
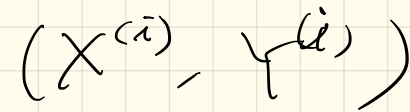


$$MSE = E \left[\left(Y - \hat{f}(x; \theta) \right)^2 \right]$$

$$= \frac{1}{n} \sum_{i=1}^n \left(\underset{\sim}{Y^{(i)}} - \underset{\sim}{f}(\underset{\sim}{X^{(i)}}; \theta) \right)^2$$



 $(X^{(i)}, Y^{(i)})$

i -th data point.

$$L_1 \text{ (ABS)} = E \left[\left| Y - f(x; \theta) \right| \right]$$



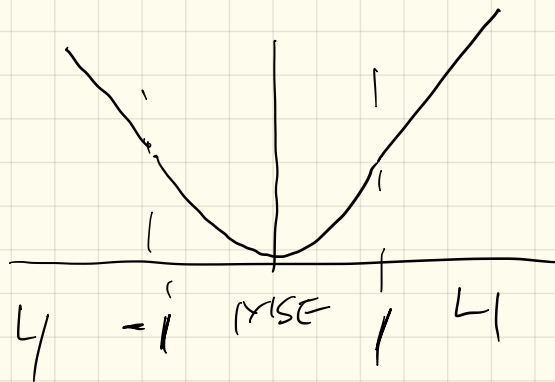
example

$1, 1, 1, \dots, 1, 1, 1, 100$

$$MSE = \frac{1}{10} (1 + 1 + \dots + 100^2)$$

$$L1 = \frac{1}{10} (1 + 1 + \dots + 100)$$

MSE + L1



$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{\frac{1}{2}}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & \dots & \dots & a_{mn} \end{bmatrix}$$

LDA $\{1, \dots, K\}$ class

$$\begin{aligned} \circ P(Y=k | x) &= \frac{P(k) P(X|Y=k)}{\sum_{i=1}^K P(i) P(X|Y=i)} \\ &\quad \text{class probability} \\ &= \frac{\pi_k \cdot e^{-\frac{1}{2}(x - \underline{\mu}_k)^T \Sigma_k^{-1} (x - \underline{\mu}_k)}}{\sum_{i=1}^K \pi_i e^{-\frac{1}{2}(x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i)}} \end{aligned}$$

Machine Learning

Teacher : Data

Student : model

||

parametric function.

