



HUST

ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

ONE LOVE. ONE FUTURE.



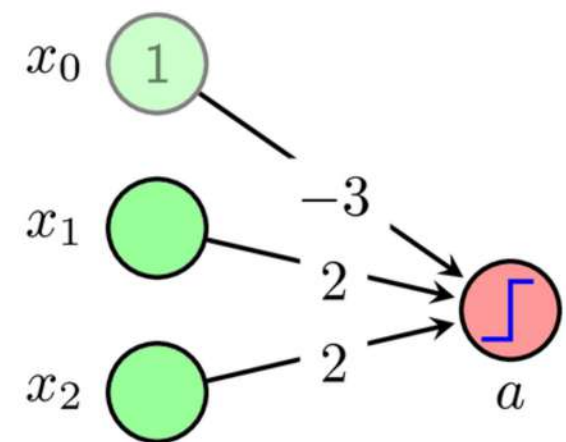
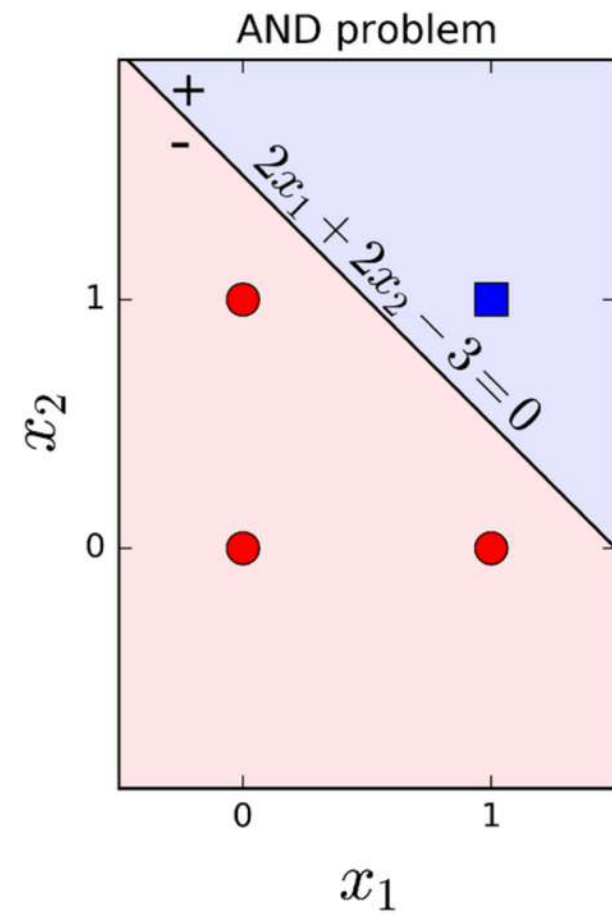
ĐẠI HỌC
BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Multilayer Perceptron

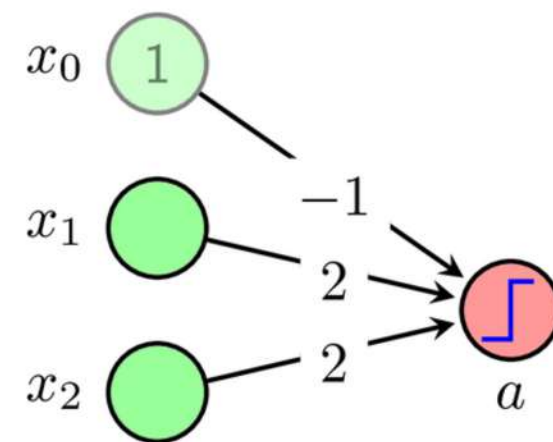
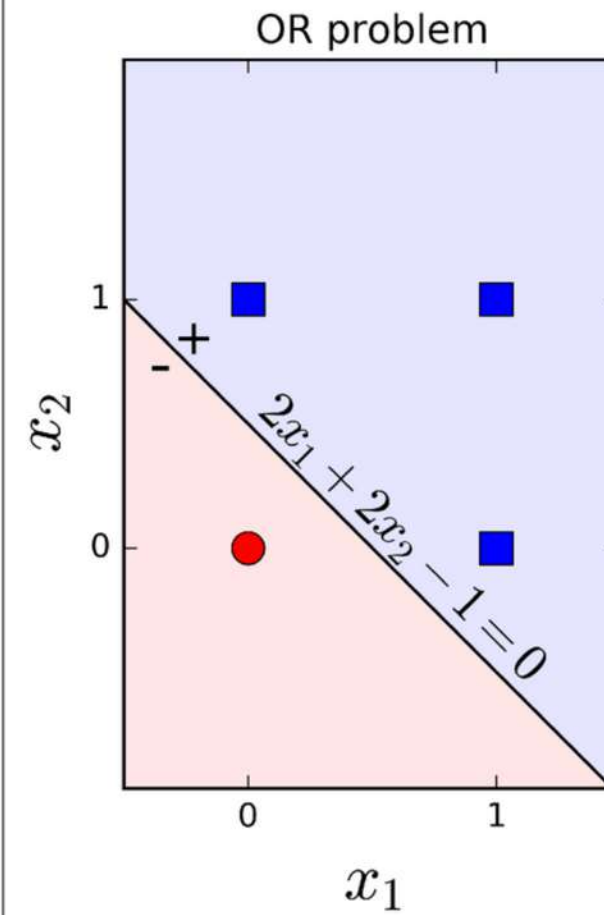
ONE LOVE. ONE FUTURE.

