

Machine Learning Notes: 02/10/24

? Phân nhóm các thuật toán ML: Dựa trên Phytik học (1) - learning style

Chú ý (2) - function

(1) Supervised Learning

$$X = \{x_1, x_2, \dots, x_N\}; Y = \{y_1, \dots, y_N\}$$

x_i, y_i là các vector

$(x_i, y_i) \in X \times Y$: training data

• Cài $y = f(x)$ thuật toán

↳ new input

→ Nhận dạng chủ nghĩa tay, Tay FB, Một mì vs. f

Classification
Regression

Semi-Supervised Learning

• Lg lén dữ liệu \times nhج chỉ một phần
trong chúng để gần nhau. "Internet"

Machine Learning

Unsupervised learning

④ Dựa vào cấu trúc của dữ liệu

để clustering hoặc dimension reduction
trong lưu trữ và tính toán.

Clustering

Association: bài toán Khám phá ra
một quy luật dữ liệu trên dữ liệu cho trước.

e.g.: Recommendation system

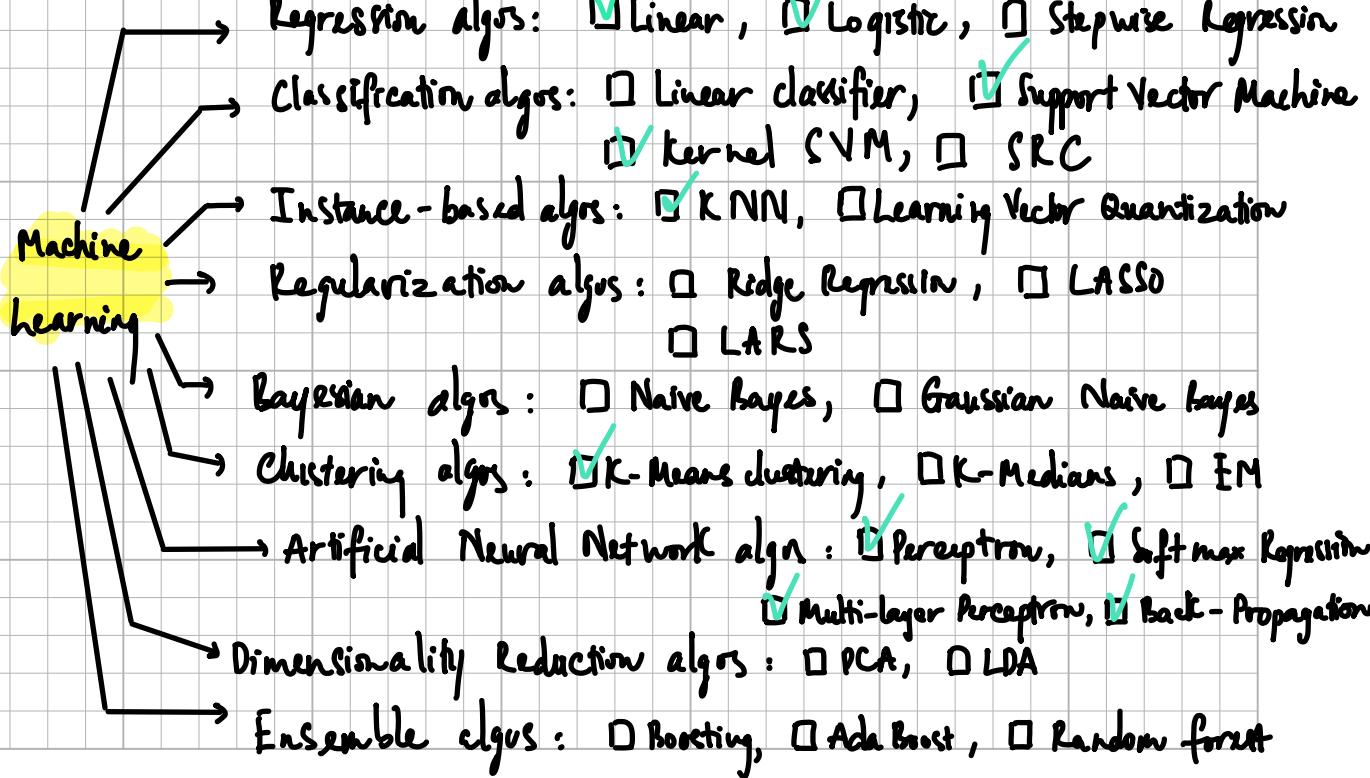
Reinforcement learning

④ bài toán giúp hệ thống tự xác hơi để
maximizing the performance.

