Hồi qui logistic: phân lớp

Ngô Minh Nhựt

Bộ môn Công nghệ Tri thức

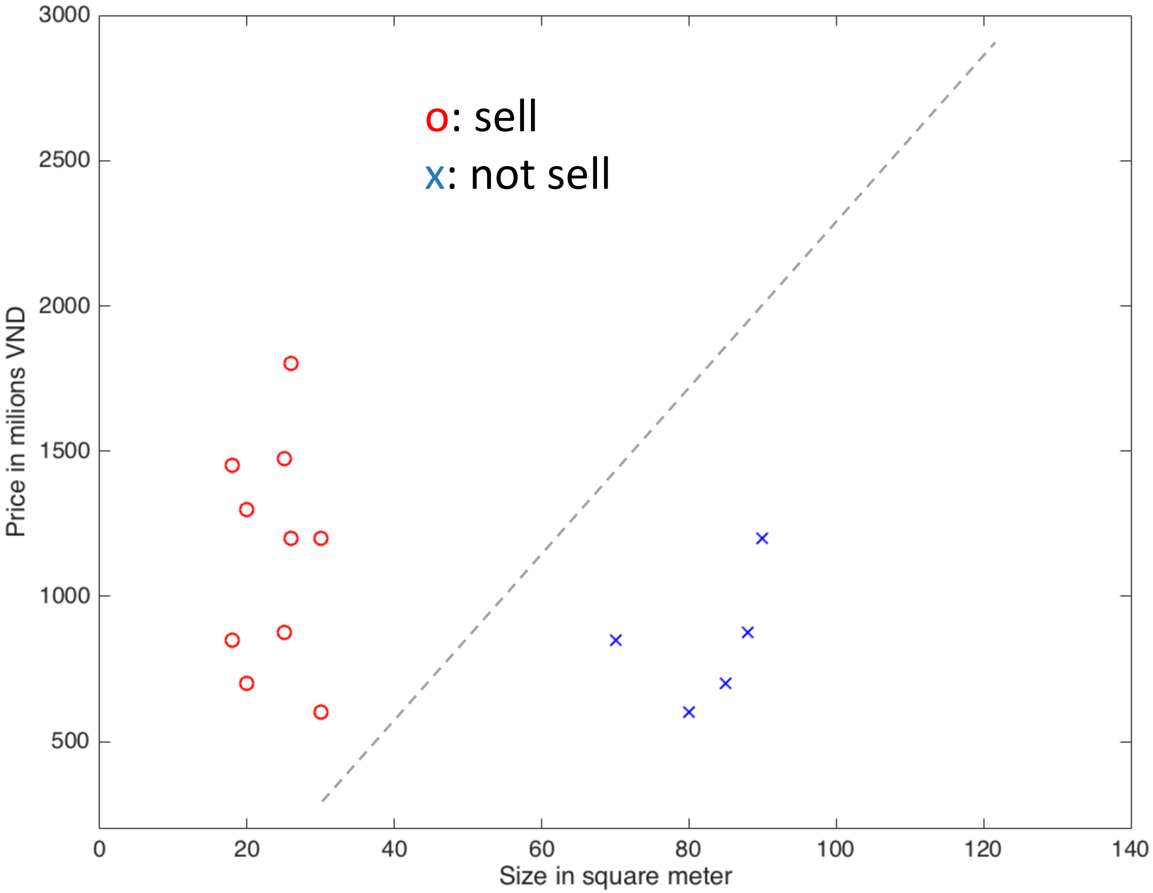
2021

Bài toán phân lớp

- Trả lời câu hỏi với **yes** hoặc **no**
  - Kiểm tra email spam
  - Kiểm tra giao dịch bất thường
  - Kiểm tra nguy cơ mắc bệnh
  - Kiểm tra 1 vùng ảnh có phải là mặt người
  - Kiểm tra 1 vùng ảnh có phải là ký tự ‘0’
  - …