- Perceptron learning
- Multilayer Perceptron
 - Activation Function
 - Forward, Backward Propagation
- Ứng dụng bài toán thực tế

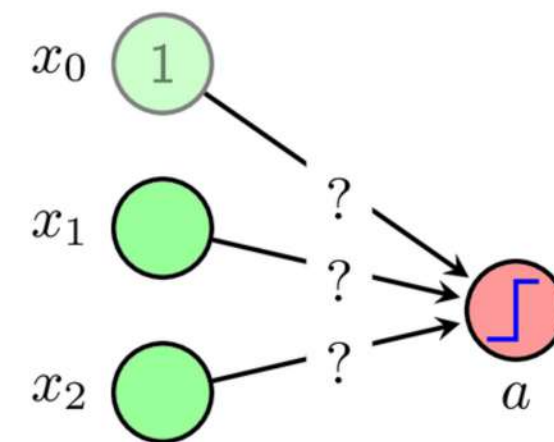
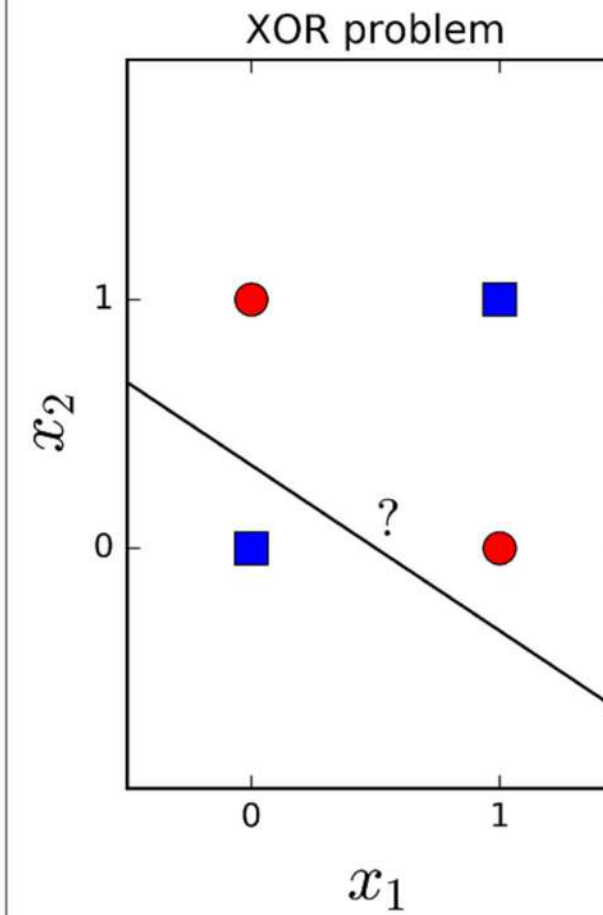
Perceptron Learning



