



Khoa Công Nghệ Thông Tin
Trường Đại Học Cần Thơ



Mạng nơ-ron nhân tạo Neural networks

Đỗ Thanh Nghị
dtngghi@cit.ctu.edu.vn

Cần Thơ
12-02-2019

Nội dung

- ★ Giới thiệu
- ★ Mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

Nội dung

- ★ Giới thiệu

- ★ Mô hình nơ-ron McCulloch & Pitts

- ★ Perceptron

- ★ Multi-layer perceptron (MLP)

- ★ Kết luận

Giới thiệu về mạng nơ-ron nhân tạo

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Mạng nơ-ron

- ★ Lớp các giải thuật máy học tự động mô hình tính toán dựa trên các mạng nơ-ron sinh học
- ★ Mạng nơ-ron là một nhóm các nơ-ron nhân tạo (nút) nối với nhau và xử lý thông tin bằng cách truyền tín hiệu theo các kết nối và tính giá trị mới tại các nút
- ★ Mạng nơ-ron nhân tạo được huấn luyện bằng các ví dụ (mẫu)
- ★ Lĩnh vực ứng dụng: nhận dạng chữ viết tay, nhận dạng mặt người, nhận dạng tiếng nói, phân loại văn bản và xử lý ngôn ngữ tự nhiên, điều khiển robot, điều khiển xe tự động, etc.

Giới thiệu về mạng nơ-ron nhân tạo

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Lịch sử phát triển

- ★ Mô phỏng suy luận của bộ não người
- ★ Nghiên cứu từ năm 1943: mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron (Rosenblatt, 1958): mô phỏng thị giác người
- ★ Giữa thập niên 1980: Mạng nơ-ron
- ★ Giải thuật lan truyền ngược: Rumelhart, Williams, Hinton
- ★ Mạng SOM: (Kohonen, 1984)
- ★ Mạng nơ-ron tích chập: Hinton, LeCun, Bengio
- ★ Mạng SVM: (Vapnik et al., 1995)
- ★ Deep learning: (Hinton et al., 2006)

Giới thiệu về mạng nơ-ron nhân tạo

★ Giới thiệu

★ Nơ-ron McCulloch & Pitts

★ Perceptron

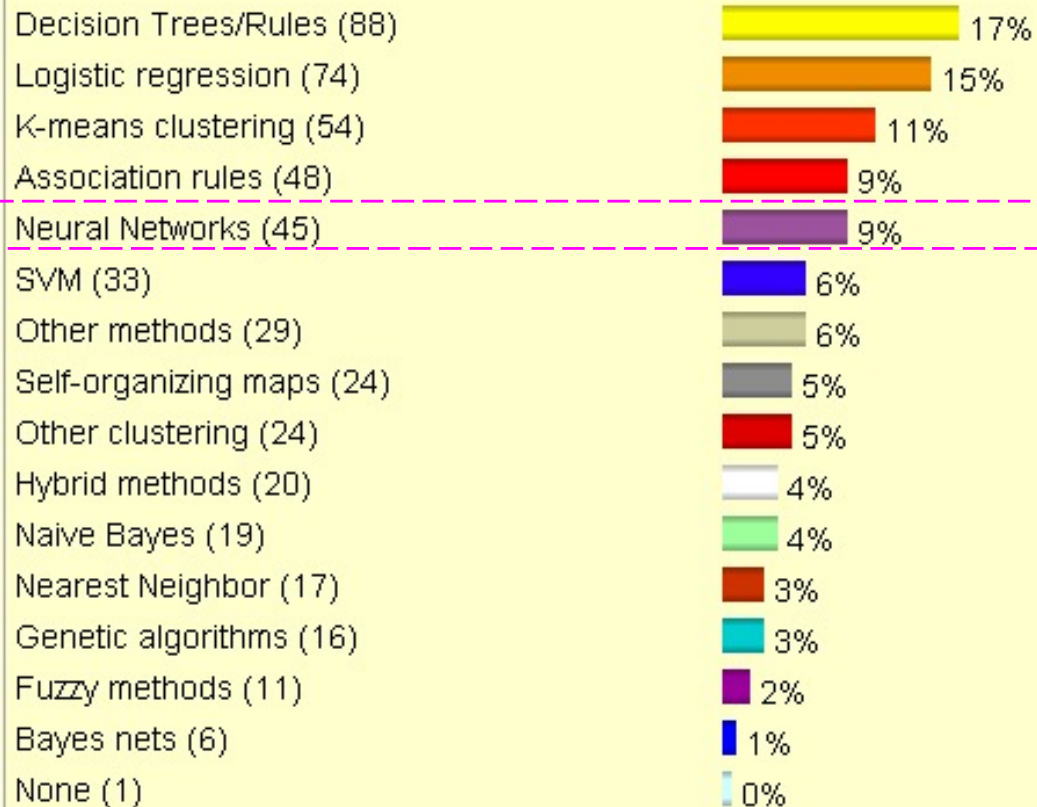
★ Multi-layer perceptron (MLP)