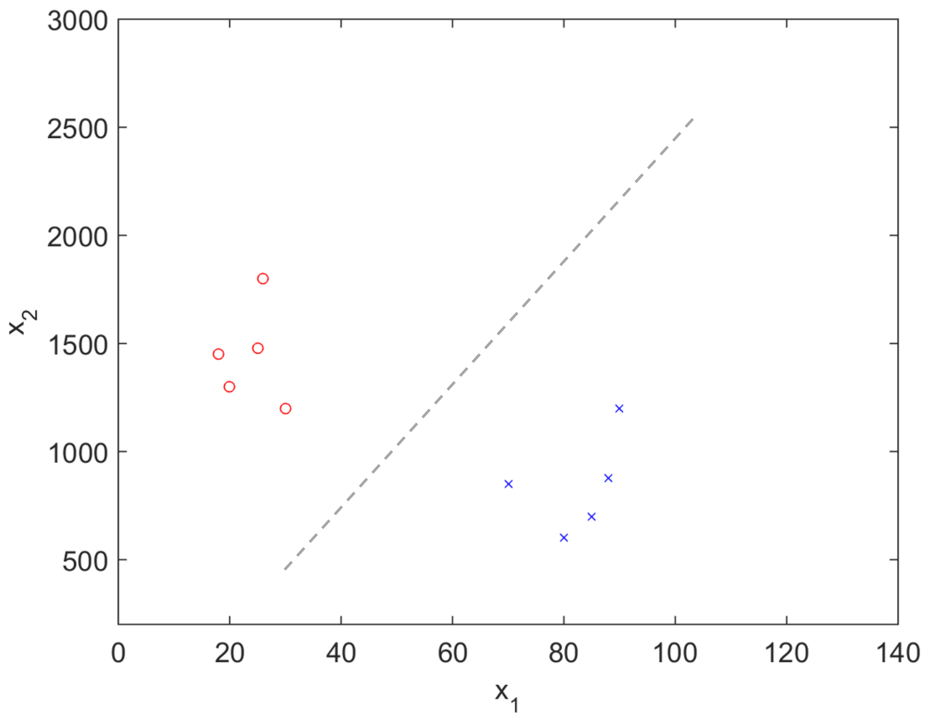
Bài toán phân lớp

Size	Price	Sell?
80	600	No
30	1200	Yes
70	850	No
26	1200	No
…	…	…



Output

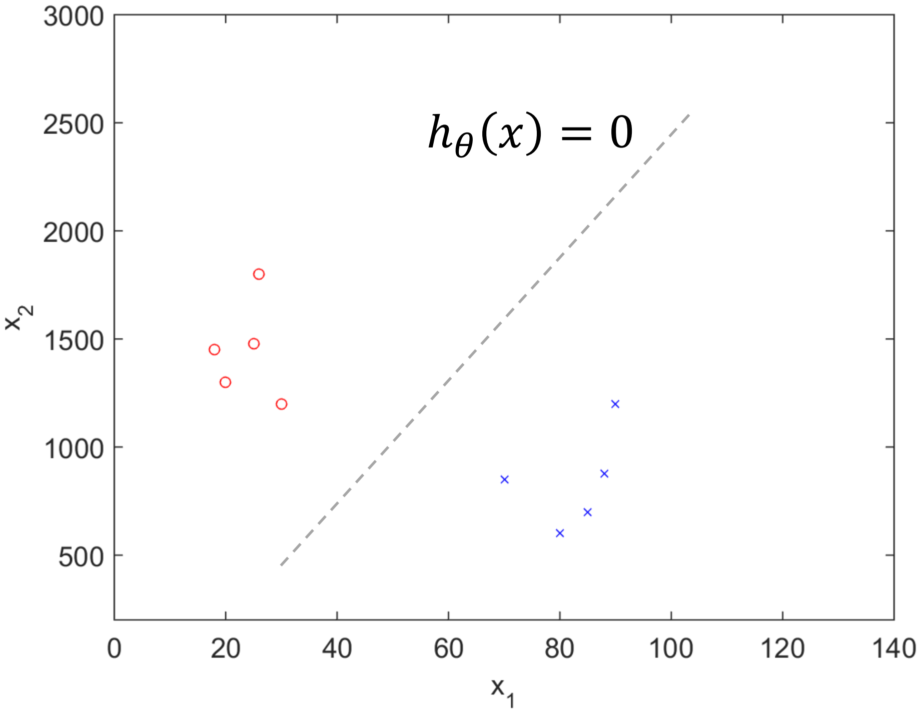
- Output = {yes, no}
- Y = {1, 0}
  - 1: lớp dương
  - 0: lớp âm