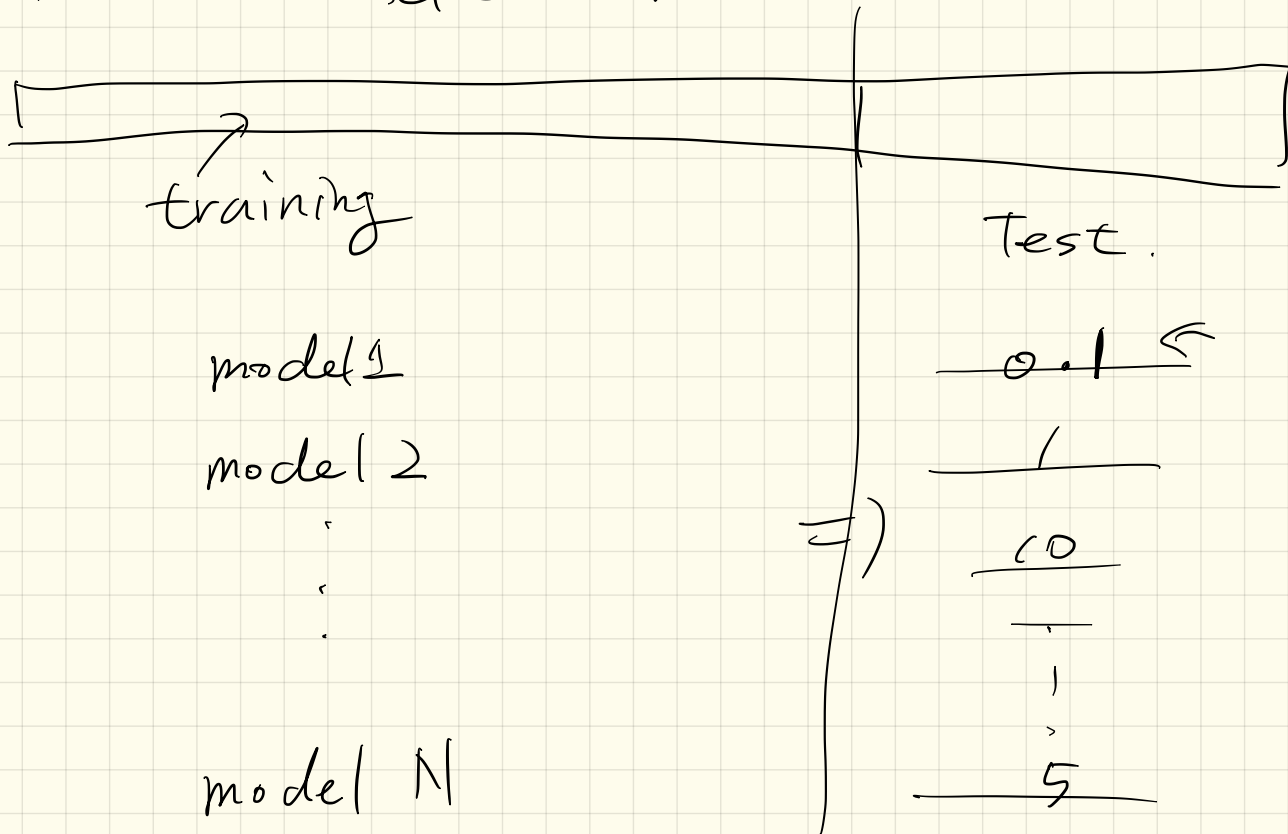
① define model \Leftarrow

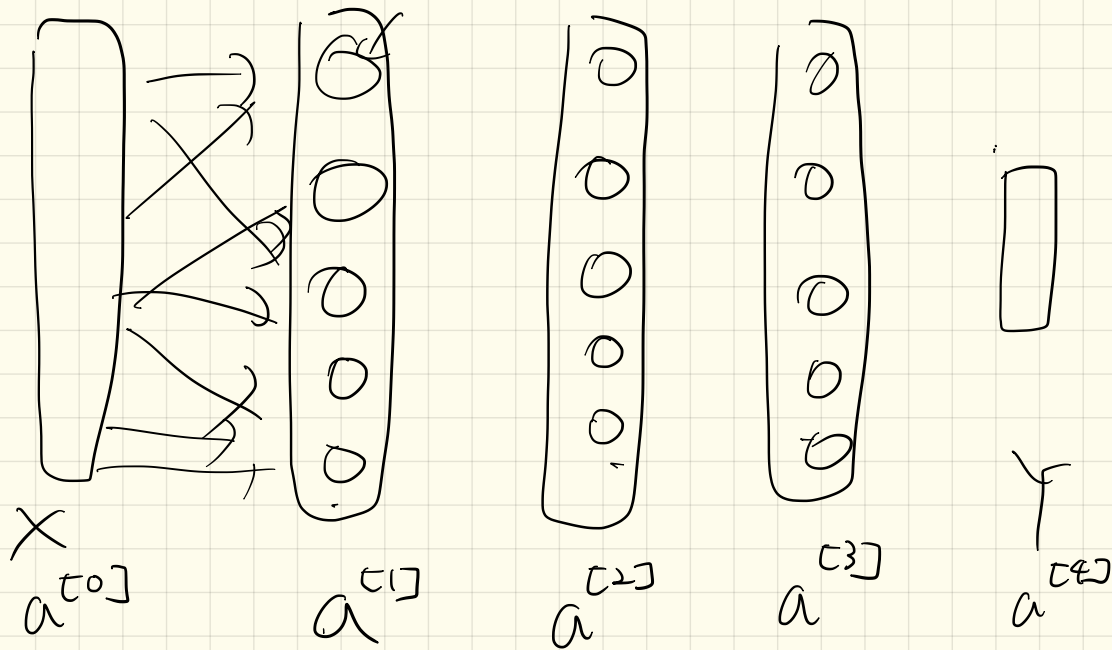
② define performance measure.

\Rightarrow MSE, ABS(L1), MSE + ABS

\Rightarrow Cross entropy, hinge loss

How to select model.





$$\underline{w}^{[1]} = \begin{bmatrix} w^{[1]}_{ij} \\ \text{input} \quad \text{output} \end{bmatrix}$$

$$\underline{b}^{[1]}$$

$$a^{[1]} = \sigma^{[1]}(w^{[1]} \cdot x + b^{[1]})$$

$$a^{[2]} = \sigma^{[2]}(w^{[2]} \cdot a^{[1]} + b^{[2]})$$

$$= \sigma^{[2]}(w^{[2]} \sigma^{[1]}(w^{[1]} \cdot x + b^{[1]}) + b^{[2]})$$

0
0
0



0

a^{π}



general

