**Phần 2\_4\_Lớp phi tuyến (activation)**

**Notes:**

* *Về format cấu trúc soạn như sau:*
  + *Lý thuyết…*
  + *Bộ code mấu/ ví dụ …*
  + *Ứng dụng (nếu có)...*
* *Mems làm nhớ note tên để mn dễ contact*

**Mục lục**

[**I. Nội dung chính 1**](#_70tfckp0agcb)

[**II. Nội dung biên soạn chi tiết 1**](#_wmm387ki8dm1)

[1. ReLU (Rectified Linear Unit): 1](#_avrhn5b6bcnm)

[2. ELU (Exponential Linear Unit): 2](#_4x6higa0zaha)

[3. Sigmoid: 2](#_aczsn6mizolq)

[4. Tanh (Hyperbolic Tangent) 2](#_qs0sk72ribyu)

[5. Softmax: 3](#_2a8u7ox6rd3s)

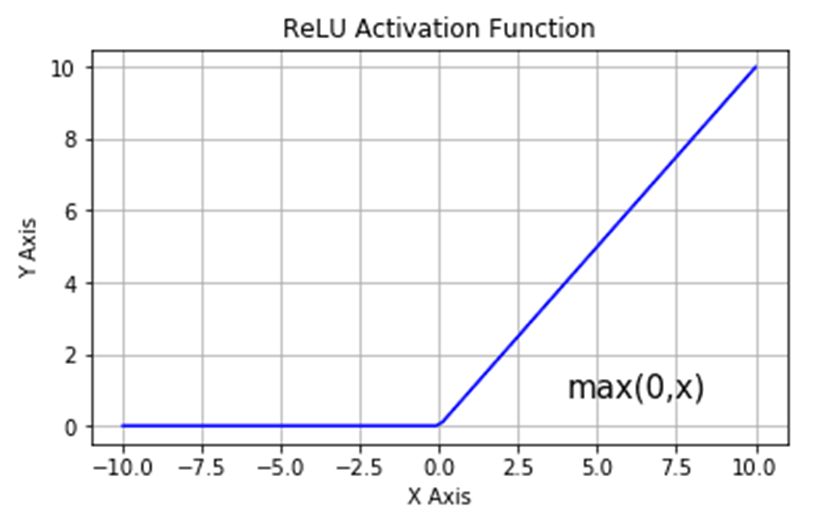
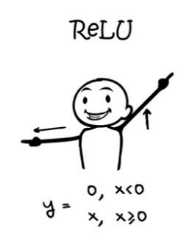
### **I. Nội dung chính**

Về các phép tính toán (lớp tính toán) để xây dựng mô hình học sâu

* + (4) Lớp phi tuyến (**activation**): **relu**, **elu**, **sigmoid**, **tanh**, **softmax**
    - Nguyên lý làm việc
    - Số tham số có thể học
    - Độ rộng của vùng nhận thức (receptive fields)
    - Hình dạng kích thước của bản đồ đặc trưng đầu vào và đầu

### **II. Nội dung biên soạn chi tiết**

#### **1. ReLU (Rectified Linear Unit):**

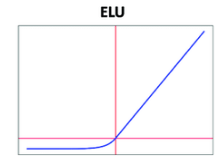
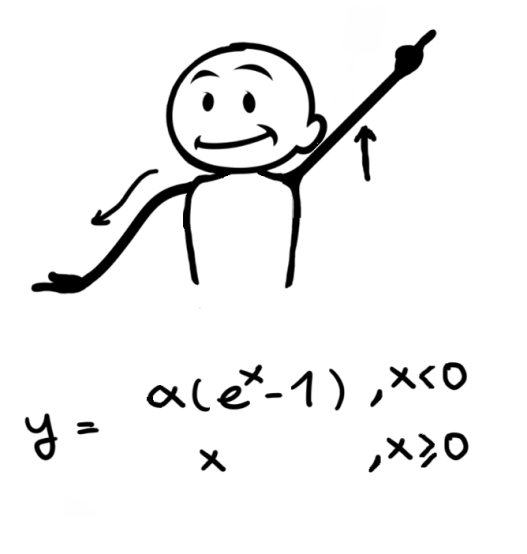
****

Hàm ReLU được định nghĩa là f(x) = max(0, x).

Nó trả về **0 cho các giá trị âm** và **giá trị đầu vào cho các giá trị dương**.

ReLU được sử dụng rộng rãi như một hàm kích hoạt do tính đơn giản và **khả năng giảm vấn đề gradient biến mất** (vanishing gradient problem).

