

# Giải thích Toán học về xấp xỉ hàm phi tuyến bằng Activation function trong Machine learning

LÊ VIỆT ĐỨC

November 2025

## 1 Định lý Xấp xỉ Toàn cầu (Đang tổng quát)

Phương trình toán học cơ bản của một mạng nơ-ron 1 lớp ẩn (với  $N$  nơ-ron) xấp xỉ hàm  $y(x)$  là:

$$y(x) = \sum_{i=1}^N c_i \cdot \sigma(a_i x + b_i) + d \quad (1)$$

Trong đó:

- $x$  là đầu vào (input).
- $y(x)$  là hàm xấp xỉ (output).
- $\sigma$  là hàm kích hoạt phi tuyến (ví dụ: ReLU, tanh).
- $a_i, b_i$  là trọng số và bias của nơ-ron ẩn thứ  $i$ .
- $c_i, d$  là trọng số và bias của lớp output.

Quá trình training là quá trình tìm các giá trị  $a_i, b_i, c_i, d$  tối ưu.

## 2 Ví dụ 1: Dùng ReLU (Tuyến tính từng khúc)

Mục tiêu là cộng các "khúc gãy" (hình chữ L) để tạo ra một hình phức tạp hơn.

### 2.1 Định nghĩa hàm ReLU

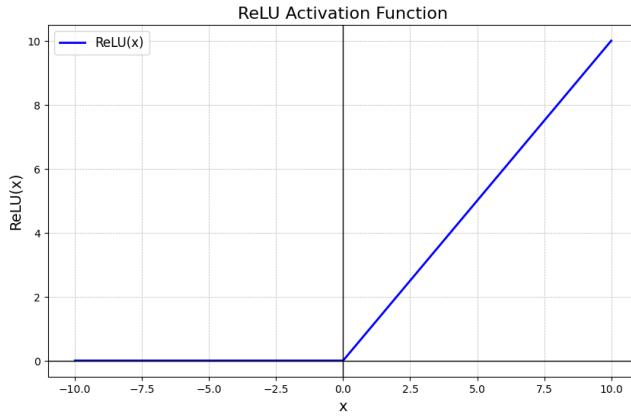
Hàm kích hoạt ReLU (Rectified Linear Unit) được định nghĩa là:

$$\text{ReLU}(z) = \max(0, z) \quad (2)$$

Hay viết dưới dạng:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

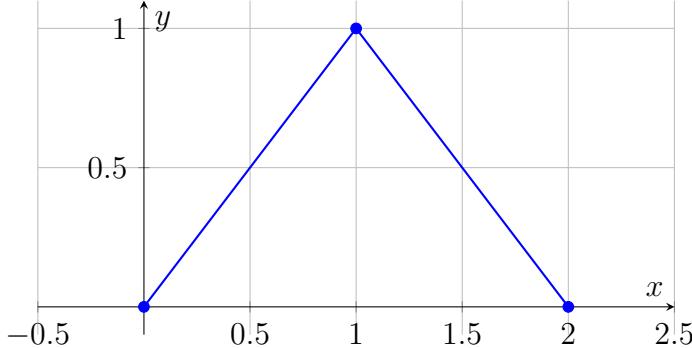
Với khoảng output:  $[0, \infty)$



Hình 1: Minh họa đồ thị ReLU

## 2.2 Bài toán: Tạo "hình tam giác"

Chúng ta muốn tạo một hàm  $y(x)$  bắt đầu từ 0, tăng tuyến tính lên 1 (tại  $x = 1$ ), và giảm tuyến tính về 0 (tại  $x = 2$ ).