Đường phân lớp

- Phân lớp: Y = {1, 0}
- Đường phân lớp: $h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 = 0$
- Luật phân lớp

$$\begin{cases} \text{Nếu } h_\theta(x) \geq 0, y = 1 \\ \text{Nếu } h_\theta(x) < 0, y = 0 \end{cases}$$

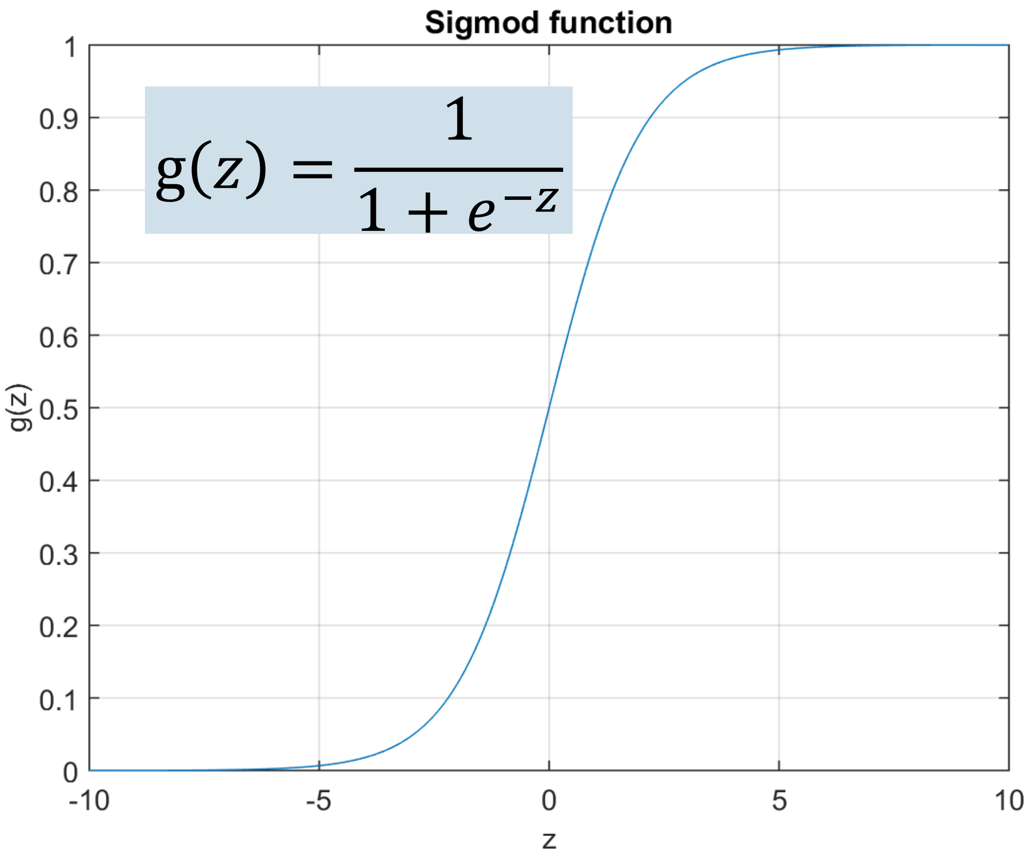


Hypothesis

- Đường phân lớp: $\theta^T x = 0$
  - $\begin{cases} \text{Nếu } \theta^T x \geq 0, y = 1 \\ \text{Nếu } \theta^T x < 0, y = 0 \end{cases}$
- Chúng ta cần: $\begin{cases} y = 1, h_\theta(x) \rightarrow 1 \\ y = 0, h_\theta(x) \rightarrow 0 \end{cases}$
- Mô hình mới: $h_\theta(x) = g(\theta^T x)$
  - $\theta^T x$ càng lớn hơn 0 thì $g(\theta^T x)$ càng tiến tới 1
  - $\theta^T x$ càng nhỏ hơn 0 thì $g(\theta^T x)$ càng tiến tới 0

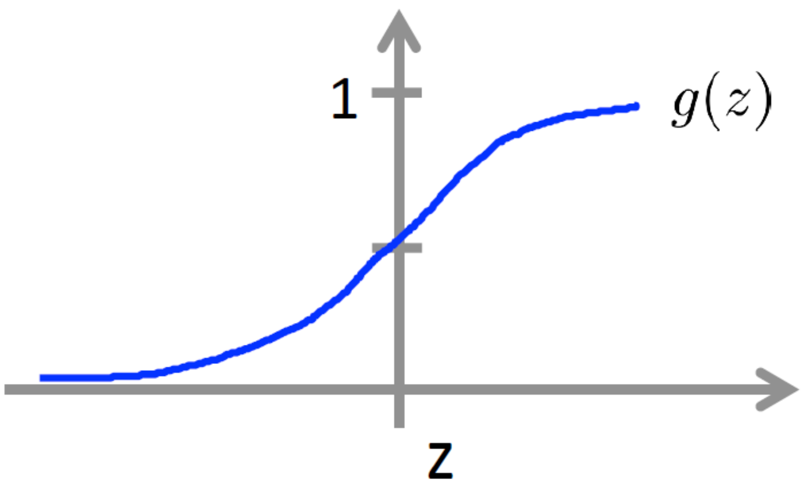
Hypothesis: hàm sigmoid

$$h_\theta(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}$$



Đường phân lớp

- $h_\theta(x) = g(\theta^T x)$
- $g(z) = \frac{1}{1+e^{-z}}$
- Đường phân lớp:
  - y = 1 nếu $h_\theta(x) \geq 0.5$
    $\rightarrow \theta^T x \geq 0$
  - y = 0 nếu $h_\theta(x) < 0.5$
    $\rightarrow \theta^T x < 0$