★ Kết luận

KDnuggets : Polls : Deployed data mining techniques

Poll

Which data mining techniques you used in a successfully deployed application?
[173 voters, 509 votes total]



Nội dung

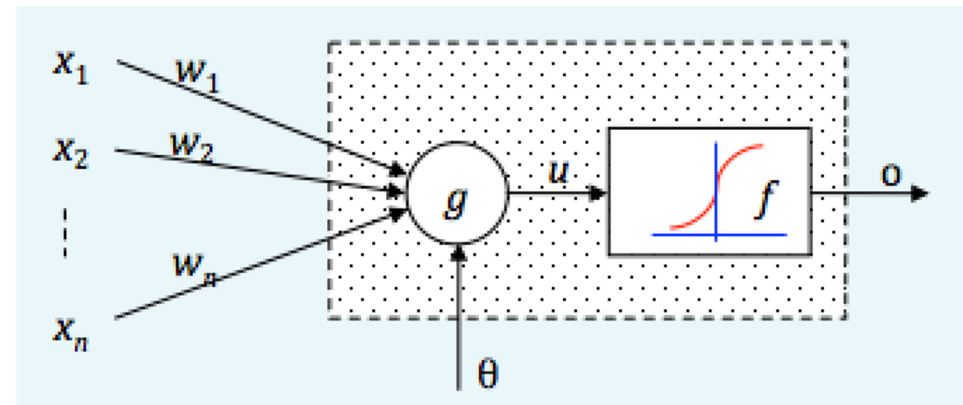
- ★ Giới thiệu
- ★ Mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

Neuron McCulloch & Pitts

- ★ Giới thiệu
- ★ Neuron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Neuron: đơn vị tính toán cơ bản

- ★ n ngõ vào, 1 giá trị ngưỡng (θ) và 1 ngõ ra (o)
- ★ w_i : trọng số
- ★ g : hàm kết hợp
- ★ f : hàm kích hoạt



$$u = g(x) = \sum_{i=1} w_i x_i + \theta$$

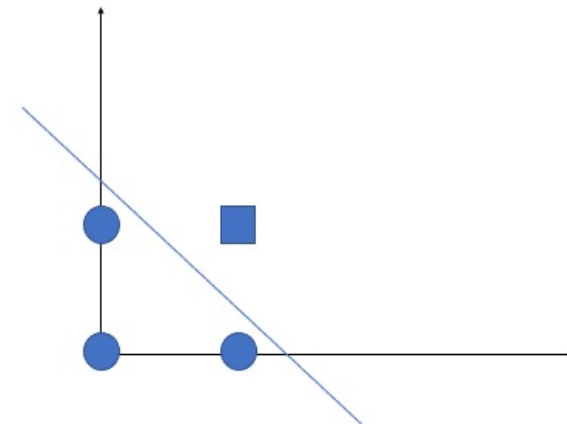
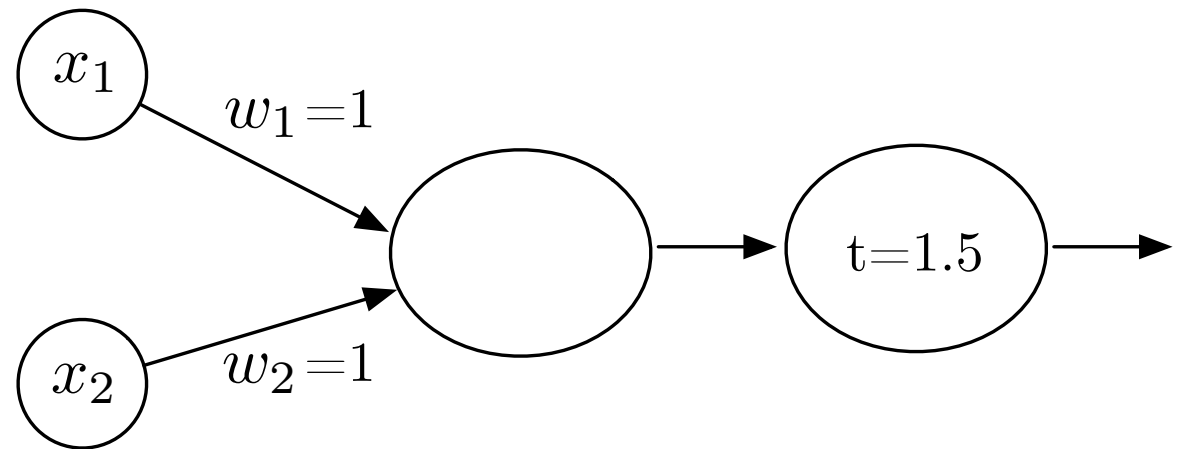
$$o = f(u) = \frac{1}{1 + e^{-u/T}}$$