Chúng ta dùng 3 nơ-ron ẩn với các tham số (giả sử đã học được) như sau:

$$\begin{aligned} f_1(x) &= \text{ReLU}(x) && (\text{Nơ-ron 1, với trọng số output } c_1 = 1) \\ f_2(x) &= \text{ReLU}(x - 1) && (\text{Nơ-ron 2, với trọng số output } c_2 = -2) \\ f_3(x) &= \text{ReLU}(x - 2) && (\text{Nơ-ron 3, với trọng số output } c_3 = 1) \end{aligned}$$

Hàm cuối cùng  $y(x)$  là tổng của chúng (giả sử  $d = 0$ ):

$$y(x) = 1 \cdot \text{ReLU}(x) - 2 \cdot \text{ReLU}(x - 1) + 1 \cdot \text{ReLU}(x - 2) \quad (3)$$

## 2.3 Phân tích từng bước (Từng khúc)

**1. Khi  $x < 0$  (ví dụ  $x = -1$ ):**

$$\begin{aligned} y(-1) &= 1 \cdot \max(0, -1) - 2 \cdot \max(0, -2) + 1 \cdot \max(0, -3) \\ &= 1 \cdot 0 - 2 \cdot 0 + 1 \cdot 0 \\ &= 0 \end{aligned}$$

**2. Khi  $0 \leq x < 1$  (ví dụ  $x = 0.5$ ):**

$$\begin{aligned} y(0.5) &= 1 \cdot \max(0, 0.5) - 2 \cdot \max(0, -0.5) + 1 \cdot \max(0, -1.5) \\ &= 1 \cdot (0.5) - 2 \cdot 0 + 1 \cdot 0 \\ &= 0.5 \end{aligned}$$

(Tổng quát trong khoảng này:  $y(x) = 1(x) - 0 - 0 = x$ )

**3. Khi  $1 \leq x < 2$  (ví dụ  $x = 1.5$ ):**

$$\begin{aligned}y(1.5) &= 1 \cdot \max(0, 1.5) - 2 \cdot \max(0, 0.5) + 1 \cdot \max(0, -0.5) \\&= 1 \cdot (1.5) - 2 \cdot (0.5) + 1 \cdot 0 \\&= 1.5 - 1.0 = 0.5\end{aligned}$$

(Tổng quát:  $y(x) = 1(x) - 2(x-1) = x - 2x + 2 = -x + 2$ )

**4. Khi  $x \geq 2$  (ví dụ  $x = 3$ ):**

$$\begin{aligned}y(3) &= 1 \cdot \max(0, 3) - 2 \cdot \max(0, 2) + 1 \cdot \max(0, 1) \\&= 1 \cdot 3 - 2 \cdot 2 + 1 \cdot 1 \\&= 3 - 4 + 1 = 0\end{aligned}$$

(Tổng quát:  $y(x) = 1(x) - 2(x-1) + 1(x-2) = x - 2x + 2 + x - 2 = 0$ )

---

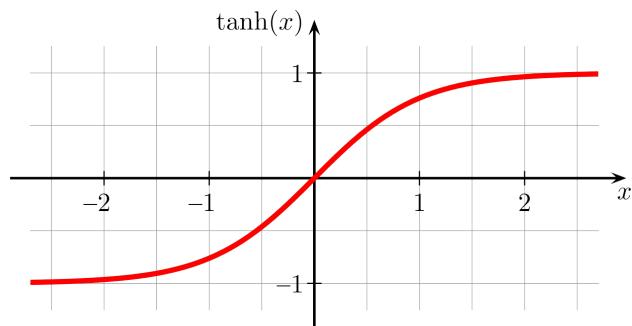
### 3 Ví dụ 2: Dùng Tanh (Hàm mượt)

Mục tiêu là cộng các "đường cong chữ S" (S-curves) để "nắn" ra một hình dạng mượt mà.

#### 3.1 Định nghĩa hàm Tanh

Hàm kích hoạt Tanh (Hyperbolic Tangent) được định nghĩa là:

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (4)$$



Hình 2: Minh họa đồ thị hàm Tanh

Giá trị của  $\tanh(z)$  luôn bị chặn trong khoảng  $(-1, 1)$ .