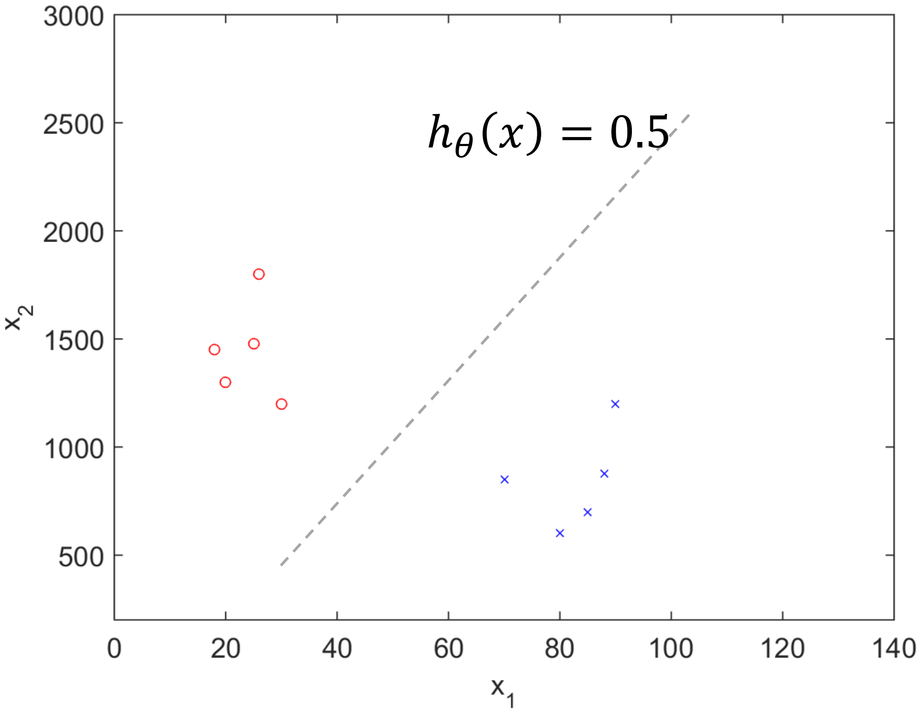


- $h_\theta(x)$ chính là xác suất y=1
  $h_\theta(x) = P(y = 1|x;\theta)$

Source: Andrew Ng

Đường phân lớp

- **Đường phân lớp:**
  - y = 1 nếu $h_\theta(x) \geq 0.5$
    ➔ $\theta^T x \geq 0$
  - y = 0 nếu $h_\theta(x) < 0.5$
    ➔ $\theta^T x < 0$