hàm sigmoid

Neuron McCulloch & Pitts

- ★ Giới thiệu
- ★ Neuron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

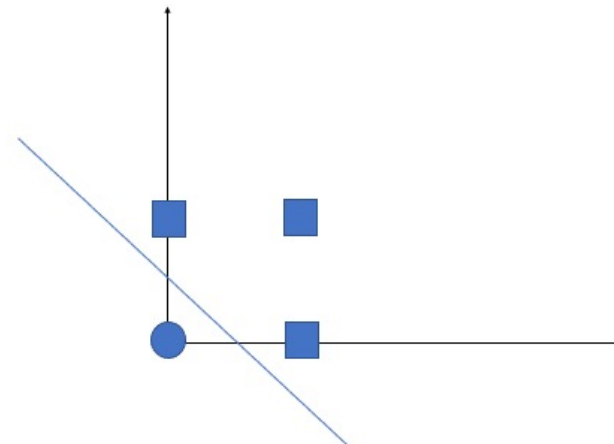
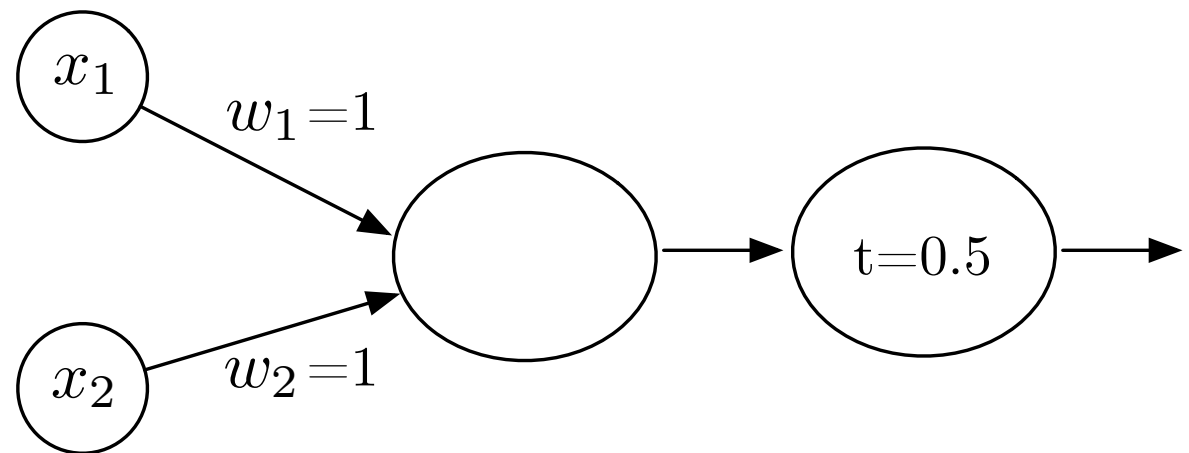
x_1	x_2	Out
0	0	0
0	1	0
1	0	0
1	1	1



Neuron McCulloch & Pitts

- ★ Giới thiệu
- ★ Neuron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

x_1	x_2	Out
0	0	0
0	1	1
1	0	1
1	1	1



Nội dung

- ★ Giới thiệu
- ★ Mô hình nơ-ron McCulloch & Pitts
- ★ **Perceptron**
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

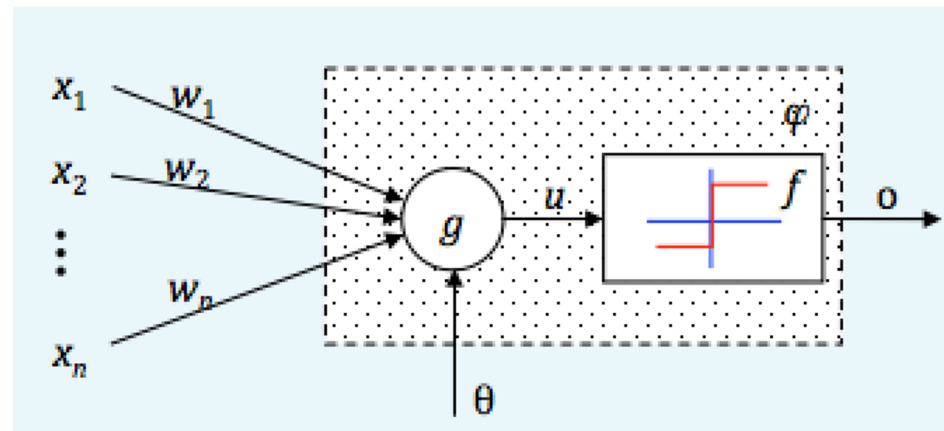
Perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Perceptron: 1 nơ-ron

- ★ Đề xuất bởi (Rosenblatt, 1958)
- ★ Tuân theo mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron với giá trị ngưỡng (θ)
 - ★ n ngõ vào, 1 giá trị ngưỡng và 1 ngõ ra
 - ★ Hàm kích hoạt: hàm step

$$u = g(x) = \sum_{i=1} w_i x_i + \theta$$
$$o = f(u) = \begin{cases} 1 & u \geq 0 \\ 0 & u < 0 \end{cases}$$