$$a = \text{sgn}(2x_1 + 2x_2 - 3)$$

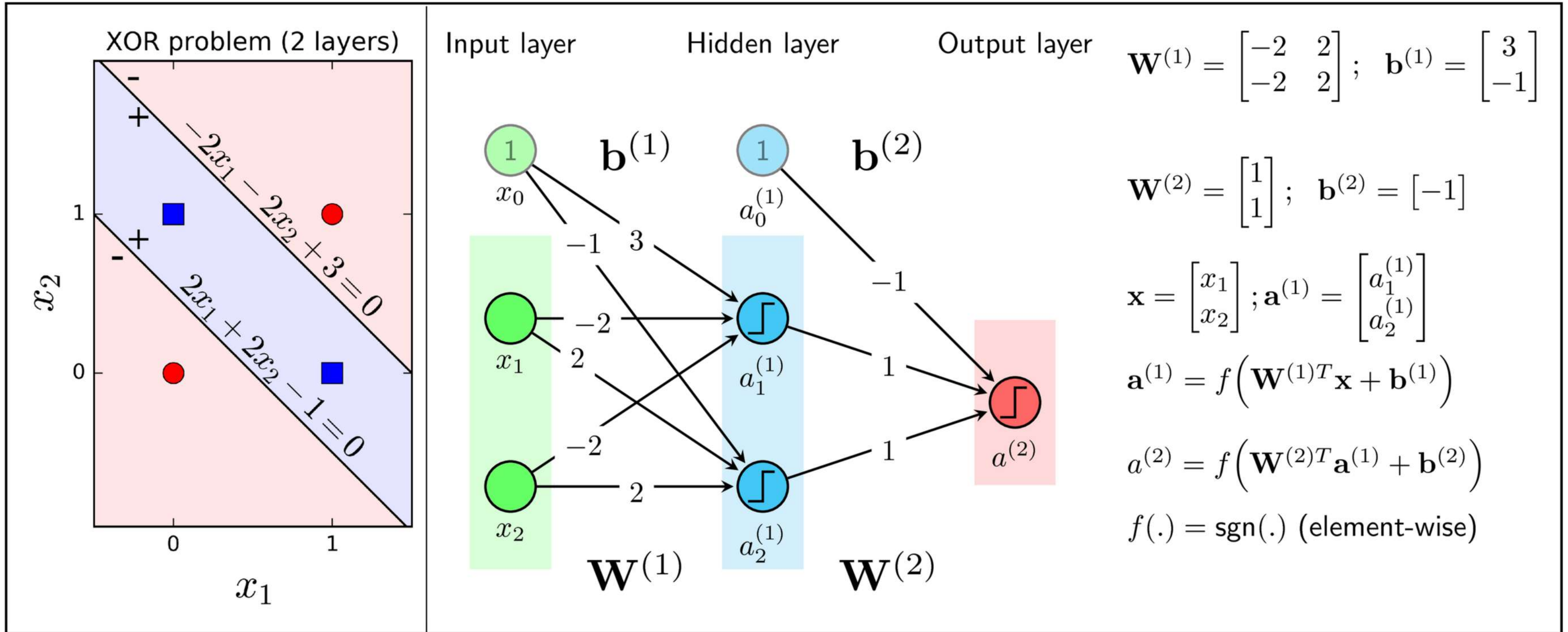


$$a = \text{sgn}(2x_1 + 2x_2 - 1)$$

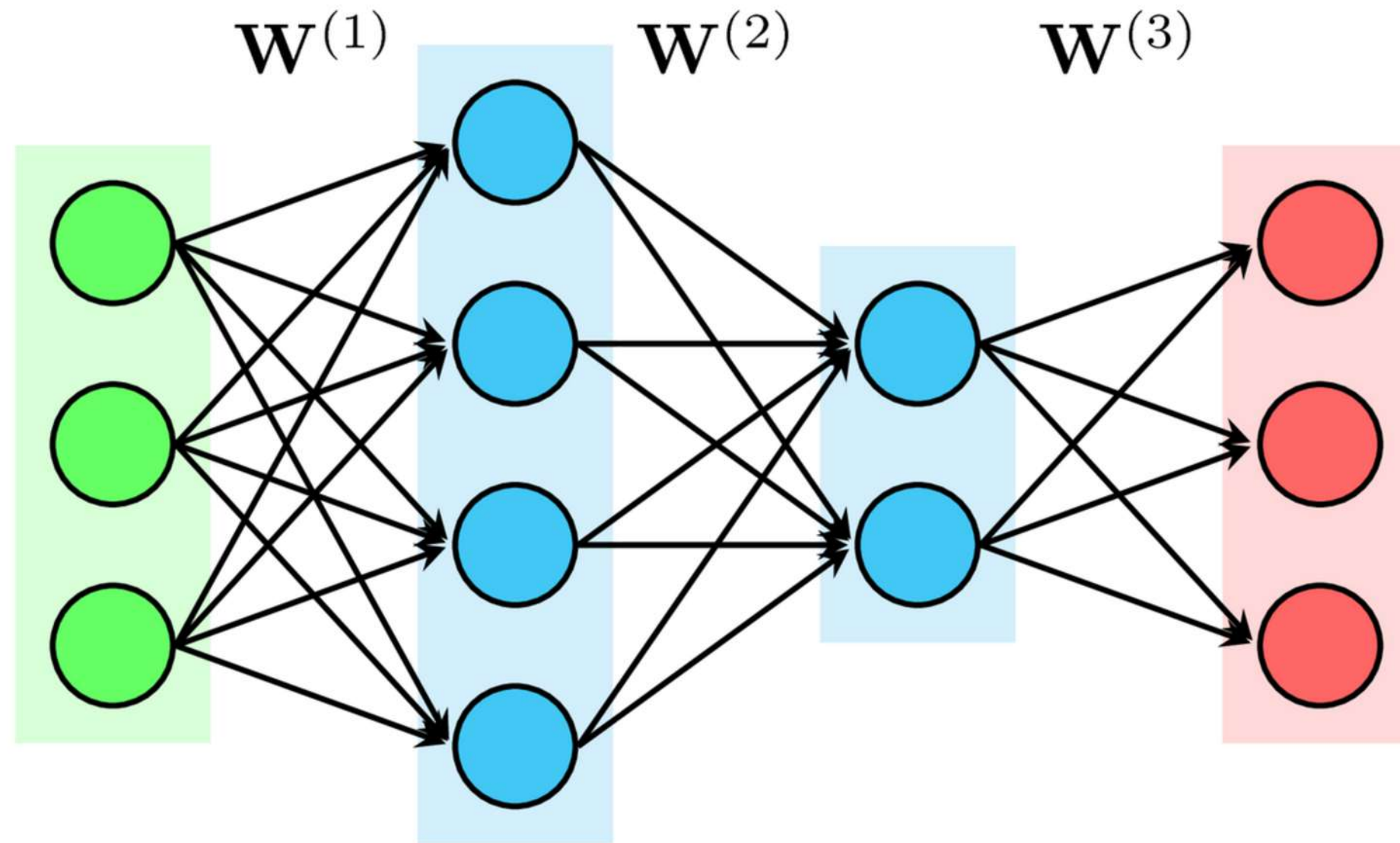


No solution!