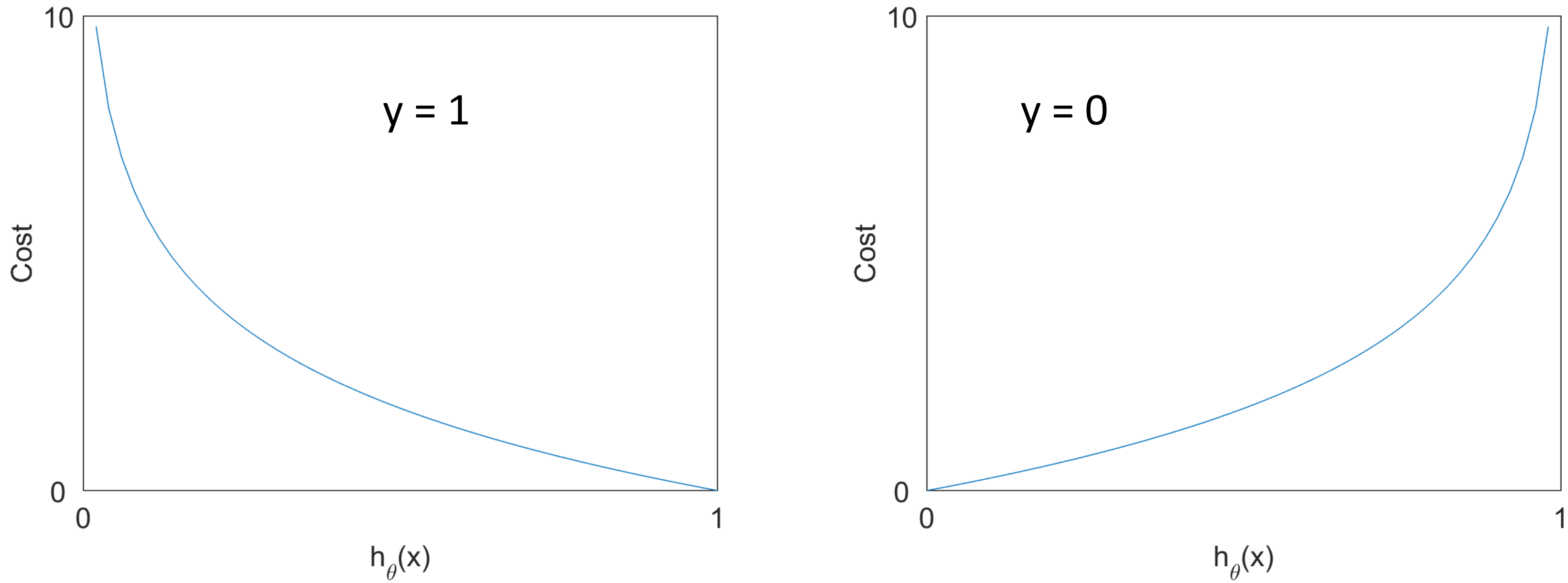


Hàm chi phí

- Hypothesis: $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$
  - $0 \leq h_\theta(x) \leq 1$
- $\mathrm{Cost}(h(x), y) = \begin{cases} -\log h_\theta(x) & \text{if } y = 1 \\ -\log\big(1 - h_\theta(x)\big) & \text{if } y = 0 \end{cases}$
  - y =1: nếu $h_\theta(x)$ càng gần 1 thì chi phí → 0, nếu $h_\theta(x)$ càng gần 0 thì chi phí → vô cùng
  - y =0: nếu $h_\theta(x)$ càng gần 0 thì chi phí → 0, nếu $h_\theta(x)$ càng gần 1 thì chi phí → vô cùng

Hàm chi phí

$$\text{Cost}(h(x), y) = \begin{cases} -\log h_\theta(x) & \text{if } y = 1 \\ -\log\big(1 - h_\theta(x)\big) & \text{if } y = 0 \end{cases}$$



Hàm chi phí

- Hypothesis: $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$
- $\text{Cost}(h(x), y) = \begin{cases} -\log h_\theta(x) & \text{if } y = 1 \\ -\log\left(1 - h_\theta(x)\right) & \text{if } y = 0 \end{cases}$
- Viết lại
- $\text{Cost}(h(x), y) = -y^{(i)} \log h\left(x^{(i)}\right) - \left(1 - y^{(i)}\right) \log\left(1 - h\left(x^{(i)}\right)\right)$
  - y = 1, 1 - y = 0 ➔ Cost(h(x), y) = ?
  - y = 0, 1 - y = 1 ➔ Cost(h(x), y) = ?

Hàm chi phí

- Hypothesis: $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$
- Hàm chi phí cho 1 mẫu

$$\mathrm{Cost}(h(x), y) = -y^{(i)} \log h\left(x^{(i)}\right) - \left(1 - y^{(i)}\right) \log\left(1 - h\left(x^{(i)}\right)\right)$$

- Hàm chi phí

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log h\left(x^{(i)}\right) + \left(1 - y^{(i)}\right) \log\left(1 - h\left(x^{(i)}\right)\right)$$

Đạo hàm riêng

- $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$
- $J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log h_\theta\left(x^{(i)}\right) + \left(1 - y^{(i)}\right) \log\left(1 - h_\theta\left(x^{(i)}\right)\right)$
- $\frac{dJ}{d\theta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$

Thuật toán hạ dốc

- Cost function:

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log h_\theta\left(x^{(i)}\right) + \left(1 - y^{(i)}\right) \log\left(1 - h_\theta\left(x^{(i)}\right)\right)$$

- Tìm θ sao cho $J(\theta)$ đạt cực tiểu
- Dự đoán cho input x mới
  - Output: $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$

Thuật toán hạ dốc

- Vector gradient:
  - $\frac{dJ}{d\theta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)}$
  - $j = 0, 1, 2, \dots, n$
- Repeat until convergence

