward

- Quá trình tính toán tại lớp ẩn:

$$\begin{aligned} z_1 &= w_{11} \cdot a_1 + w_{21} \cdot a_2 = 1 \\ a_3 &= \sigma(z_1) = 0.731 \\ z_2 &= w_{12} \cdot a_1 + w_{22} \cdot a_2 \\ a_4 &= \sigma(z_2) = 0.731 \end{aligned} \tag{1}$$

- Quá trình tính toán tại lớp đầu ra:

$$\begin{aligned} z_3 &= w_3 \cdot a_3 + w_4 \cdot a_4 = 0.731 \\ a_5 &= \sigma(z_3) = 0.675 \end{aligned} \tag{2}$$

## Backpropagation

Hàm mất mát: $L = -[y \cdot log(a_5) + (1-0) \cdot log(1-a_5)] = 1.124$
Cập nhật trọng số: $w_3, w_4, w_{11}, w_{12}, w_{21}, w_{22}$ sử dụng Gradient Descent: $w = w - \eta \cdot \frac{\partial L}{\partial w}$

- Áp dụng quy tắc chuỗi (chain rule) trong việc đạo hàm

$$\frac{\partial L}{\partial a_5} = \frac{a_5 - y}{a_5(1-a_5)} \tag{3}$$

$$\frac{\partial \sigma(z)}{\partial z} = \sigma(z)(1-\sigma(z)) \tag{4}$$

- Áp dụng chain rule cho $w_3, w_4$:

$$\begin{aligned}
\frac{\partial L}{\partial w_3} &= \frac{\partial L}{\partial z_3} \cdot \frac{\partial z_3}{\partial w_3} \\
\frac{\partial L}{\partial w_3} &= \frac{\partial L}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_3} \cdot \frac{\partial z_3}{\partial w_3} \\
\frac{\partial L}{\partial w_3} &= \frac{a_5 - y}{a_5 \cdot (1-a_5)} \cdot a_5 \cdot (1-a_5) \cdot a_3 \\
\frac{\partial L}{\partial w_3} &= (a_5 - y) \cdot a_3
\end{aligned} \tag{5}$$

$$\begin{aligned}
\frac{\partial L}{\partial w_4} &= \frac{\partial L}{\partial z_3} \cdot \frac{\partial z_3}{\partial w_4} \\
\frac{\partial L}{\partial w_3} &= \frac{\partial L}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_3} \cdot \frac{\partial z_3}{\partial w_4} \\
\frac{\partial L}{\partial w_4} &= \frac{a_5 - y}{a_5 \cdot (1-a_5)} \cdot a_5 \cdot (1-a_5) \cdot a_4 \\
\frac{\partial L}{\partial w_4} &= (a_5 - y) \cdot a_4
\end{aligned} \tag{6}$$

- Áp dụng chain rule cho $w_{11}, w_{21}, w_{12}, w_{22}$ :

$$\begin{aligned}
\frac{\partial L}{\partial w_{11}} &= \frac{\partial L}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{11}} \\
\frac{\partial L}{\partial w_{11}} &= \frac{\partial L}{\partial a_3} \cdot \frac{\partial a_3}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{11}} \\
\frac{\partial L}{\partial w_{11}} &= \frac{\partial L}{\partial a_5} \cdot \frac{\partial a_5}{\partial z_3} \cdot \frac{\partial z_3}{\partial a_3} \cdot \frac{\partial a_3}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_{11}} \\
\frac{\partial L}{\partial w_{11}} &= \frac{a_5 - y}{a_5 \cdot (1-a_5)} \cdot a_5(1-a_5) \cdot w_3 \cdot a_3 \cdot (a_3 - 1) \cdot a_1 \\
\frac{\partial L}{\partial w_{11}} &= (a_5 - y) \cdot w_3 \cdot a_3 \cdot (1-a_3) \cdot a_1
\end{aligned} \tag{7}$$

$$\frac{\partial L}{\partial w_{21}} = (a_5 - y) \cdot w_3 \cdot a_3 \cdot (1-a_3) \cdot a_2 \tag{8}$$

$$\frac{\partial L}{\partial w_{12}} = (a_5 - y) \cdot w_4 \cdot a_4 \cdot (1-a_4) \cdot a_1 \tag{9}$$

$$\frac{\partial L}{\partial w_{22}} = (a_5 - y) \cdot w_4 \cdot a_4 \cdot (1-a_4) \cdot a_2 \tag{10}$$

