

MASSP

Nguyễn Thiện Nhân
Phổ Thông Năng Khiếu - T1619

2 July 2019

1 Mô hình chung

Mô hình chung là tìm hàm $f : X \rightarrow Y$ với X là bộ cơ sở dữ liệu và Y là các đầu ra sao cho hàm f với bộ dữ liệu X thì cho đầu ra Y hiệu quả nhất.

Một mô hình chung có thể biểu diễn như sau :

$$\begin{aligned} X &\xrightarrow{h_1} Z_{h_1} \xrightarrow{h_2} Z_X \textbf{(1)} \\ Y &\xrightarrow{g_1} Z_{g_1} \xrightarrow{g_2} Z_Y \textbf{(2)} \end{aligned}$$

Với h_1, g_1, h_2, g_2 là các **basis function**.

Hàm h_2, g_2 dùng để đưa các Z_{g_1}, Z_{h_1} và tạo ra được tương ứng Z_X, Z_Y vào không gian hàm chung. Từ đó đưa ra so sánh giữa Z_X, Z_Y . Để so sánh ta dùng

$$d(Z_X, Z_Y) = P(X, Y_{g_1, g_2, h_1, h_2})$$

Trong đó P là hàm để đo **Performance Measure**

Có nhiều cách tính d và P , chẳng hạn

$$P(X_1, X_2) = \sqrt{\sum (x_{ij_1} - x_{ij_2})^2} \textbf{(3)}$$

$$P(X_1, X_2) = \sqrt{\sum x_{ij_1} \cdot x_{ij_2}} \textbf{(4)}$$

2 Linear Regression

Ta sẽ lấy ví dụ theo mô hình trên.

Lấy X là tập hợp các tính chất của một căn nhà ví dụ vị trí địa lý, kích thước, độ bền,....

Y là tập hợp giá trị của căn nhà.

Như mô tả ở mô hình chung, khi đưa X qua basis function h_1 sẽ cho ta được các tọa độ (coordinates).

Từ các tọa độ, bước h_2 sẽ là từ Z ta nhân vô hướng với một vector w cột ví dụ như

$$w = [r_1, r_2, \dots, r_N]^T$$

để dự đoán giá nhà \hat{y} và so sánh nó với giá trị thực (linear). Ở dạng công thức toán ta có thể viết là

$$\hat{y} = Z^T w$$

Từ đây ta điều chỉnh $d(y, \hat{y})$ (**Performance Measure**) sao cho sai số của đầu ra đạt giá trị nhỏ nhất. Hay nói cách khác là ta cần tìm w^* sao cho

$$w^* = \operatorname{argmin} \sum_{i=1}^N d(y, \hat{y})$$

Với tập dữ liệu $D = \{(x_i, y_i)\}_{i=1}^N$

3 Linear Classification

Một ví dụ của Linear Classification là phân loại các hình theo dạng người, thú nuôi, trái cây,... Với đầu ra là vector

$$\hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \cdot \\ \cdot \\ \hat{y}_N \end{bmatrix}$$

Với \hat{y} được tính bằng **softmax** thông qua hàm **sigmoid** và mỗi phần tử trong \hat{y} được đảm bảo rằng

$$\hat{y}_i = q_i = \frac{\sigma(q_k)}{\sum_{i=1}^N \sigma(q_i)}$$

Và **cross_entropy** được dùng để tính Performance Measure của việc so sánh

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