$$\{$$

$$\theta_j = \theta_j - \alpha \frac{dJ}{d\theta_j}$$

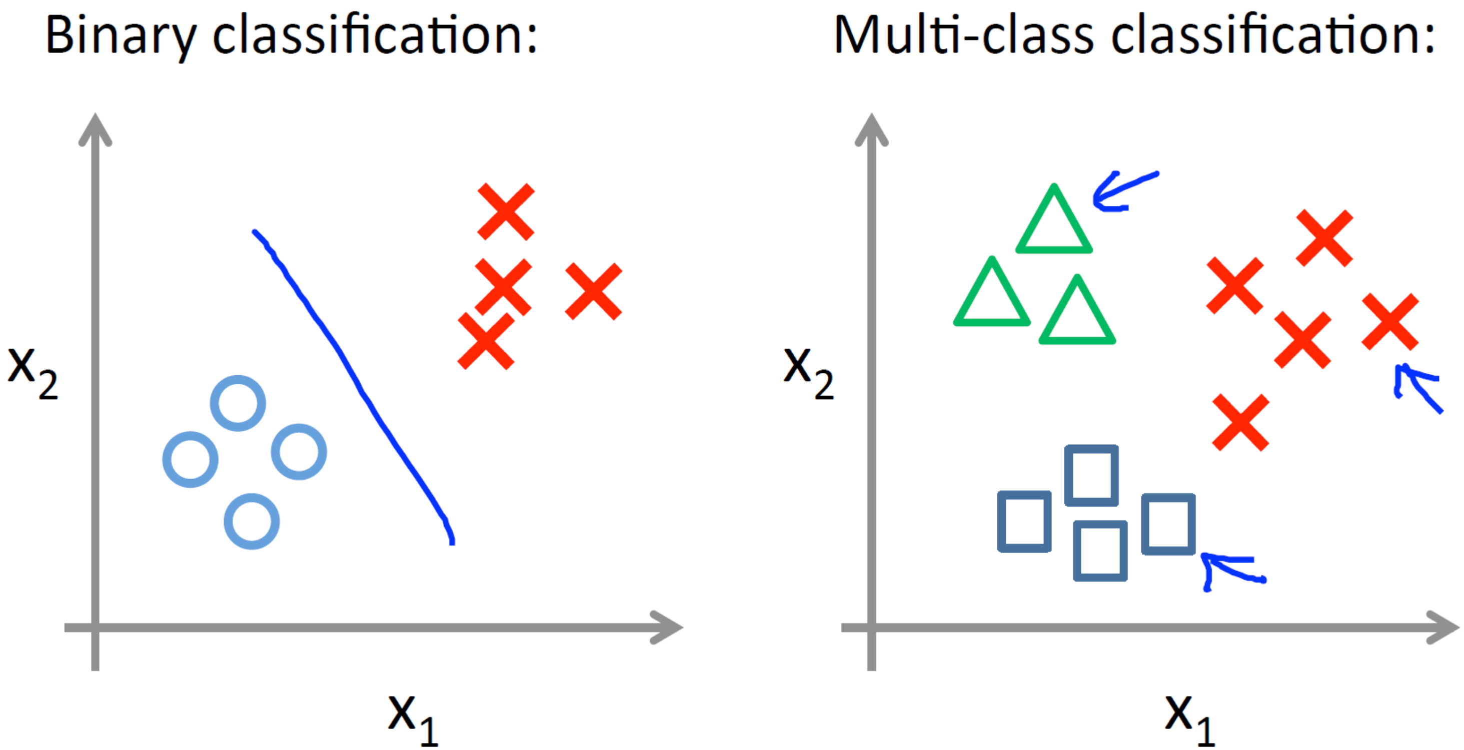
$$\}$$

Hồi qui logistic với nhiều lớp

- Thời tiết: sunny, cloudy, rain, heavy rain
- Chữ số: 0, 1, …, 9
- Đối tượng: human, cat, house, landscape

