Moshe Binieli - 311800668

1) $h_{(a_1,b_1,a_2,b_2)}(X_1,X_2) = \begin{cases} 1 & \text{if } a_1 \le x_1 \le b_1 \text{ and } a_2 \le x_2 \le b_2 \\ 0 & \text{else} \end{cases}$

$A = \{$ Smallest rectangle that covers all train set $\}$

A is ERM?

__Answer:__ By (עקרון מימושיות), there is some function $f^*(x) \in H$ such that

$\mathbb{E}_{x\sim D}[\ell(y, f^*(x))] = 0$

"A" algorithm creates __Best__ fitting rectangle on the dataset, it separate perfectly positive examples and negative examples, __therefore it is ERM__

2)

We need to find $E_{S\sim D}[L_S(h)] = L_D(h)$

$$\boxed{* \quad E[\mathbb{I}(X)] = 1 \cdot P(X) + (1 - P(X)) \cdot 0 = P(X) \; *}$$

$\mathbb{E}_{S\sim D}[L_S(h)] = \mathbb{E}_{S\sim D}\left[\frac{1}{m}\sum_{1}^{m} \mathbb{I}(h(x) \neq f(x))\right] =$ take constants out

$\frac{1}{m} \cdot \mathbb{E}_{S\sim D}\left[\sum_{1}^{m} \mathbb{I}(h(x) \neq f(x))\right] =$

$= \frac{1}{m} \cdot \sum_{1}^{m} \left[\mathbb{E}_{x\sim D}[\mathbb{I}(h(x) \neq f(x))]\right] = \frac{1}{m} \cdot \sum_{1}^{m} \cdot P_{x\sim D}(h(x) \neq f(x)) = \frac{1}{m} \cdot \sum_{1}^{m} \cdot L_D(h) =$

$= L_D(h)$

Scanned by CamScanner