Multilayer Perceptron



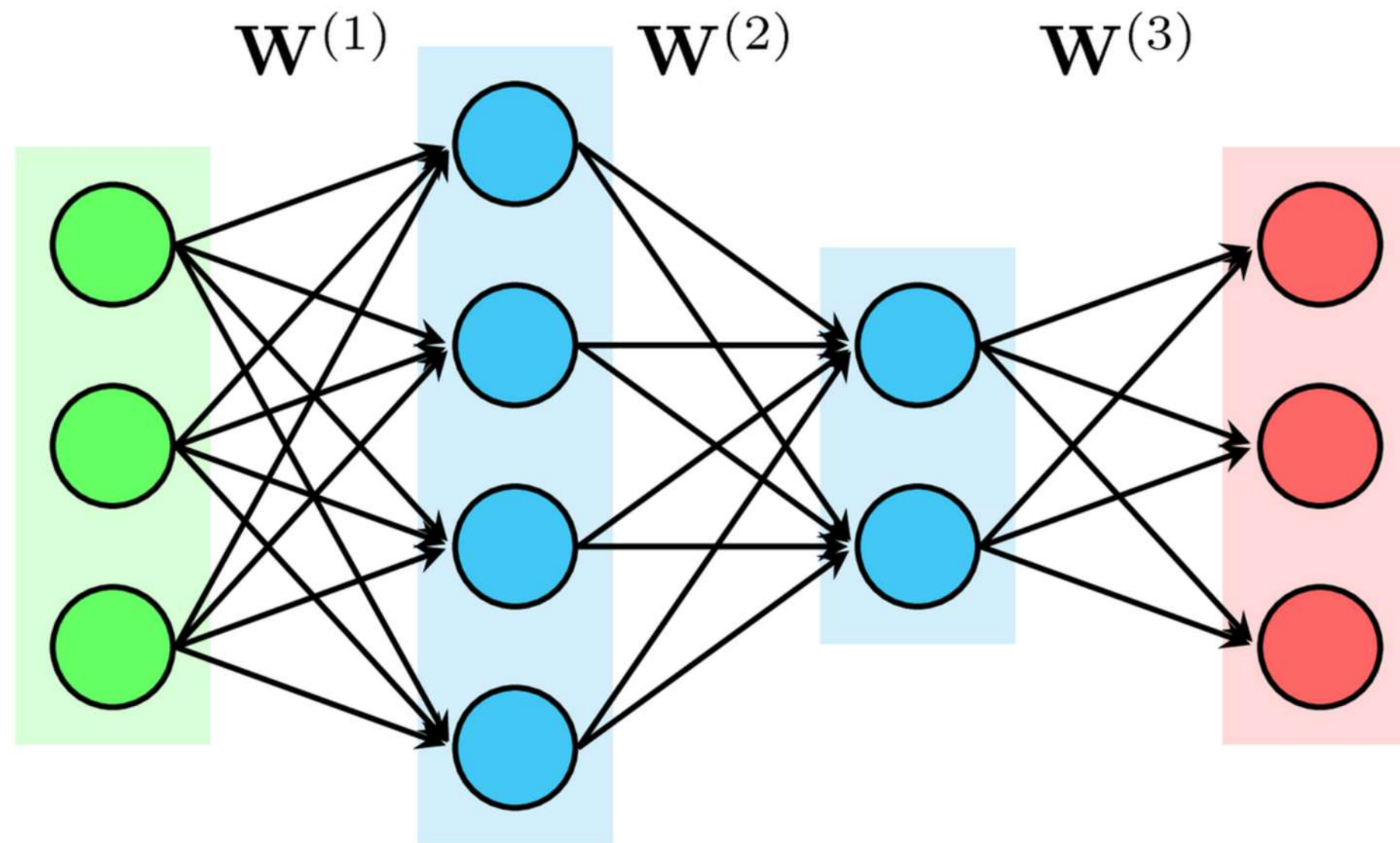
Multilayer Perceptron



$$H = XW^{(1)} + b^{(1)}$$
$$O = HW^{(2)} + b^{(2)}$$

$$O = \left(XW^{(1)} + b^{(1)} \right) W^{(2)} + b^{(2)} = XW^{(1)}W^{(2)} + b^{(1)}W^{(2)} + b^{(2)} = XW + b$$