#### **2. ELU (Exponential Linear Unit):**

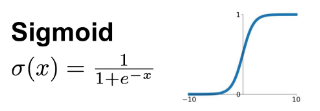
****

Hàm ELU được định nghĩa là f(x) = x nếu x >= 0, và f(x) = alpha \* (exp(x) - 1) nếu x < 0, trong đó alpha là một siêu tham số.

Tương tự như ReLU, ELU tránh vấn đề gradient biến mất.

ELU có sự chuyển tiếp mượt mà hơn cho các giá trị âm so với ReLU và có thể cung cấp hiệu suất học tốt hơn trong một số trường hợp.

#### **3. Sigmoid:**

****

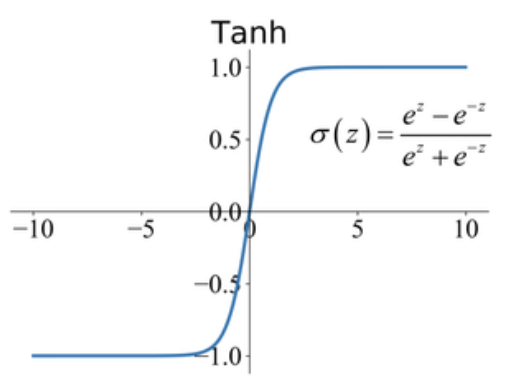
Hàm sigmoid được định nghĩa là f(x) = 1 / (1 + exp(-x)).

Nó ánh xạ đầu vào thành một phạm vi từ **0 đến 1**.

Sigmoid thường được sử dụng trong các bài toán phân loại nhị phân**(0 hoặc 1)** vì nó có thể tạo ra đầu ra giống như xác suất.

#### **4. Tanh (Hyperbolic Tangent)**

**Y= Tanh(x)**

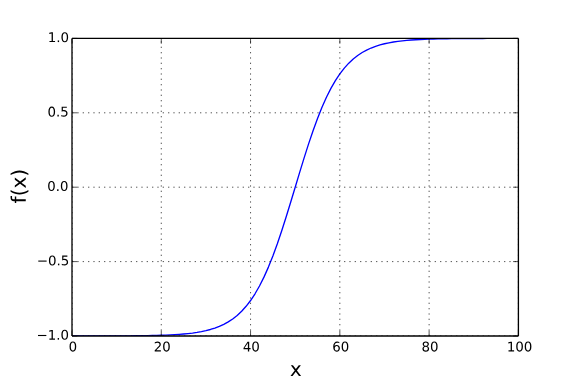
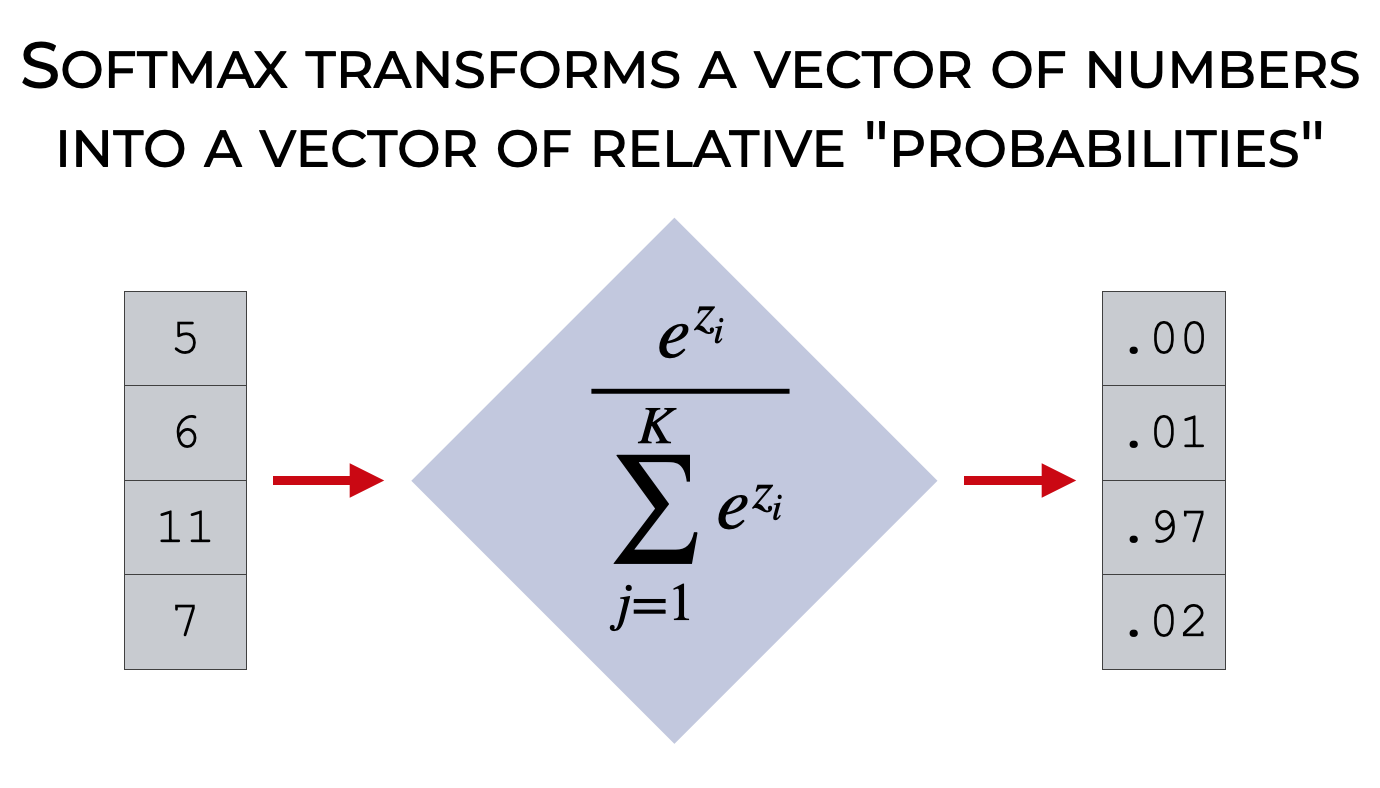
****

