

## second day

Phạm Quốc Trung

02/07/2019

### 1 mô hình chung của machine learning

- $f(x) = y$
- Đầu vào  $x$  được đặt trong bởi những không gian hàm số
- đầu ra  $y$  (coordinate) được biểu diễn bằng phương trình tọa độ, từ đầu ra có thể hình thành những khái niệm mới
- Hàm  $f(x)$  được gọi là hàm đặc trưng (basic function) - Magic model: lấy chiết xuất đặc trưng từ 2 tập input và output rồi đưa đến một không gian có thể so sánh 2 hàm rồi chọn ra output phù hợp
- Xét hàm  $f = (f_1, f_2, f_3, \dots)$
- thì  $X = y_1 * x_1 + y_2 * x_2 + y_3 * x_3 + \dots$
- > chỉ ra số phần của  $x_i$  trong  $X$

### 2 Principal Component Analysis

- $X = x_0 + a_1 * x_1 + a_2 * x_2 + \dots + a_k * x_k$
- $y = x_0 + a_{21} * x_1 + a_{22} * x_2 + \dots + a_{2k} * x_k$
- $x_0$  là mean
- khi đó tọa độ
- $X = (a_1, a_2, a_3, \dots)$
- $Y = (a_{21}, a_{22}, a_{23}, \dots)$
- bài toán tính  $p(X, Y)$
- Cách 1:
- tính  $\langle X, Y \rangle$  : inner product = Tổng:  $\sqrt{\sum_{ij}^n (X_{ij} * Y_{ij})}$
- cách 2:
- tính  $|X - Y| = \Delta x = \sqrt{\sum_{ij}^n (X_{ij} - Y_{ij})^2}$

### 3 Linear regression

- Mục tiêu:
- Thường được dùng để giải những bài toán có output là 1 số thực

Lấy ví dụ về định giá một ngôi nhà với độ lớn  $x_1$ , độ mới  $x_2$  và có hiện trạng hư hỏng  $x_3$

- Ta nhận xét với  $x_1, x_2$  càng lớn và  $x_3$  càng thấp thì giá của căn nhà càng cao  
 $\Rightarrow$  có dạng chung

Ta có  $y \approx f(x) = W_1 * x_1 + W_2 * x_2 + w_3 * x_3 + w_0$

Mục tiêu là tìm ra  $W = [W_0, w_1, w_2, w_3]$  tối ưu

### 3.1 Công thức toán

-Cho  $W = [W_0, w_1, w_2, w_3]^T$  (ct)

-Cho  $X = [1, x_1, x_2, x_3]$  (hàng)

thì  $y \approx \bar{x} * W = \hat{y}$

Lưu ý : Bài toán Linear regression có thể giải được bằng phương pháp đạo hàm

## 4 Logistic regression

- là mô hình có đầu ra là dạng xác suất và thích hợp cho việc giải các bài toán Binary classification -Mô hình đầu ra dự đoán Logistic regression :

$-f(x) = \theta((w^t) * x)$

Hàm sigmoid được sử dụng :  $f(s) = \frac{1}{1+e^{-s\sigma(s)}}$  để đưa giới hạn các giá trị về dạng các giá trị (0,1)

## 5 Softmax regression

-Cho ví dụ là bài toán phân loại ảnh với input là ảnh và output là nhóm mà ảnh được phân vào được biểu diễn dưới dạng vecto xác suất

ví dụ như :

1  
0  
0  
0

1 .. -Điều kiện của vecto xác suất :

$$\begin{cases} 0 \leq p_i \leq 1 \\ \sum_{i=1}^n p_i = 1 \end{cases}$$

Để giải bài toán phân loại ảnh trên ta sẽ tìm softmax *function* để chuyển dữ liệu đầu vào thành 1 vecto xác suất

Ở đây ta có thể dùng softmax regression với công thức "

$$a_i = \frac{\exp(z_i)}{\sum_{j=1}^C \exp(z_j)}$$

sau đó có thể dùng cross entropy để so sánh với vecto xác suất của nhóm ảnh với công thức :  $-\sum_{i=1}^n p_i \log P_{\hat{y}}^i$

## 6 Multilayer Perceptron

-Mục đích : được sử dụng trong những trường hợp khó có thể phân lớp được bằng những đường thẳng tuyến tính

cho  $\gamma(W * a)$  là 1 hàm phi tuyến tính

trong đó  $a = \gamma(W * Z)$  tức là :

từ  $Z$  qua một hàm phi tuyến tính là  $\gamma(W * Z)$  để trở thành  $a$  , rồi từ  $a$  qua softmax function để trở thành vecto xác suất cũng là output của bài toán  $y$

- $Z$  được gọi là input layer ,  $a$  là hidden layer ,  $y$  là output layer . sao khi đã có  $y$  ta làm tương tự như softmax regression , dùng cross entropy để so sánh với các class probability vector , rồi tìm đường decision boundary tối ưu nhất