Multilayer Perceptron

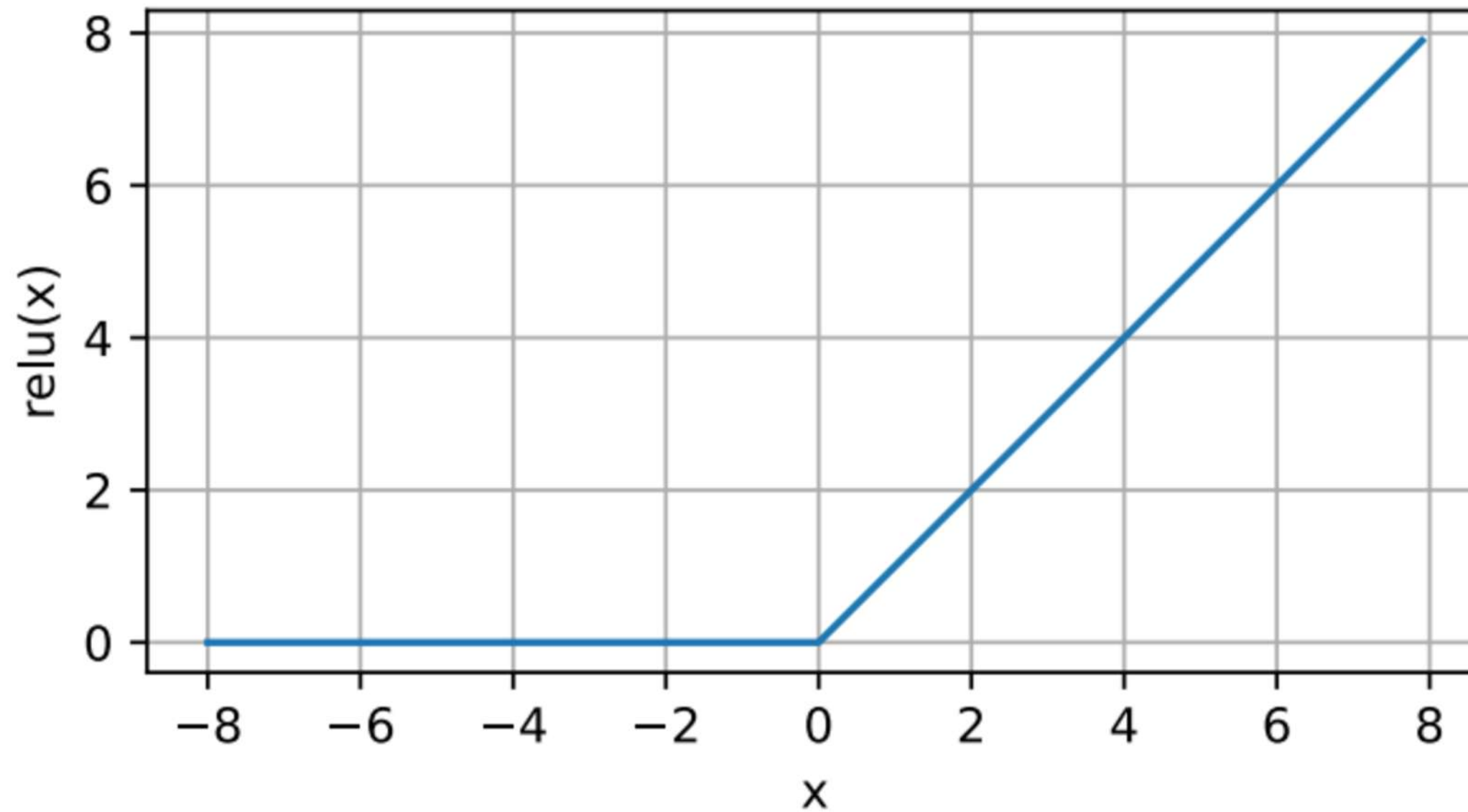


$$H = f \left(XW^{(1)} + b^{(1)} \right)$$
$$O = HW^{(2)} + b^{(2)}$$

- Hàm kích hoạt $f(.)$ đưa tính phi tuyến vào mô hình

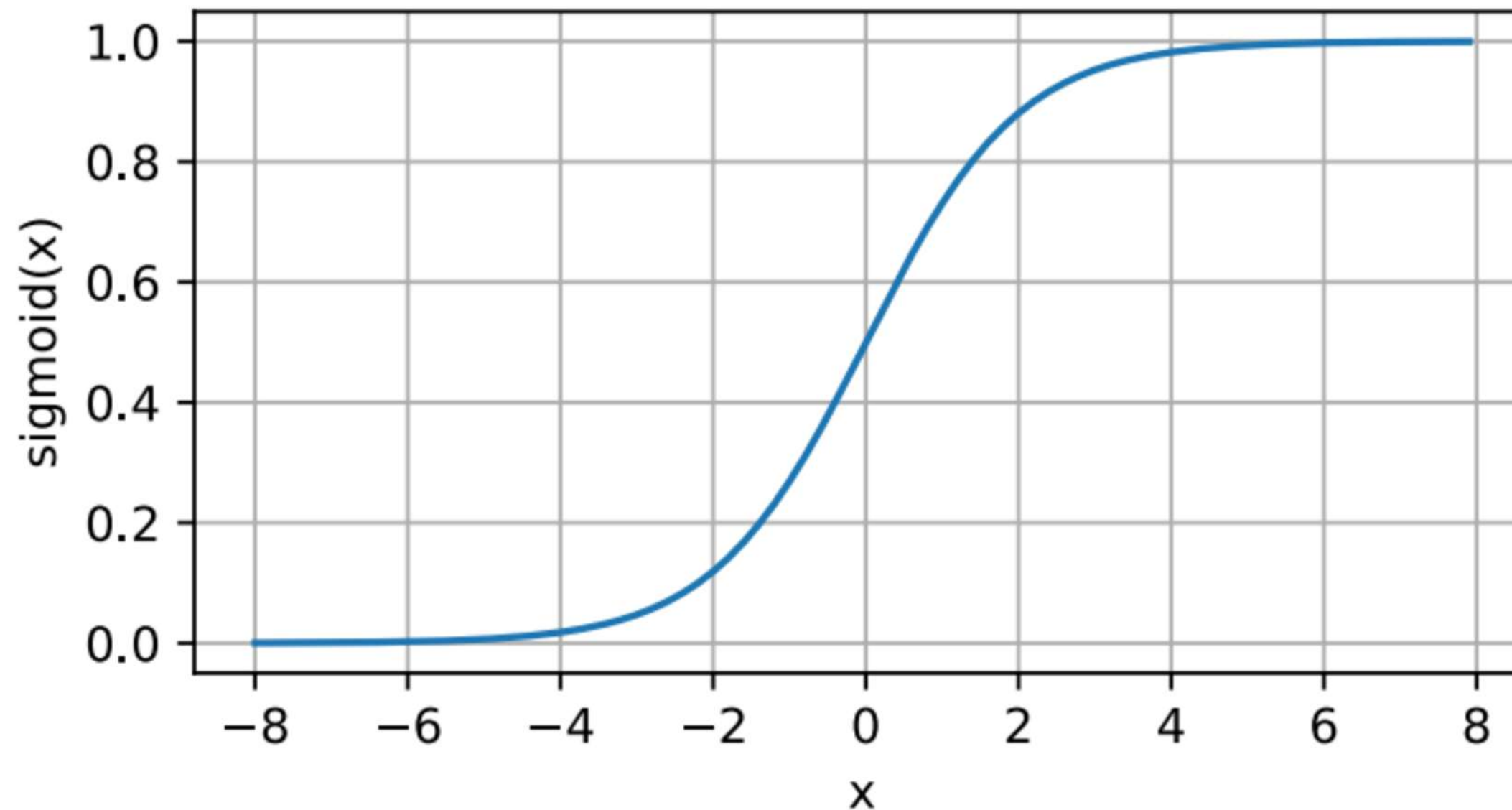
Activation Function

$$\text{ReLU}(x) = \max(0, x)$$



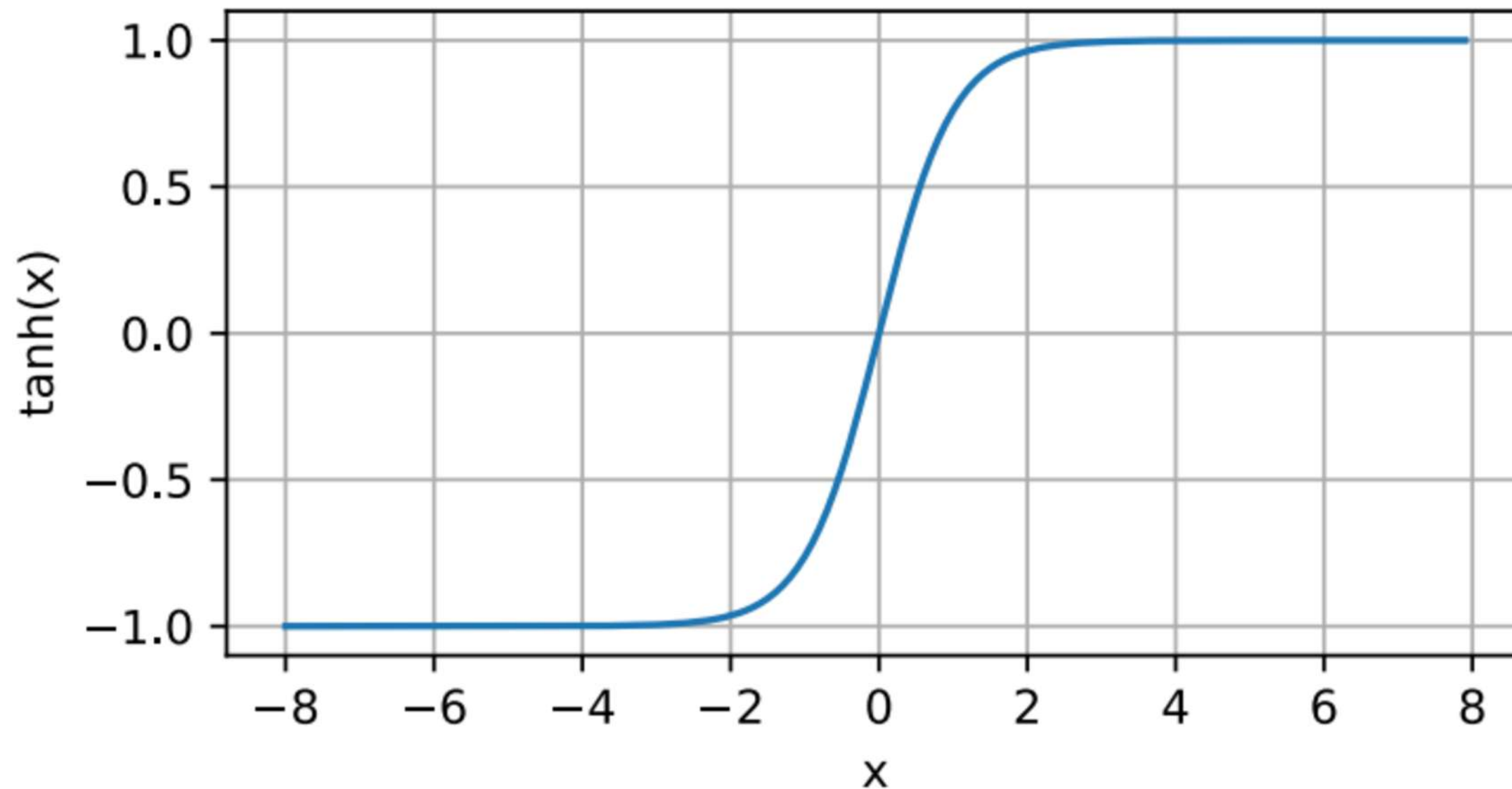
Activation Function

$$\text{sigmoid}(x) = \frac{1}{1 + \exp(-x)}$$



Activation Function

$$\tanh(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}$$



- Thực hiện tính toán lần lượt từ lớp đầu tiên đến lớp cuối cùng

$$\begin{aligned}a^{(0)} &= x \\z^{(i)} &= a^{(i-1)}W^{(i)} + b^{(i)}; i = 1, \dots, n \\a^{(i)} &= f\left(z^{(i)}\right); i = 1, \dots, n\end{aligned}$$