↳ Game theory

E.g: AlphaGo, MarI/O

(2)



• Linear Regression → Supervised Learning

$$\nabla_x (f(x)^T g(x)) = \nabla_x f^T g + \nabla_x g^T f$$

$$\begin{matrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \\ \hline x_{\text{bar}} & & & \end{matrix} \quad \begin{matrix} 1 & 147 \\ 1 & 150 \\ \vdots & \vdots \\ 1 & 183 \end{matrix} \quad \begin{matrix} A = \frac{1}{n} \sum x_i \\ b = \frac{\sum y_i}{n} \\ \sum x_i \cdot y_i \end{matrix}$$

$$\begin{matrix} y_1 \\ \vdots \\ y_n \end{matrix} \quad \begin{matrix} A^T & b \\ \begin{pmatrix} (0)(1) & \sum y \\ (1)(1) & \sum x_i y_i \end{pmatrix} & \begin{matrix} w_0 \\ w_1 \end{matrix} \end{matrix} \quad \begin{matrix} : \\ \frac{(0).\sum y + (1).\sum x_i y_i}{(1).\sum y + (3).\sum x_i y_i} \end{matrix}$$

$$2 \times 2 \quad 2 \times 1$$

$$= 2 \times 1$$

• K-means Clustering → Unsupervised Learning

→ Từ dữ kiện đưa vào và số lượng nhãn, chỉ ra center của mỗi nhãn
và phân các dữ kiện vào các簇.

Ghi thêm: mỗi cf chỉ là chi tiêu vào tf whom.

• K-nearest Neighbours → Supervised Learning

• Gradient Descent

• Perceptron Learning Algorithm

• Logistic Regression

$$\frac{\partial J(w; x_i, y_i)}{\partial w} = \frac{z_i - y_i}{z_i(1-z_i)} \cdot z_i \cdot (1-z_i) \cdot x$$

$$= (z_i - y_i) \cdot x$$

$$w^T x = -\frac{w_0}{w_1}$$

• Feature Engineering

oblivious Classifiers do not train Classification

• Softmax Regression

$$\begin{aligned} J_i(w) &\triangleq J(w; x_i, y_i) = \sum_{j=1}^C y_{ji} \log \left(\frac{\exp(w_j^T x_i)}{\sum_{k=1}^C \exp(w_k^T x_i)} \right) \\ &= - \sum_{j=1}^C y_{ji} w_j^T x_i + \log \sum_{k=1}^C \exp(w_k^T x_i) \end{aligned}$$

$$\frac{\partial J_i(w)}{\partial w_j} = e_j \cdot x_i \quad (\text{Note } e_j = q_j - y_j)$$

• Multi-Layer Perception \checkmark Backpropagation

Units: from raw to z , then to a

$$w^{(l)} \in \mathbb{R}^{d^{(l-1)} \times d^{(l)}}$$

$$b^{(l)} \in \mathbb{R}^{d^{(l)} \times 1}$$

$$z_j^{(l)} = w_j^{(l)} T_a^{(l-1)} + b_j^{(l)}$$

$$z^{(l)} = w^{(l)} + a^{(l-1)} + b^{(l)}$$

$$a^{(l)} = f(z^{(l)})$$

• Mahalanobis norm

$$\|x\|_A = \sqrt{x^T A^{-1} x}$$

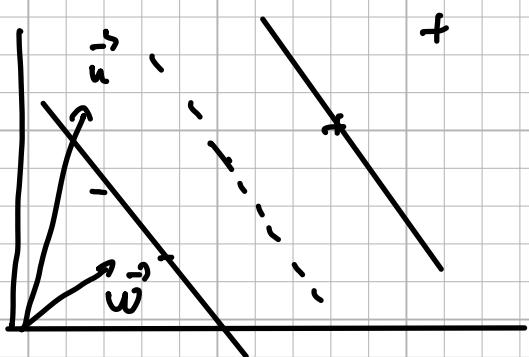
• Xuất hiện alg: $x^T A^{-1} x \geq 0$

• Nét xuất hiện: $x^T B x \geq 0$

$\leq \rightarrow$ element-wise

• Bases from duality (duality)
infimum và supremum

• Support Vector Machine



$$\bar{w} \cdot \bar{w} \geq c$$

$$\bar{w} \cdot \bar{w} + b \geq 0 \text{ Then } +$$

$$b = -c$$

$$y_i = \begin{cases} 1, & +\text{samples} \\ -1, & -\text{samples} \end{cases}$$

$$y_i (\bar{w} \cdot \vec{x}_i + b) - 1 \geq 0$$

$$\text{Width} = (\vec{x}_+ - \vec{x}_-)^T \cdot \frac{\bar{w}}{\|\bar{w}\|} = \frac{2}{\|\bar{w}\|}$$

$$\rightarrow \text{Max} \quad \frac{1}{2} \|\bar{w}\|^2$$

$$L = \frac{1}{2} \|\bar{w}\|^2 - \sum \alpha_i [y_i (\bar{w} \cdot \vec{x}_i + b) - 1]$$

$$\frac{\partial L}{\partial \bar{w}}, \frac{\partial L}{\partial b}$$

$$\bar{w} = \sum_i \alpha_i y_i \vec{x}_i$$

$$\sum \alpha_i y_i = 0$$

$$K(x_i, x_j) = \phi(\bar{x}_i) \cdot \phi(\bar{x}_j)$$

$$\textcircled{1} (\bar{u}, \bar{v} + 1)^w \quad \textcircled{2} e^{-\frac{\|\bar{x}_i - \bar{x}_j\|}{\sigma}}$$

- Kernel Support Vector Machines
- Multi-class SVM