➔ y = {1, 2, 3 , …}

Hồi qui logistic với nhiều lớp

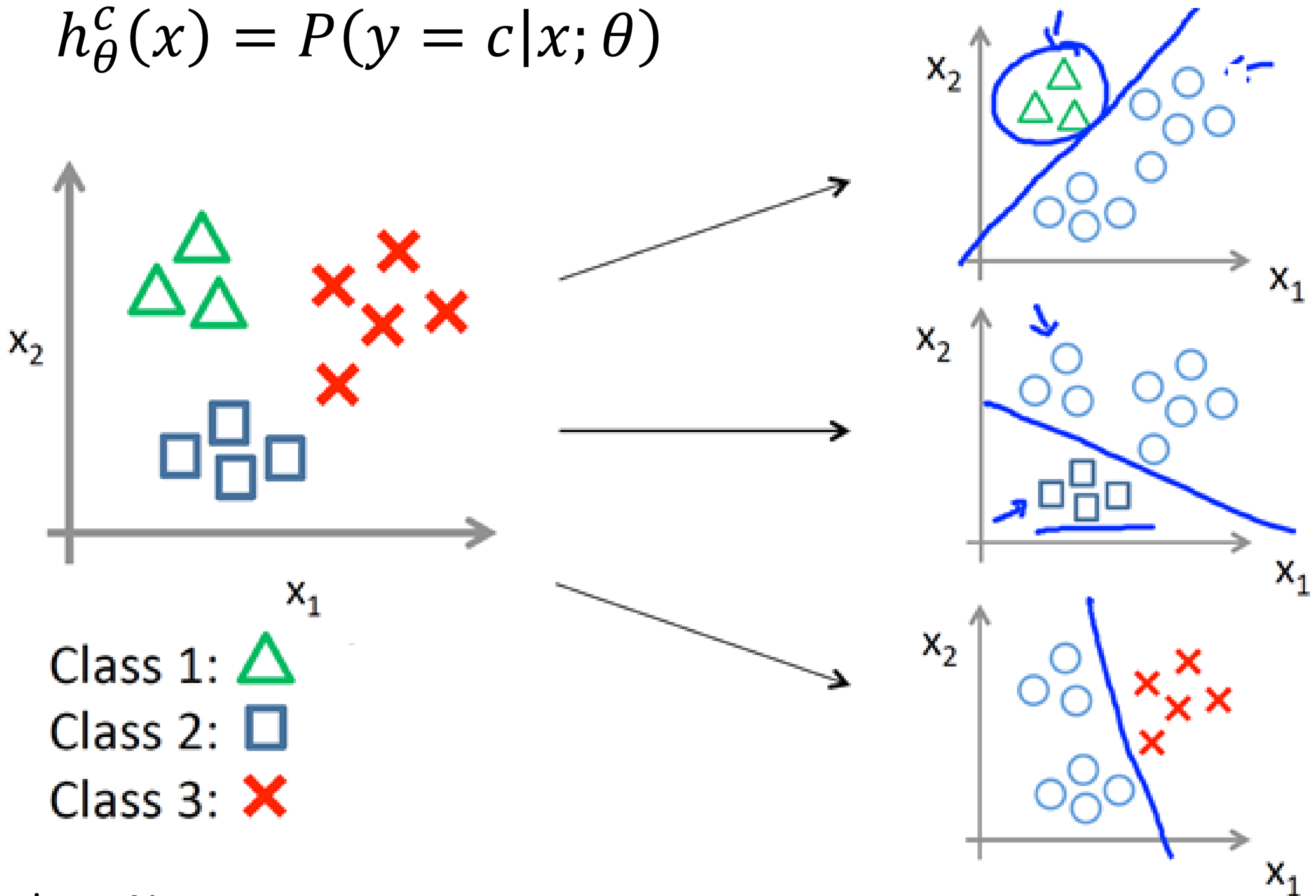


Source: Andrew Ng

Hồi qui logistic với nhiều lớp

- Huấn luyện bộ phân lớp cho từng lớp $h_\theta^c(x)$

$$h_\theta^c(x) = P(y = c|x; \theta)$$



Source: Andrew Ng

Hồi qui logistic với nhiều lớp

- Huấn luyện bộ phân lớp cho từng lớp $h_\theta^c(x)$

$$h_\theta^c(x) = P(y = c|x;\theta)$$

- Dự đoán cho input mới
  - $y = \max_c h_\theta^c(x)$