- Backpropagation giúp tính gradient ngược từ layer cuối cùng đến layer đầu tiên bằng cách sử dụng quy tắc chuỗi:

$$\frac{df}{dx} = \frac{df}{du} \frac{du}{dx}$$

Backward Propagation

$$\begin{aligned}\frac{dJ}{dW^{(i)}} &= \frac{dJ}{dz^{(i)}} \frac{dz^{(i)}}{dW^{(i)}} = \frac{dJ}{dz^{(i)}} a^{(i-1)} \\&= \frac{dJ}{da^{(i)}} \frac{da^{(i)}}{dz^{(i)}} a^{(i-1)} \\&= \frac{dJ}{dz^{(i+1)}} \frac{dz^{(i+1)}}{da^{(i)}} f' \left(z^{(i)} \right) a^{(i-1)} \\&= \frac{dJ}{dW^{(i+1)}} \frac{1}{a^{(i)}} W^{(i+1)} f' \left(z^{(i)} \right) a^{(i-1)}\end{aligned}$$

$$\frac{d}{dx} \text{sigmoid}(x) = \text{sigmoid}(x) (1 - \text{sigmoid}(x)) \leq 0.25$$

$$\frac{d}{dx} \tanh(x) = 1 - \tanh^2(x) \leq 1$$

$$\frac{d}{dx} \text{ReLU}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

- Universal Approximation Theorem: Mô hình MLP với 1 hidden layer và 1 activation function phù hợp có thể xấp xỉ bất kỳ hàm liên tục nào trên tập dữ liệu giới hạn
- Trong các bài toán thực tế, mô hình MLP với 2 hoặc 3 hidden layer là đủ