## Quá trình tối ưu

Cập nhật trọng số:

$$
\begin{aligned}
w_3^{new} &= w_3 - \eta \cdot \frac{\partial L}{\partial w_3} = w_3 - \eta \cdot (a_5 - y) \cdot a_3 = 0.5 - 1 \cdot 0.675 \cdot 0.731 = 0.007 \\
w_4^{new} &= w_4 - \eta \cdot \frac{\partial L}{\partial w_4} = w_3 - \eta \cdot (a_5 - y) \cdot a_4 = 0.5 - 1 \cdot 0.675 \cdot 0.731 = 0.007 \\
w_{11}^{new} &= w_{11} - \eta \cdot \frac{\partial L}{\partial w_{11}} = w_{11} - \eta \cdot (a_5 - y) \cdot w_3 \cdot a_3 \cdot (1 - a_3) \cdot a_1 = 0.5 - 1 \cdot 0.675 \cdot 0.5 \cdot 0.731 \cdot (1 - 0.731) \cdot 1 = 0.434 \\
w_{21}^{new} &= w_{21} - \eta \cdot \frac{\partial L}{\partial w_{21}} = w_{21} - \eta \cdot (a_5 - y) \cdot w_3 \cdot a_3 \cdot (1 - a_3) \cdot a_2 = 0.5 - 1 \cdot 0.675 \cdot 0.5 \cdot 0.731 \cdot (1 - 0.731) \cdot 1 = 0.434 \\
w_{12}^{new} &= w_{12} - \eta \cdot \frac{\partial L}{\partial w_{12}} = w_{12} - \eta \cdot (a_5 - y) \cdot w_4 \cdot a_4 \cdot (1 - a_4) \cdot a_1 = 0.5 - 1 \cdot 0.675 \cdot 0.5 \cdot 0.731 \cdot (1 - 0.731) \cdot 1 = 0.434 \\
w_{22}^{new} &= w_{22} - \eta \cdot \frac{\partial L}{\partial w_{22}} = w_{22} - \eta \cdot (a_5 - y) \cdot w_4 \cdot a_4 \cdot (1 - a_4) \cdot a_2 = 0.5 - 1 \cdot 0.675 \cdot 0.5 \cdot 0.731 \cdot (1 - 0.731) \cdot 1 = 0.434
\end{aligned}
\tag{11}
$$

Tính toán lại Feedforward:

$$
\begin{aligned}
z_1 &= w_{11} \cdot a_1 + w_{21} \cdot a_2 = 0.434 \cdot 1 + 0.434 \cdot 1 = 0.868 \\
z_2 &= w_{12} \cdot a_1 + w_{22} \cdot a_2 = 0.434 \cdot 1 + 0.434 \cdot 1 = 0.868 \\
a_3 &= \sigma(z_1) = 0.704 \\
a_4 &= \sigma(z_2) = 0.704 \\
z_3 &= w_3 \cdot a_3 + w_4 \cdot a_4 = 0.07 \cdot 0.704 + 0.07 \cdot 0.704 = 0.098 \\
a_5 &= \sigma(z_3) = 0.524 \\
L &= 0.742
\end{aligned}
\tag{12}
$$

Cập nhật lại trọng số Backpropagation:

$$
\begin{aligned}
w_3 &= -0.361 \\
w_4 &= -0.361 \\
w_{11} &= 0.433 \\
w_{21} &= 0.433 \\
w_{12} &= 0.433 \\
w_{22} &= 0.433
\end{aligned}
\tag{13}
$$

Tính toán lại Feedforward:

$$
\begin{aligned}
z_1 &= 0.866 \\
z_2 &= 0.866 \\
a_3 &= 0.703 \\
a_4 &= 0.703 \\
z_3 &= -0.507 \\
a_5 &= 0.375 \\
L &= 0.470
\end{aligned}
\tag{14}
$$

Lặp lại quá trình trên cho đến khi hàm mất mát hội tụ.
**Code Python: Click here**