

MaSSP Lecture 2

July 2019

1 Supervised-learning:

Dùng để dự đoán output dựa vào input đã biết

Example: Định giá nhà

- input: ảnh \rightarrow output: price (R)
- Basis function: vị trí, tuổi, kích thước... \Rightarrow Mỗi basis biểu diễn trên 1 trục
 \Rightarrow Từ các tọa độ tạo thành vector $z \rightarrow$ Từ z map vào trục price để định giá

2 Linear Regression

- Linear function: Ma trận biến đổi tuyến tính bằng cách nhân với 1 vector:
 $Z^T.W = \hat{y}$
- So sánh difference: $d^2 = (y - \hat{y})^2$
- Performance measure:
 $W = \operatorname{argmin} \sum_{T=1}^N d(y, \hat{y}) = \operatorname{argmin} \sum_{T=1}^N (y - \hat{y})^2$

3 Linear Classification:

Example: Nhận dạng ảnh

- input: ảnh
- output: nhận dạng người/động vật/trái cây

Sử dụng vector xác suất(p):

- người : $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

- thú: $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

- trái cây: $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Ví dụ:

$$p = \begin{bmatrix} 0.3 \\ 0.6 \\ 0.1 \end{bmatrix}$$

$$\Rightarrow \text{Tổng quát: } p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

input đi qua basis function G_1 (designed) $\Rightarrow Z_{G_1}$.

$$Z_{G_1} \text{ đi qua } G_2 : w.z = \hat{y} \Rightarrow \hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \dots \\ \hat{y}_d \end{bmatrix} \text{ so sánh với: } y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_d \end{bmatrix}$$

Tính difference:

- $d(y, \hat{y}) = \sqrt{\sum_{i=1}^d (\hat{y}_i - y_i)^2}$

- $d(y, \hat{y}) = \sum_{i=1}^d y_i \hat{y}_i$

- $d(y, \hat{y}) = - \sum_{i=1}^d p_{y_i} \cdot \log p_{\hat{y}_i}$ (cross-entropy)
trong đó p_{y_i} là xác suất thực, $p_{\hat{y}_i}$ là xác suất tính được