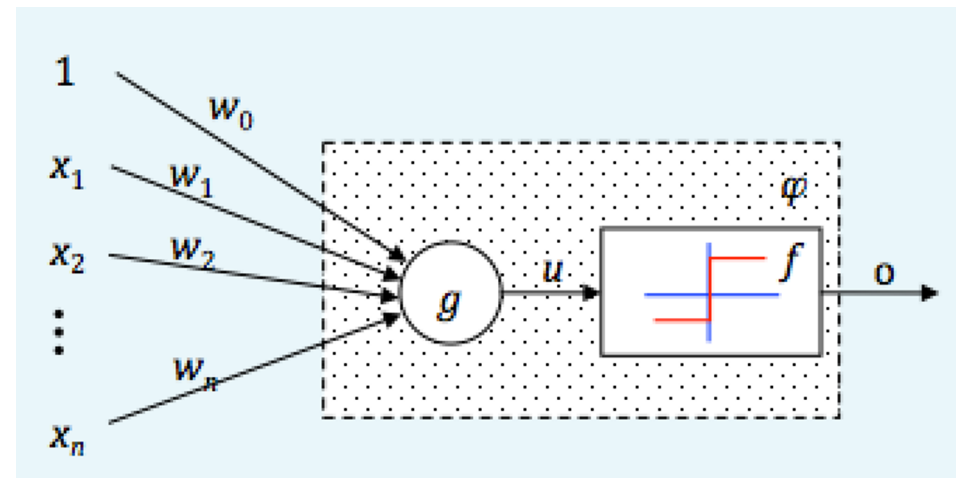
Perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Perceptron không giá trị ngưỡng

- ★ Đề đơn giản trong tính toán: xem ngưỡng như trọng số w_0 với ngõ vào giả x_0 bằng 1
- ★ $(n + 1)$ ngõ vào và 1 ngõ ra

$$u = g(x) = \sum_{i=0}^n w_i x_i$$

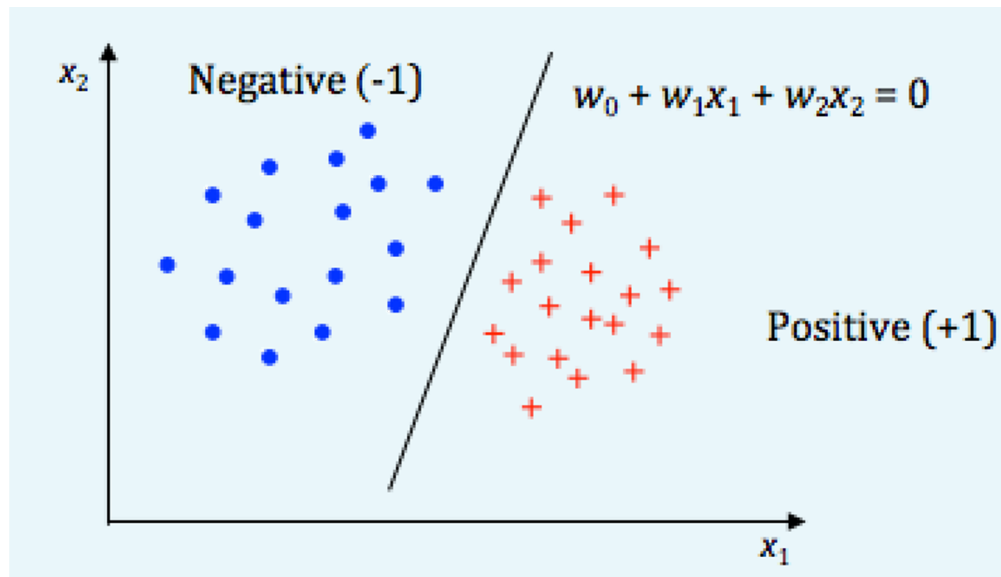


Perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Diễn giải hình học

- ★ Cho 1 véc-tơ $x = (x_1, x_2, \dots, x_n)$, perceptron tính toán ngõ ra o có thể bằng 0 hoặc 1 cho phân lớp nhị phân
- ★ Siêu phẳng = mô hình phân lớp



$$u = g(x) = \sum_{i=0}^n w_i x_i = 0$$

Perceptron

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Giải thuật huấn luyện

- ★ Tập huấn luyện: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$
- ★ Tìm các trọng số w_j sao cho perceptron nhận vào $x^{(i)}$ thì tính toán ngõ ra $o = y^{(i)}$ cho mọi $i = 1, 2, \dots, m$

★ Diễn giải hình học

- ★ Tìm siêu phẳng tách dữ liệu sao cho mỗi lớp dữ liệu nằm về 1 phía của siêu phẳng

Perceptron

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Giải thuật huấn luyện

- ★ Tập huấn luyện: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$
- ★ Tham số η : tốc độ học

★ Trường hợp tách rời tuyến tính

- ★ Khởi tạo ngẫu nhiên các trọng số w_j
- ★ Mỗi cặp $(x^{(i)}, y^{(i)})$, perceptron nhận $x^{(i)}$ và tính toán ngõ ra o tương ứng
- ★ Nếu $o \neq y^{(i)}$ thì cập nhật w_j

$$w_j = w_j + \eta \cdot (y_i - o_i) \cdot x_{ij}, \forall j = 0..n$$

Perceptron

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Trường hợp không khả tách

★ Tìm siêu phẳng tốt nhất có thể

★ Cực tiểu hóa lỗi phân lớp

★ Hàm lỗi:

$$E(w) \equiv E(w_0, w_1, \dots, w_n) = \frac{1}{2} \sum_{i=1}^m \left(y^{(i)} - g(x^{(i)}) \right)^2$$