Hàm tanh được định nghĩa là f(x) = (exp(x) - exp(-x)) / (exp(x) + exp(-x)).

Nó **ánh xạ đầu vào thành một phạm vi từ -1 đến 1.**

Tanh thường được sử dụng làm **hàm kích hoạt trong các mạng neural hồi quy** (RNNs) và có thể mô hình cả **giá trị dương và âm**.

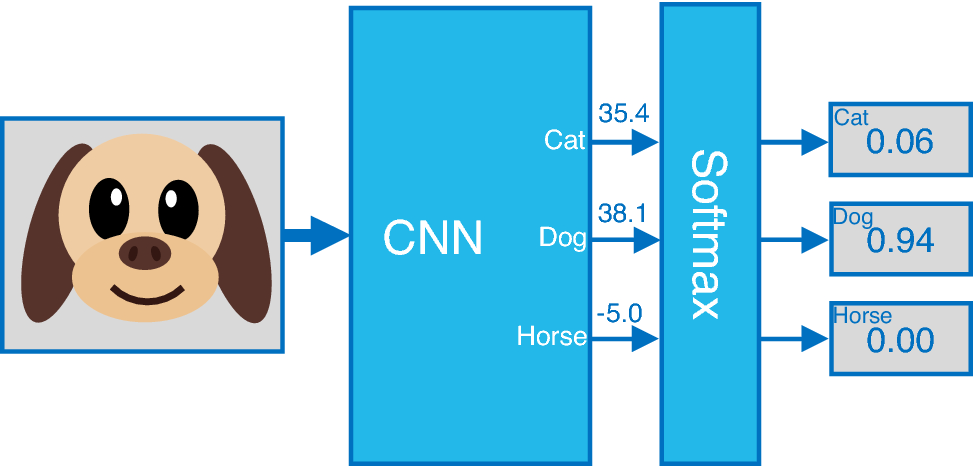
#### **5. Softmax:**

****

Hàm softmax được sử dụng để chuyển đổi một **vector số thực thành một phân phối xác suất**.

Nó lấy một vector đầu vào và **chuẩn hóa nó để tạo ra một phân phối xác suất** trong đó **tổng của tất cả các xác suất bằng 1**.

Softmax thường được sử dụng làm **hàm kích hoạt trong lớp đầu ra của các bài toán phân loại đa lớp**.

****

VD như hình trên : cat output = exp(35.4) / (exp(35.4)+exp(38.1)+exp(-5)) =0.06 (6%)

dog output = exp(38.1) / (exp(35.4)+exp(38.1)+exp(-5)) = 0.94 (94%)

horse output = exp(-5) / (exp(35.4)+exp(38.1)+exp(-5)) = 0.00 (0%)