

# Notes 2 tháng 7

Nguyễn Nga Nhi

Ngày 2 tháng 7 năm 2019

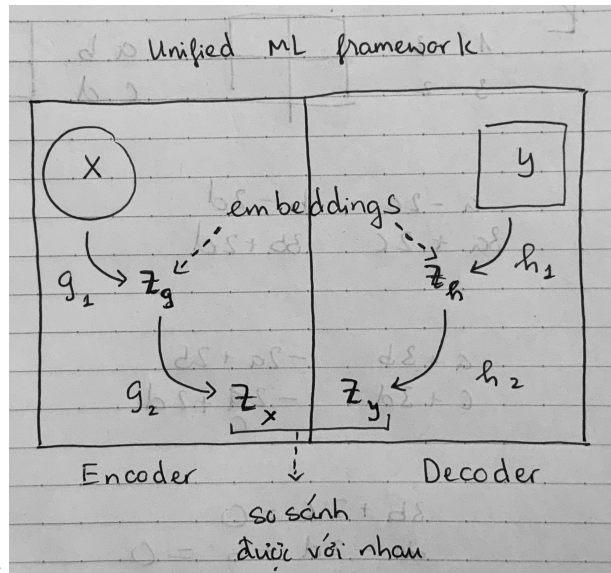
## 1 Machine Learning

Mô hình tổng quát: TEFPA

- + T = Task
- + E = Experience
- + F = Function
- + P = Performance
- + A = Algorithm

Machine Learning Framework: Simplified + Unified

- + Simplified = đơn giản hóa
- + Unified = đa dạng input



ML Framework.jpg

- + X = data
- + Y = label

- +  $g_1, h_1$  = hàm trích xuất
- +  $Z_g, Z_h$  = embedding / vector coordinate
- +  $Z_x, Z_y$  = thành phần tương ứng có thể so sánh
- +  $g_2, h_2$  = thuật toán sao cho  $Z_x, Z_y$  gần nhau nhất

## 2 Basis functions

Hàm đặc trưng: Trích xuất đặc trưng của input

+ output: vector coordinate Z

(1 basis  $\implies$  1 axis  $\implies$  1 coordinate)

## 3 Principle Component Analysis

PCA

Giảm chiều dữ liệu: Lựa chọn k thành phần chính

Biểu diễn:

$$X = X_0 + z_1X_1 + z_2X_2 + \dots + z_kX_k$$

+  $X_0$  = TBC training data

+  $X_i$  = basis function

+  $z_i$  = coordinate

So sánh khác biệt giữa 2 vector  $X_1, X_2$

+ khoảng cách =  $\sqrt{\sum (x_{1_{ij}} - x_{2_{ij}})^2}$

+ góc =  $\sqrt{\sum x_{1_{ij}}x_{2_{ij}}}$