★ Tìm w để cực tiểu hóa E

Perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Trường hợp không khả tách

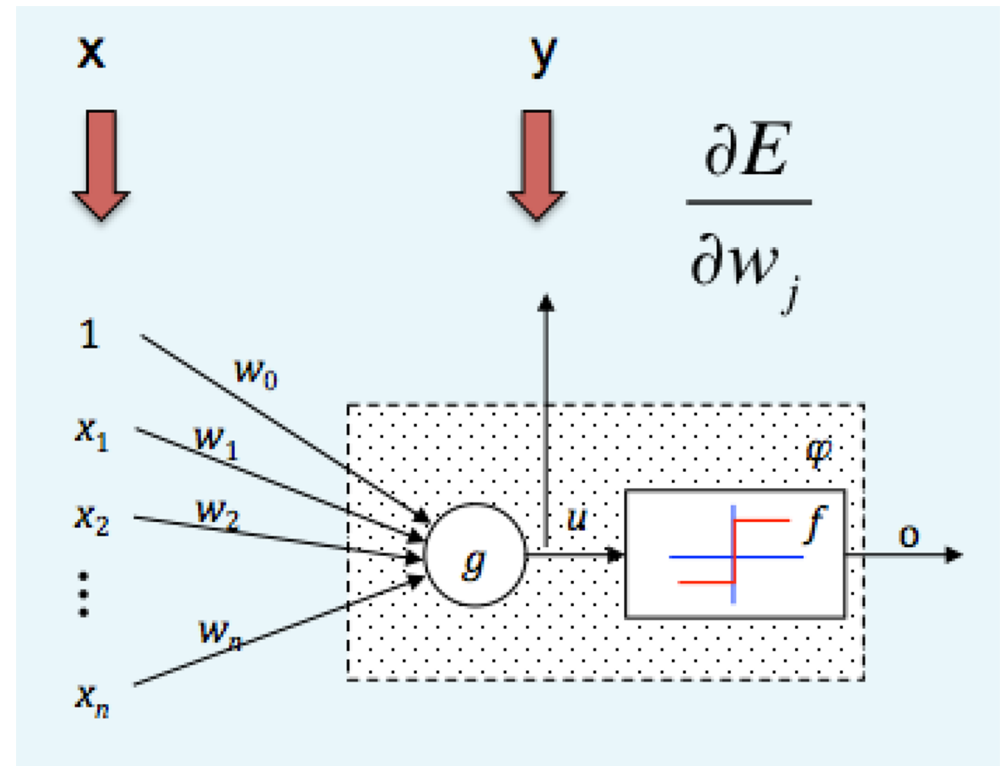
★ Giảm gradient (standard)

★ Khởi tạo ngẫu nhiên w_j

★ Cập nhật w_j

$$w_j = w_j - \eta \frac{\partial E}{\partial w_j}$$

$$\frac{\partial E}{\partial w_j} = - \sum_{i=1}^m (y^{(i)} - g(x^{(i)})) x_j^{(i)}$$



Perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Trường hợp không khả tách

★ Giảm gradient ngẫu nhiên

★ Thay đổi luật cập nhật w_j (luật Delta)

1. Chọn ngẫu nhiên 1 phần tử $x^{(i)}$ và cập nhật w_j :

$$w_j = w_j - \eta \frac{\partial E}{\partial w_j}$$

$$\frac{\partial E}{\partial w_j} = -\left(y^{(i)} - g(x^{(i)})\right) \cdot x_j^{(i)}$$

