

# Huấn luyện perceptron

## Trường hợp dữ liệu **không** khả tách tuyến tính

- Cố gắng tìm một siêu phẳng “tốt” nhất
  - **Tốt = lỗi (trên tập học) nhỏ nhất có thể**
- Định nghĩa hàm lỗi  $E(w)$  theo các trọng số  $w$  trên tất cả các phần tử của tập học:

$$E(w) \equiv E(w_0, w_1, \dots, w_n) = \frac{1}{2} \sum_{i=1}^m (y_i - \varphi(x_i))^2$$

- **Bài toán huấn luyện trở thành tìm  $w$  sao cho  $E(w)$  nhỏ nhất**

26

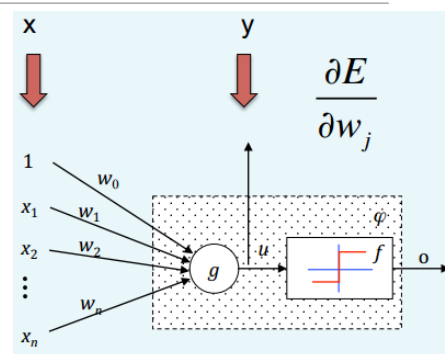
# Huấn luyện perceptron

Chú ý:

- Hàm kích hoạt trong trường hợp này được thay bằng hàm tuyến tính  $f(u) = u$  hay:

$$\varphi(x_i) = g(x_i) = \sum_{j=0}^n w_j \cdot x_{ij}$$

Ta tối ưu  $E(w)$  bằng phương pháp gradient descent



$$o = f(u) = f(g(x)) = \begin{cases} 1 & g(x) \geq 0 \\ 0 & g(x) < 0 \end{cases}$$

27

## Huấn luyện perceptron

Giải thuật:

- Khởi động ngẫu nhiên  $w$
- Lặp cho đến khi điều kiện dừng thoả mãn:

- $w = w - \eta \Delta E(w)$   
 $= w + \Delta w$  (Với  $\Delta w = -\eta \Delta E(w)$ )

Gradient của  $E(w)$ :

$$\frac{\partial E}{\partial w_j} = - \sum_{i=1}^m (y_i - g(x_i)) \cdot x_{ij}$$

$$\nabla E(w) = \begin{pmatrix} \frac{\partial E}{\partial w_0} \\ \frac{\partial E}{\partial w_1} \\ \vdots \\ \frac{\partial E}{\partial w_n} \end{pmatrix}$$

28

## Huấn luyện perceptron

Luật cập nhật  $w$ :

- **Luật gradient chuẩn**: cập nhật sau khi xem xét tất cả các phần tử của tập học, cộng dồn  $\Delta w$
- **Luật Delta**: cập nhật sau khi xét mỗi phần tử của tập học

29

# Huấn luyện perceptron

Luật cập nhật w:

- **Luật gradient chuẩn:** cập nhật sau khi xem xét tất cả các phần tử của tập học, cộng dồn  $\Delta w$

**Đầu vào:**  
Tập mẫu huấn luyện  $X = \{(x_i, y_i)\}_{i=1, m}$  với  $x_i \in R^n$  và  $y_i \in \{1, -1\}$   
Tốc độ học:  $\eta$   
Dung sai:  $\varepsilon$  (lỗi lớn nhất có thể chấp nhận được)

**Giải thuật:**  
Khởi tạo ngẫu nhiên các trọng số  $w_i, \forall i = 0..n$   
**repeat**  
   $E = 0$   
  **for**  $j = 0$  to  $n$  **do**  
     $\Delta w_j = 0$   
  **end for**  
  Xáo trộn ngẫu nhiên tập mẫu huấn luyện  $X$   
  **for**  $i = 1$  to  $m$  **do**  
     $E = E + (y_i - \phi(x_i))^2$   
    // Tích lũy lượng cần cập nhật  $\Delta w_j$  cho  $w_j$   
    **for**  $j = 0$  to  $n$  **do**  
       $\Delta w_j = \Delta w_j + \eta \cdot (y_i - \phi(x_i)) \cdot x_i$   
    **end for**  
  **end for**  
  // Cập nhật trọng số  $w_j$   
  **for**  $j = 0$  to  $n$  **do**  
     $w_j = w_j + \Delta w_j$   
  **end for**  
**until**  $E < \varepsilon$

# Huấn luyện perceptron

Luật cập nhật w:

- **Luật Delta:** cập nhật sau khi xét mỗi phần tử của tập học

**Đầu vào:**  
Tập mẫu huấn luyện  $X = \{(x_i, y_i)\}_{i=1, m}$  với  $x_i \in R^n$  và  $y_i \in \{1, -1\}$   
Tốc độ học:  $\eta$   
Dung sai:  $\varepsilon$  (lỗi lớn nhất có thể chấp nhận được)

**Giải thuật:**  
Khởi tạo ngẫu nhiên các trọng số  $w_i, \forall i = 0..n$   
**repeat**  
   $E = 0$   
  Xáo trộn ngẫu nhiên tập mẫu huấn luyện  $X$   
  **for**  $i = 1$  to  $m$  **do**  
     $E = E + (y_i - \phi(x_i))^2$   
    // Cập nhật  $w_j$   
    **for**  $j = 0$  to  $n$  **do**  
       $w_j = w_j + \eta \cdot (y_i - \phi(x_i)) \cdot x_i$   
    **end for**  
  **end for**  
**until**  $E < \varepsilon$