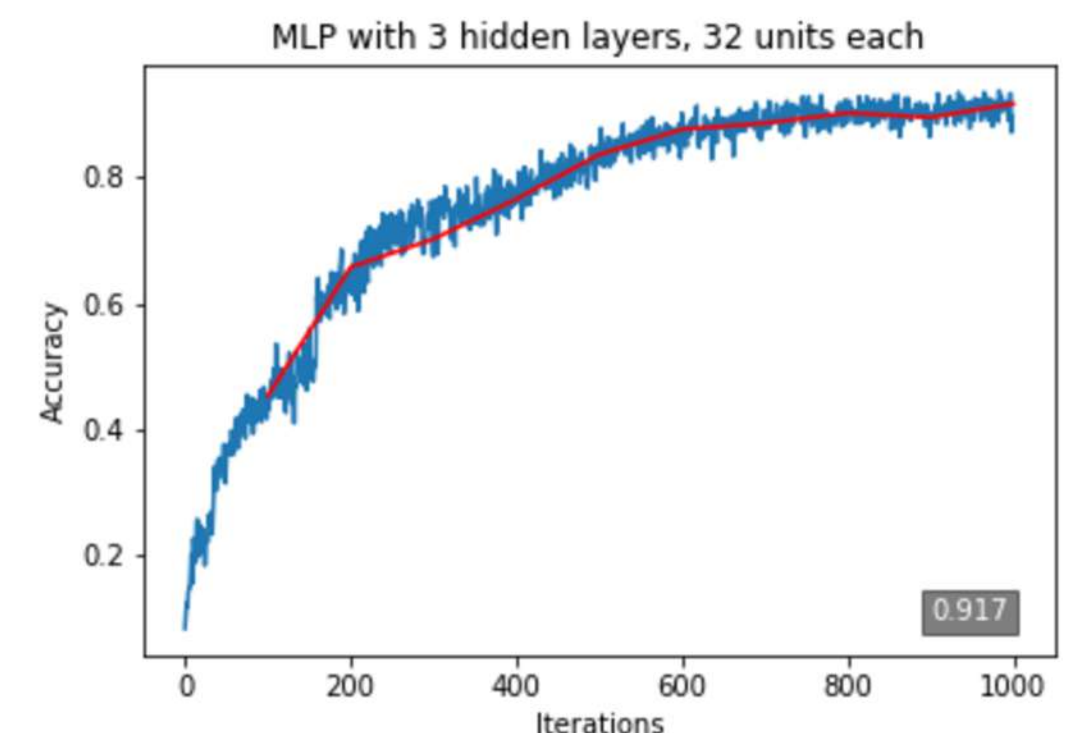
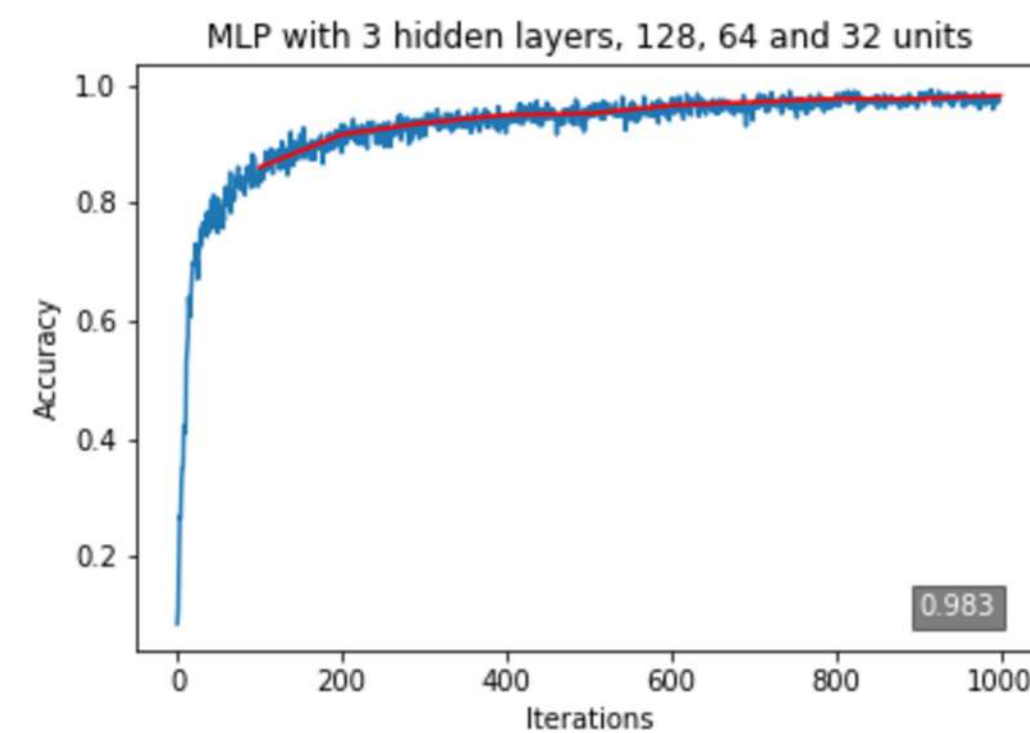
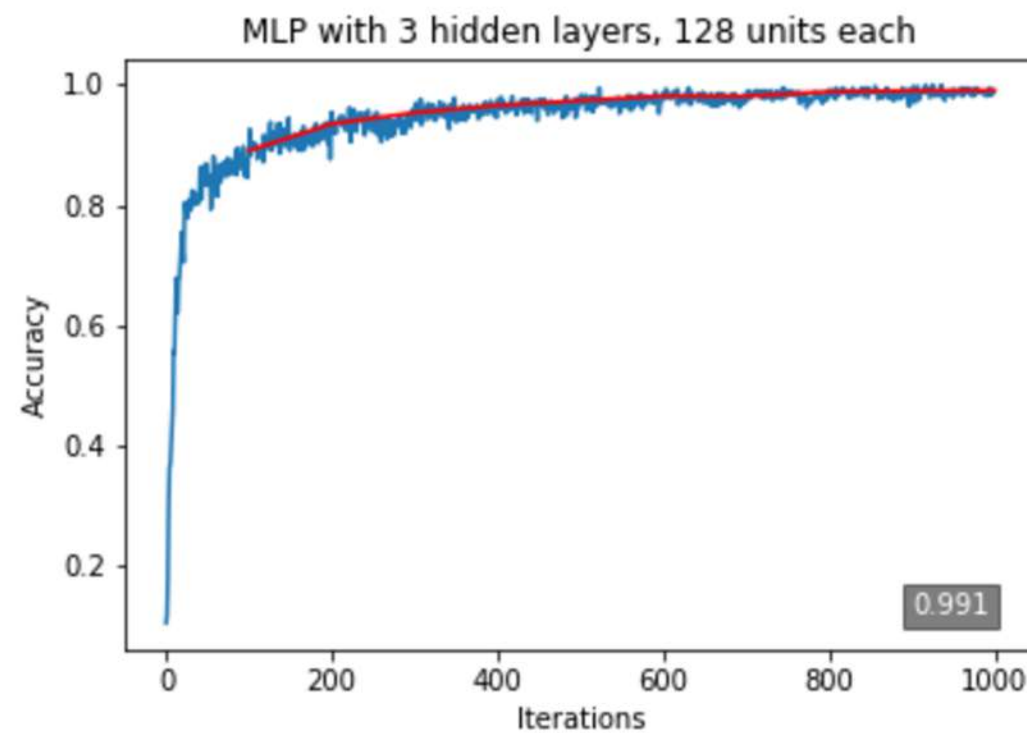
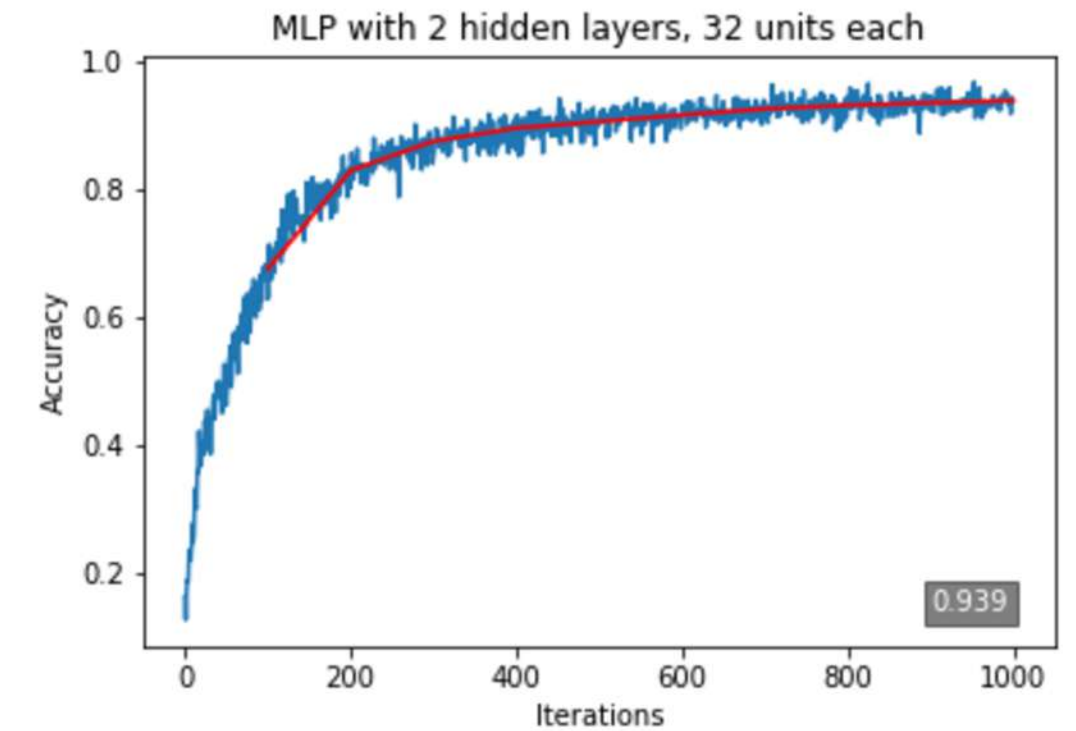
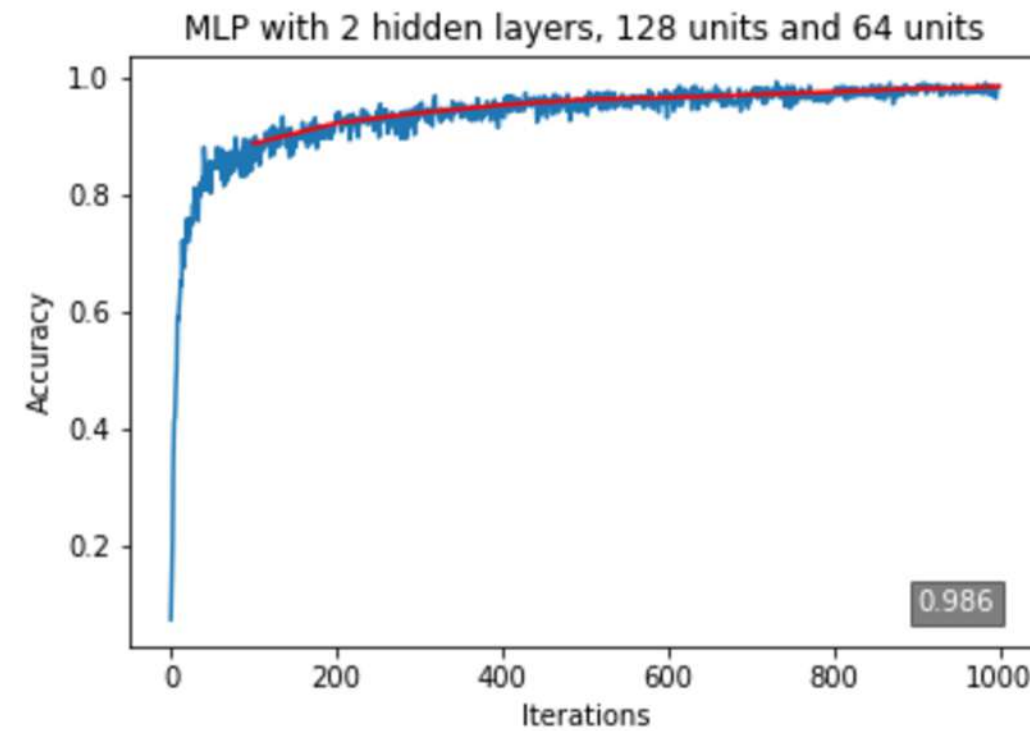
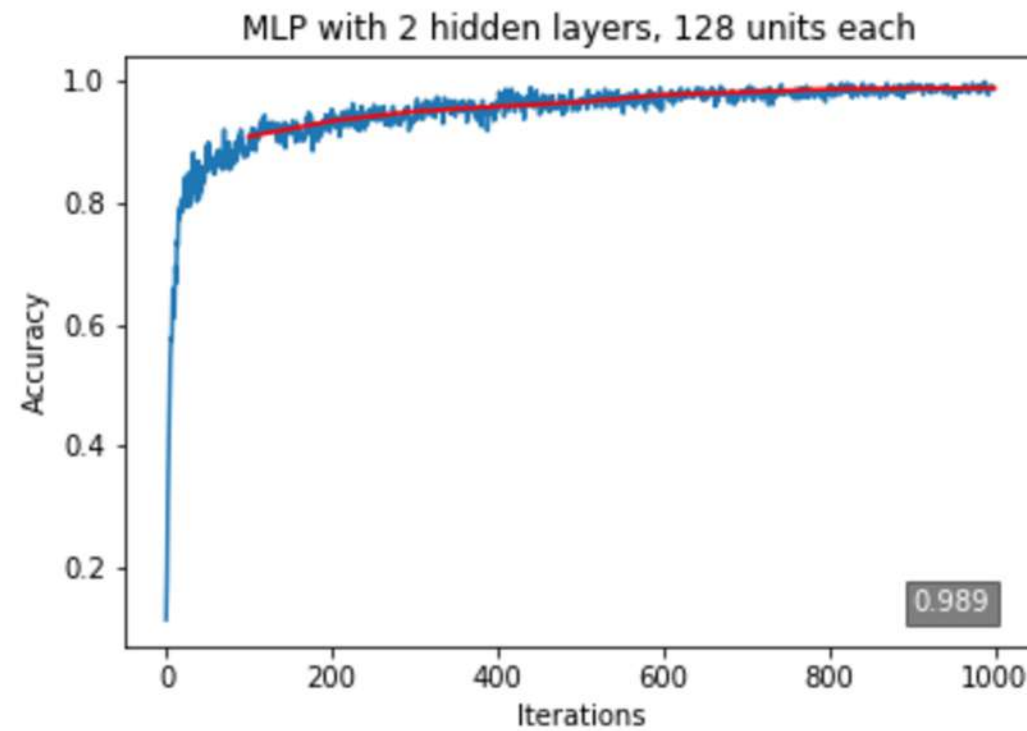
- Điểm mạnh của MLP:
 - Có thể dùng trong cả bài toán phân loại lẫn hồi quy
 - Có thể xấp xỉ mọi hàm liên tục (trên lý thuyết)
- Điểm yếu của MLP:
 - Khó mở rộng, tăng hidden layer làm tăng độ phức tạp của mô hình, dẫn đến tăng thời gian, kéo theo overfitting
 - Khó giải thích

Ứng dụng trong bài toán thực tế

- Bộ dữ liệu: Digit Recognizer
- Gồm dữ liệu 70000 dữ liệu của bức ảnh các chữ số viết tay
- Tỉ lệ train:test là 3:2



Ứng dụng trong bài toán thực tế



A large graphic on the left side of the slide. It features a dark blue background with a pattern of red dots of varying sizes arranged in concentric, slightly irregular circles, creating a sense of depth and movement. The word "HUST" is centered within this graphic.

HUST

THANK YOU !



hust.edu.vn



fb.com/dhbkhn