2. Lặp lại bước 1 cho đến khi hội tụ

Nội dung

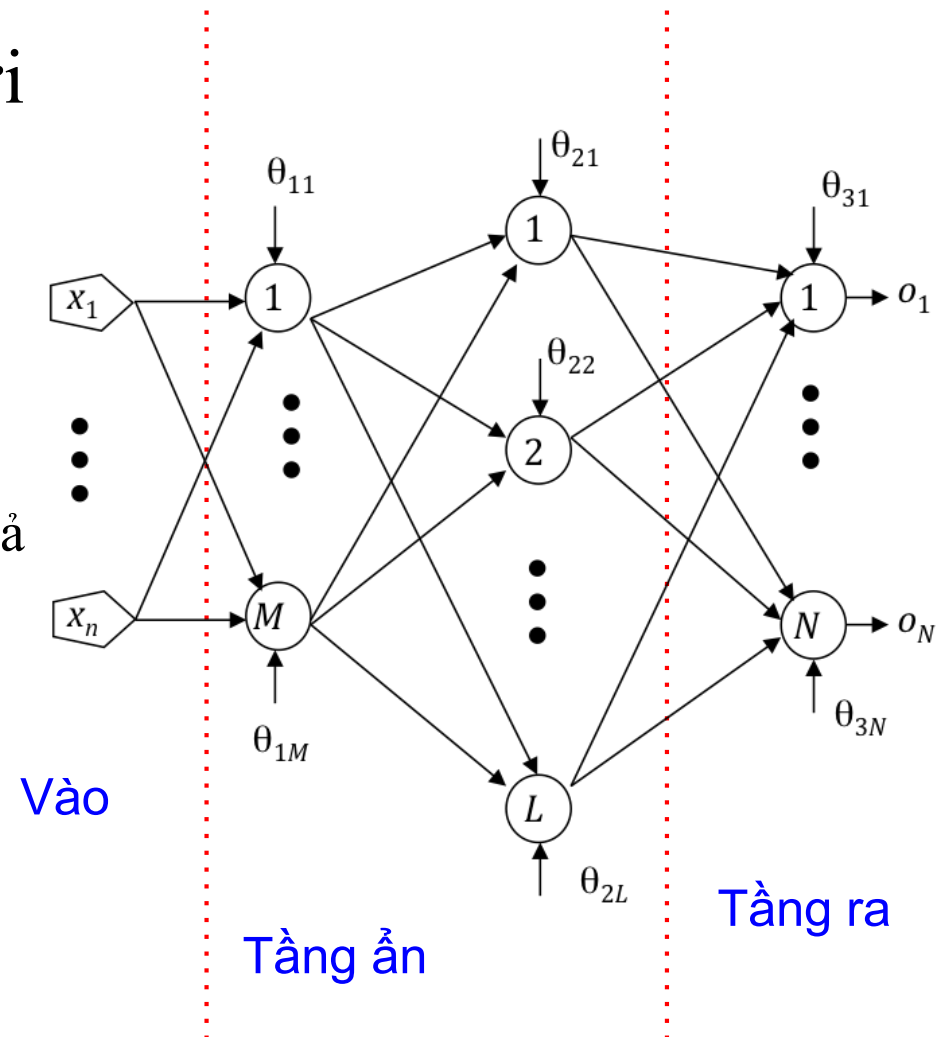
- ★ Giới thiệu
- ★ Mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

Multi-layer perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Mạng nơ-ron truyền tiến tới

- ★ 2 tầng kề nhau được kết nối đầy đủ
- ★ Mỗi nơ-ron
 - ★ Hàm kết hợp: tuyến tính
 - ★ Hàm kích hoạt: phi tuyến, khả vi liên tục
 - ★ Sigmoid, hyperbolic tangent
- ★ Các nơ-ron được bố trí trong nhiều tầng



Multi-layer perceptron

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Biểu diễn dạng ma trận

- ★ $x = (x_1, x_2, \dots, x_n)^T$: vào
- ★ k : số tầng
- ★ n_i : số nơ-ron ở tầng thứ i
- ★ $u^{(i)}$: ngõ ra trung gian của các nơ-ron tầng thứ i
- ★ $o^{(i)}$: ngõ ra của các nơ-ron tầng thứ i
- ★ $W^{(i)}$: ma trận trọng số của các nơ-ron tầng thứ i

Multi-layer perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Biểu diễn dạng ma trận

$$W^{(i)} = \begin{pmatrix} w_{10}^{(i)} & w_{11}^{(i)} & \dots & w_{1n_{i-1}}^{(i)} \\ w_{20}^{(i)} & w_{21}^{(i)} & \dots & w_{2n_{i-1}}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_i 0}^{(i)} & w_{n_i 1}^{(i)} & \dots & w_{n_i n_{i-1}}^{(i)} \end{pmatrix} \quad u^{(i)} = \begin{pmatrix} u_1^{(i)} \\ \vdots \\ u_{n_i}^{(i)} \end{pmatrix} \quad o^{(i)} = \begin{pmatrix} o_1^{(i)} \\ \vdots \\ o_{n_i}^{(i)} \end{pmatrix}$$

Ma trận trọng số của tầng thứ i

Ngõ ra trung gian u

Ngõ ra o

$$u^{(1)} = W^{(1)} \begin{pmatrix} 1 \\ x \end{pmatrix} \quad u^{(i)} = W^{(i)} \begin{pmatrix} 1 \\ o^{(i-1)} \end{pmatrix} \quad o_j^{(i)} = f(u_j^{(i)})$$

Multi-layer perceptron

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

★ Giải thuật huấn luyện mạng nơ-ron đa tầng (MLP)

★ Hàm lỗi

$$E(W) \equiv E(W^{(1)}, W^{(2)}, \dots, W^{(k)}) = \frac{1}{2} \sum_{i=1}^m \left(y^{(i)} - g(x^{(i)}) \right)^2$$

★ Tìm những W để cực tiểu hóa lỗi E

★ Giải thuật lan truyền ngược (back propagation)

Multi-layer perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Giải thuật lan truyền ngược

★ Phương pháp giảm gradient

1. Khởi tạo ngẫu nhiên các W

2. Cập nhật W theo luật

$$W = W - \eta \nabla E(W)$$

Lặp lại bước 2 cho đến khi hội tụ

★ Tính toán các đạo hàm riêng

Multi-layer perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Tính toán các đạo hàm riêng

