

Para aprender

$$E_{in}(\theta) \approx 0$$

$$E_{in}(\theta) \approx E_{out}(\theta)$$

Error en muestra

Error fuera de muestra

Solo si ocurren los 2, hay aprendizaje

$E_{in} \rightarrow$ es una muestra representativa

Nunca podemos afirmar que NO se parecen

E_{in}

E_{out}

Muchos datos

$$P[|E_{out}(\theta) - E_{in}(\theta)| > \epsilon] \leq K e^{-\epsilon^2 n}$$

Muy pequeña

En reconocimiento

La probabilidad de que no se parezcan es muy pequeña

U conjunto universo

$$\phi: \mathcal{P}^U \rightarrow \mathbb{R}$$

$$P(A) > 0$$

$$P(\emptyset) = 0$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B) \geq P(A) + P(B)$$

En aprendizaje

$$P[\bigcup_{\theta} |E_{out}(\theta) - E_{in}(\theta)| > \epsilon] \leq \sum_{\theta} P[|E_{out}(\theta) - E_{in}(\theta)| > \epsilon] \leq O[K e^{-\epsilon^2 n}]$$

Puede ser infinito

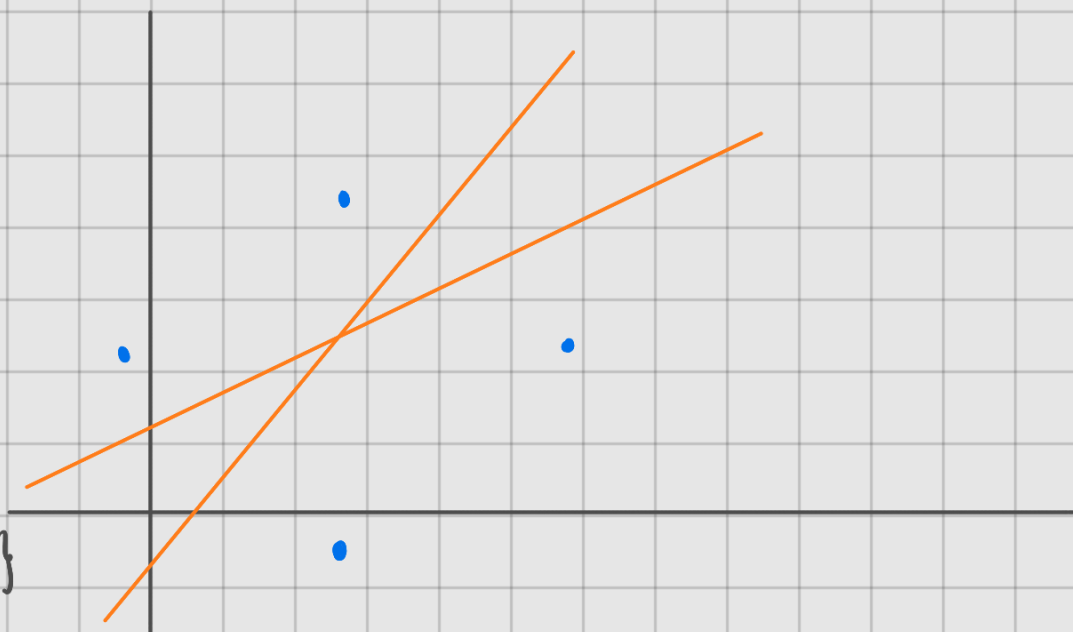
Cota superior

Ahora el problema es encontrar una cota superior más baja

$$h_{\theta}: X \rightarrow \{-1, 1\}$$

$\{p_1, p_2, p_3, p_4\}$

2^{Θ^2} Posibilidades



H conjunto de func.

$$h_{\theta}: X \rightarrow \{-1, 1\}$$

$$h_{\theta}: \{p_1, p_2, \dots, p_m\} \rightarrow \{-1, 1\}$$

cardinalidad $H = \infty$

$$\text{card } |H(p_1, \dots, p_m)| \leq 2^m$$

Función de crecimiento

$$m_H(M) = \max_{x^{(1)}, \dots, x^{(M)} \in X} (H(x^{(1)} \dots x^{(M)}))$$

↓
Máxima cardinalidad
posible

$$\text{Acotado a } m_H(M) \leq 2^M$$



Para este caso
la cardinalidad es
6

$$2^3 - 2$$

$$h(x) = \text{sign}(w_1 x_1 + w_2 x_2 + \frac{1}{k})$$

$$X = \mathbb{R}$$

$$\text{Card}(H) = \infty$$

$$m_H(1) = 2$$

$$m_H(2) = 2^2 = 4$$

$$m_H(3) = 2^3 = 8$$

$$m_H(4) = 14 < 2^4 \quad \text{A partir de 4 datos no puedo memorizar}$$

$$P[|E_{\text{out}}(\theta) - E_{\text{in}}(\theta)| \gg \epsilon] \leq \underbrace{2^{m_H(M)}}_{\substack{\downarrow \\ \text{Crece con} \\ \text{respecto a } M}} \cdot \underbrace{C(k e^{-\epsilon^2 M})}_{\substack{\downarrow \\ \text{Decrece con} \\ \text{respecto a } M}}$$

Vapnik -- Chernovenskis

$$\delta \approx M^{\text{duc}} e^{-M}$$

$$M \geq 10 \text{duc}(H)$$

Mi conjunto de datos debe ser al menos 10 veces mayor
a mi número de parámetros

↓
Crece más
rápido de lo
que decrece