03. Teorema NFL y dimensión VC

1 Problemas

Problema (5.1)**.** Utilizando el hint y el Lema B.1,

$$\begin{aligned}\mathbb{P}_{S\sim D^m}[L_D(A(S)) \geq 1/8] &= \mathbb{P}_{S\sim D^m}[L_D(A(S)) \geq 1 - 7/8] \\ &\geq \frac{\mathbb{E}[L_D(A(S)) - (1/8)]}{7/8} \geq \frac{1/8}{7/8} = \frac{1}{7}\end{aligned}$$

Problema (5.2)**.** Dada la complejidad de las clases, como $H_2 \subseteq H_5$, H_5 tiene más complejidad, un mayor error de estimación y un menor error de aproximación a comparación de H_2. Por lo tanto, si se cuenta con un conjunto reducido de entrenamiento, se prefiere aprender con la clase más pequeña.

Problema (6.5)**.** A generalizar la demostración del caso $d = 2$. Dados los números reales $a_i \leq b_i,\ i \in [d]$, se define el clasificador

$$h_{(a_1,b_1,\ldots,a_d,b_d)}(x_1,\ldots,x_d) = \prod_{i=1}^{d} 1_{[x_i \in [a_i,b_i]]} \tag{1}$$

La clase de rectángulos alineados con los ejes en $\mathbb{R}^d$

$$H^d_{rec} = \{h_{(a_1,b_1,\ldots,a_d,b_d)} : \forall i \in [d],\ a_i \leq b_i\} \tag{2}$$

Considere el conjunto $\{x_1,\ldots,x_{2d}\}$ con $x_i = \vec{e}_i,\ i \in [d]$, y $x_i = -\vec{e}_{i-d},\ i > d$. Elija

$$a_i = \begin{cases} -2 & y_{i+d} = 1 \\ 0 & \text{othw.} \end{cases}$$

Similarmente

$$b_i = \begin{cases} 2 & y_i = 1 \\ 0 & \text{othw.} \end{cases}$$

Entonces, $h(x_i) = y_i$, $\forall i \in [2d]$. Asimismo, $\mathrm{VCdim}(H^d_{rec}) \geq 2d$. Ahora, considere un conjunto C cuya cardinalidad supera a $2d$. A probar que C no puede ser roto H^d_{rec}. Por el principio de las casillas, existe un elemento $x \in C$ tal que para cada $j \in [d]$, existe $x' \in C$ con $x'_j \leq x_j$. También existe $x'' \in C$ con $x''_j \geq x_j$. Por lo tanto, la anotación en la que x es negativa y el resto de elementos en C son positivos no se puede obtener.

Problema (6.7)**.** *a)* La clase $H = \{1_{[x \geq t]:t\in\mathbb{R}}\}$ definida sobre subconjuntos no vacíos de $\mathbb{R}$ de umbrales sobre la línea es infinito, cuando su dimensión VC es igual a 1.

b) La clase de hipótesis $H = \{1_{[x\leq 1]}, 1_{[x\leq 1/2]}\}$ satisface los requerimientos.

Problema (6.9)**.** A probar que $\text{VCdim}(H) = 3$. Sea $C = \{1, 2, 3\}$. Sea $(\pm, \pm, \pm)$ los signos correspondientes a C.

1. $(-, -, -)$, $(a, b, s) = (0.5, 3.5, -1)$
2. $(-, -, +)$, $(a, b, s) = (2.5, 3.5, 1)$
3. $(-, +, -)$, $(a, b, s) = (1.5, 2.5, 1)$
4. $(-, +, +)$, $(a, b, s) = (1.5, 3.5, 1)$
5. $(+, -, -)$, $(a, b, s) = (0.5, 1.5, 1)$
6. $(+, -, +)$, $(a, b, s) = (1.5, 2.5, -1)$
7. $(+, +, -)$, $(a, b, s) = (0.5, 2.5, 1)$
8. $(+, +, +)$, $(a, b, s) = (0.5, 3.5, 1)$

Se concluye que $\text{VCdim}(H) \geq 3$. Sea $C = \{x_1, \dots, x_4\}$ y suponga que $x_1 < \cdots < x_4$. Entonces, La anotación $y_1 = y_3 = -1$ y $y_2 = y_4 = 1$ no se obtiene por ninguna hipótesis de H. Por lo tanto, $\text{VCdim}(H) = 3$.