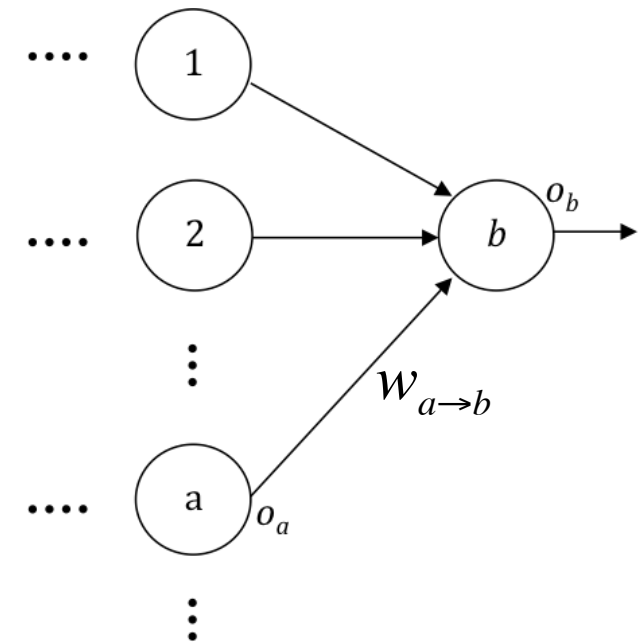
★ Các nơ-ron của tầng ra

$$E = \frac{1}{2}(y_b - o_b)^2 \quad o_b = f_b(u_b)$$

$$u_b = \sum_a (w_{a \rightarrow b} o_a)$$

$$\frac{\partial E}{\partial o_b} = \frac{\partial}{\partial o_b} \left(\frac{1}{2}(y_b - o_b)^2 \right) = -(y_b - o_b)$$

$$\frac{\partial E}{\partial w_{a \rightarrow b}} = -(y_b - o_b) \cdot f_b'(u_b) \cdot o_a$$



Multi-layer perceptron

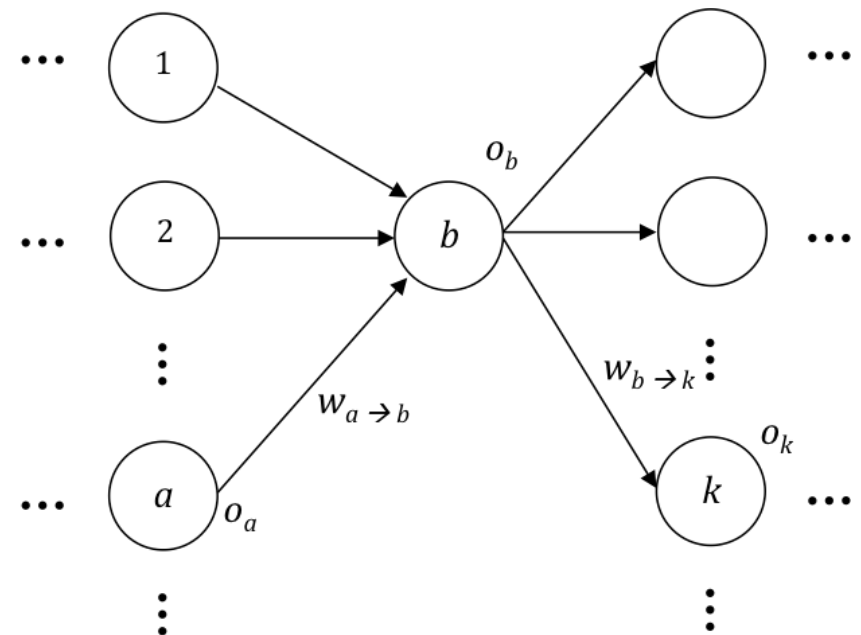
- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Tính toán các đạo hàm riêng

★ Các nơ-ron của tầng ẩn

$$\frac{\partial E}{\partial w_{a \rightarrow b}} = \frac{\partial E}{\partial o_b} \cdot f'_b(u_b) \cdot o_a$$

$$\frac{\partial E}{\partial o_b} = \sum_k \left(\frac{\partial E}{\partial o_k} \cdot f'_k(u_k) \cdot w_{b \rightarrow k} \right)$$



Multi-layer perceptron

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

★ Luật cập nhật các trọng số

★ Phương pháp giảm gradient

★ gradient được tính trên toàn bộ tập huấn luyện

$$\frac{\partial E}{\partial w_{a \rightarrow b}} = \sum_{i=1}^m \frac{\partial E}{\partial o_b} \cdot f_b'(u_b) \cdot o_a$$

★ Phương pháp giảm gradient ngẫu nhiên

★ gradient được tính trên 1 phần tử chọn ngẫu nhiên từ tập huấn luyện

Multi-layer perceptron

★ Biểu diễn dạng ma trận

★ Các nơ-ron của tầng ra

$$\frac{\partial E}{\partial W^{(k)}} = - \left[\left(y - o^{(k)} \right) .* f'(u^{(k)}) \right] * \left(o^{(k-1)} \right)^T$$