#### 3.2 Bài toán: Tạo "cục u" mượt

Chúng ta muốn tạo một hàm  $y(x)$  bắt đầu từ 0, tăng mượt lên đỉnh, và giảm mượt về 0. Chúng ta dùng 2 nơ-ron ẩn.

$$f_1(x) = \tanh(5x + 5) \quad (\text{Nơ-ron 1, } c_1 = 1. \text{ Tâm tại } x = -1)$$

$$f_2(x) = \tanh(5x - 5) \quad (\text{Nơ-ron 2, } c_2 = -1. \text{ Tâm tại } x = 1)$$

Hàm cuối cùng  $y(x)$  là tổng của chúng:

$$y(x) = \tanh(5x + 5) - \tanh(5x - 5) \quad (5)$$

### 3.3 Phân tích từng bước (Theo giới hạn)

Khi  $x \rightarrow -\infty$ :

1. Tại  $x = -3$ :

$$\begin{aligned}y(-3) &= \tanh(5(-3) + 5) - \tanh(5(-3) - 5) \\&= \tanh(-10) - \tanh(-20) \\&\approx (-1) - (-1) = 0\end{aligned}$$

2. Tại  $x = -1$  (Tâm của S-curve thứ nhất):

$$\begin{aligned}y(-1) &= \tanh(5(-1) + 5) - \tanh(5(-1) - 5) \\&= \tanh(0) - \tanh(-10) \\&\approx 0 - (-1) = 1\end{aligned}$$

3. Tại  $x = -0.5$  (Giữa "vai" và "đỉnh"):

$$\begin{aligned}y(-0.5) &= \tanh(5(-0.5) + 5) - \tanh(5(-0.5) - 5) \\&= \tanh(-2.5 + 5) - \tanh(-2.5 - 5) \\&= \tanh(2.5) - \tanh(-7.5) \\&\approx 0.987 - (-1) = 1.987\end{aligned}$$

4. Khi  $x = 0$  (Tại đỉnh):

$$\begin{aligned}y(0) &= \tanh(5(0) + 5) - \tanh(5(0) - 5) \\&= \tanh(5) - \tanh(-5) \\&\approx (1) - (-1) = 2\end{aligned}$$

Khi  $x \rightarrow +\infty$

5. Tại  $x = 0.5$ :

$$\begin{aligned}y(0.5) &= \tanh(5(0.5) + 5) - \tanh(5(0.5) - 5) \\&= \tanh(2.5 + 5) - \tanh(2.5 - 5) \\&= \tanh(7.5) - \tanh(-2.5) \\&\approx 1 - (-0.987) = 1.987\end{aligned}$$

6. Tại  $x = 1$ :

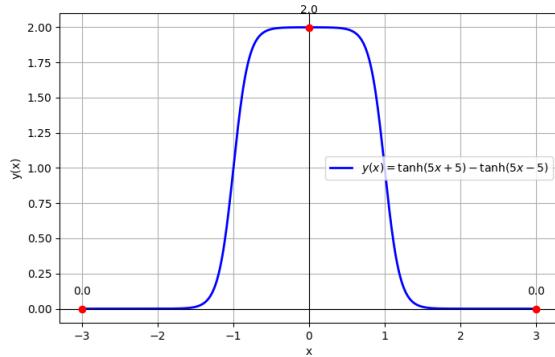
$$\begin{aligned}y(1) &= \tanh(5(1) + 5) - \tanh(5(1) - 5) \\&= \tanh(10) - \tanh(0) \\&\approx 1 - 0 = 1\end{aligned}$$

7. Tại  $x = 3$ :

$$\begin{aligned}y(3) &= \tanh(5(3) + 5) - \tanh(5(3) - 5) \\&= \tanh(20) - \tanh(10) \\&\approx (1) - (1) = 0\end{aligned}$$

Phân tích: Dối xứng hoàn hảo. Đồ thị bắt đầu "cong" xuồng từ đỉnh.

Kết quả là một "cục u" mượt, phi tuyến được tạo ra bằng cách trừ hai hàm S-curves của tanh.



Hình 3: Đồ thị 'cục u' từ hai hàm tanh

## Tài liệu

- [1] GeeksforGeeks, *ReLU Activation Function in Deep Learning*, Available at: <https://www.geeksforgeeks.org/deep-learning/relu-activation-function-in-deep-learning/>, Accessed: 16 Nov 2025.
- [2] GeeksforGeeks, *Tanh Activation in Neural Network*, Available at: <https://www.geeksforgeeks.org/deep-learning/tanh-activation-in-neural-network/>, Accessed: 16 Nov 2025.