★ y : giá trị ra muốn có (véc-tơ)

★ $o^{(k)}$: ngõ ra của các nơ-ron của tầng ra (tầng k)

★ $.*$: phép nhân element-wise (element-by-element)

★ Nếu hàm kích hoạt là sigmoid thì

$$f'(u^{(k)}) = \frac{o^{(k)} .* (1 - o^{(k)})}{T}$$

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

Multi-layer perceptron

- ★ Biểu diễn dạng ma trận
 - ★ Các nơ-ron của tầng ẩn

$$\frac{\partial E}{\partial W^{(i)}} = \left[\frac{\partial E}{\partial o^{(i)}} \cdot * f'(u^{(i)}) \right] * \left(o^{(i-1)} \right)^T$$

với

$$\frac{\partial E}{\partial o^{(i)}} = W^{(i+1)} * \left(\frac{\partial E}{\partial o^{(i+1)}} \cdot * f'(u^{(i+1)}) \right)$$

Nội dung

- ★ Giới thiệu
- ★ Mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ Kết luận

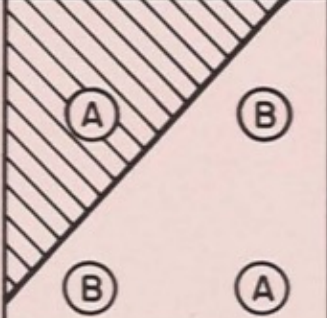
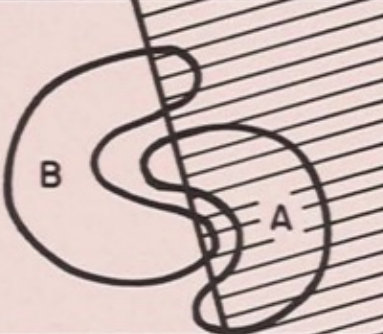
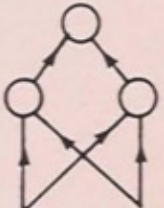
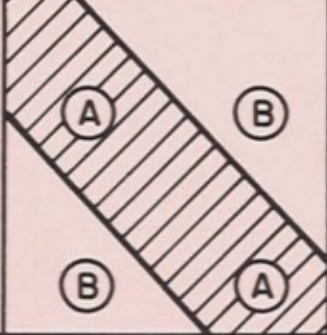
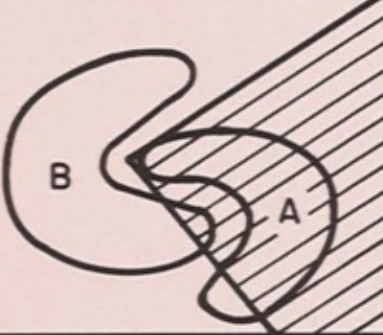
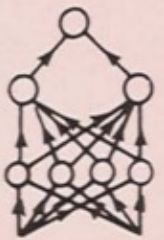
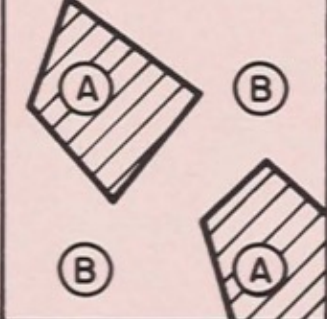
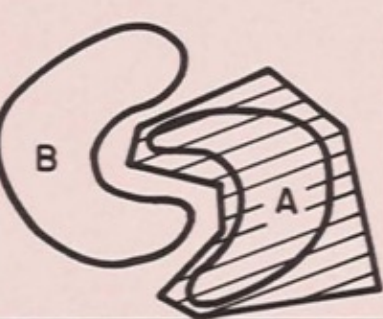
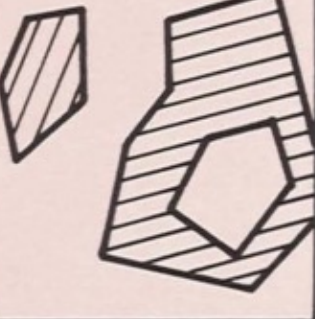
Kết luận

- ★ Giới thiệu
 - ★ Nơ-ron McCulloch & Pitts
 - ★ Perceptron
 - ★ Multi-layer perceptron (MLP)
 - ★ Kết luận
-

- ★ Mô hình nơ-ron McCulloch & Pitts
- ★ Perceptron: đơn giản và đa tầng
- ★ Giải thuật giảm gradient
- ★ Giải thuật lan truyền ngược

Kết luận

- ★ Giới thiệu
- ★ Nơ-ron McCulloch & Pitts
- ★ Perceptron
- ★ Multi-layer perceptron (MLP)
- ★ **Kết luận**

Structure	Types of Decision Regions	Exclusive-or Problem	Classes with Meshed Regions	Region Shapes
One Layer 	Half-Plane			
Two Layers 	Typically Convex			
Three Layers 	Arbitrary			



Merci !