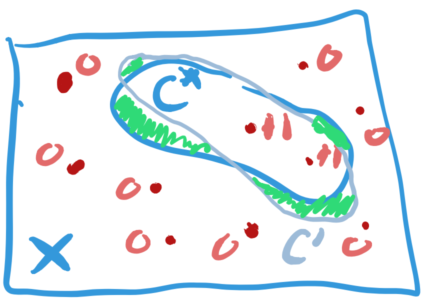
Chapter 5

concept learning (binary)

realizable $(X, \mathcal{C}, \mathcal{P})$ $\qquad x_i \in X$

Given samples $(x_1, y_1), \dots, (x_n, y_n)$ $\quad y_i \in \{0,1\}$



$c \subset X$

↑ a concept

$\mathcal{C}$ = a set of possible concepts
= a set of subsets of X

$L_P(c', c^*) = P(c' \triangle c^*)$

[Say c' is ε-accurate (for P, c^*) if

$$P(c' \triangle c^*) \leq \varepsilon.$$
]

Given $(X, \mathcal{C}, \mathcal{P})$

Definition A learning algorithm

$A: A_n((x_1, y_1), \dots, (x_n, y_n)) \to \hat{c}$ is Probably Almost Correct (PAC) if for any $\varepsilon, \delta > 0$, there exists $n(\varepsilon, \delta)$ so if $n \geq n(\varepsilon, \delta)$ then for any $c^* \subset \mathcal{C}$ and for any $P \in \mathcal{P}$ $\hat{c}$ is ε-accurate with probability at least $1-\delta$.

Theorem 5.1 If $|\mathcal{C}| < +\infty$ then there exists a PAC algorithm (i.e. the problem is PAC learnable).

Proof $\mathcal{C} = \{C_1, \dots, C_M\}$

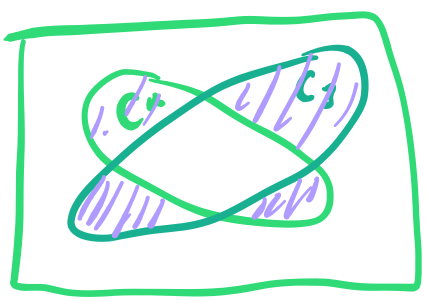
know $C^* \in \mathcal{C}$.

Fix $P \in \mathcal{P}$.

Fix $\varepsilon, \delta > 0$

$\varepsilon = 10\%$

$B = \{ C \in \mathcal{C} : P(C \triangle C^*) > \varepsilon \}$ - "bad" concepts



$P(\hat{C} = C_3) \leq (1-\varepsilon)^n$

$\hat{C} = A((x_1, y_1), \dots, (x_n, y_n))$

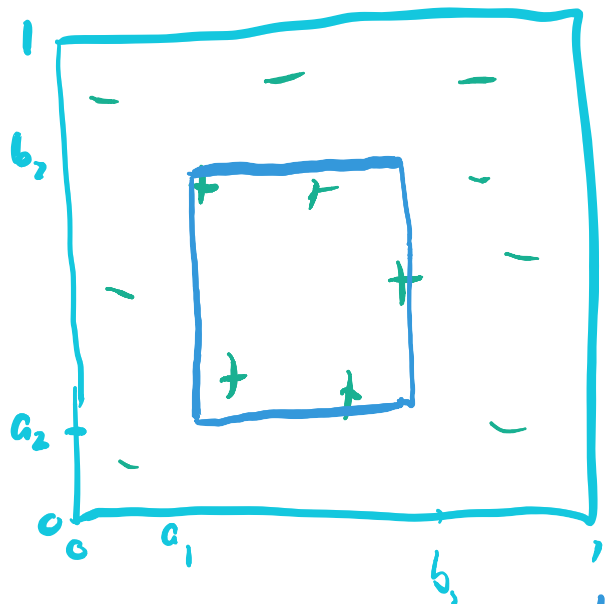
So if $M(1-\varepsilon)^n \leq \delta$ then

$P(\hat{C}$ is ε-accurate$) \geq 1-\delta$.

Example - Axis parallel rectangles
- Realizable case

$X = [0,1]^2$ $\quad \mathcal{C} = \{\text{axis parallel rectangles}\}$

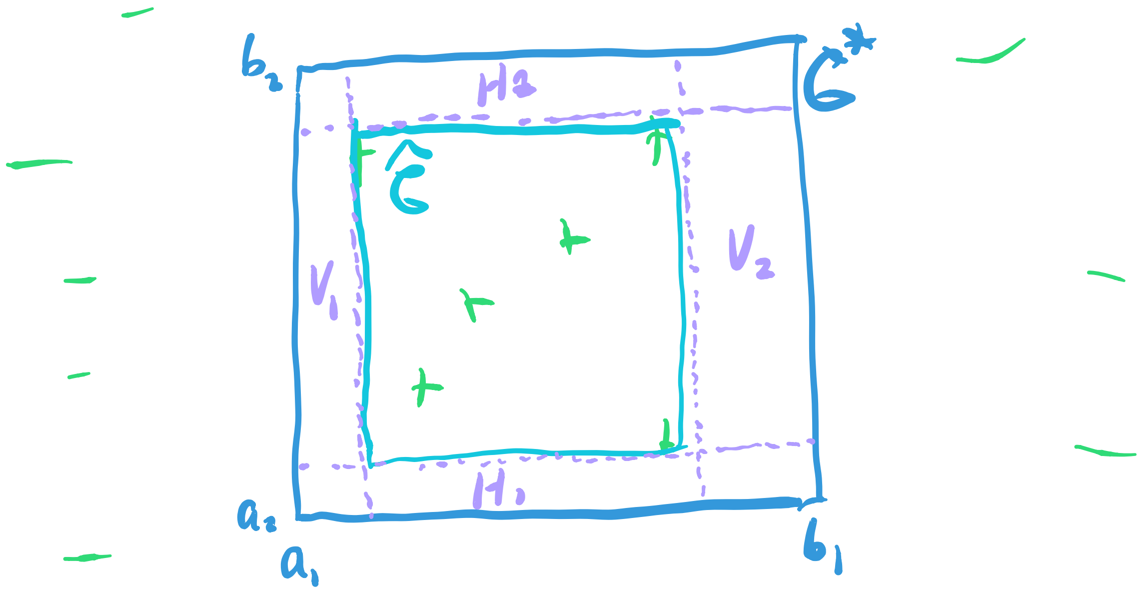
$\mathcal{P}$ = all prob. dist^{ns} on X



Let $\hat{C}$ be the smallest rectangle giving zero errors on test data

Claim This is a PAC algorithm.

New picture. Draw C^* = true rectangle

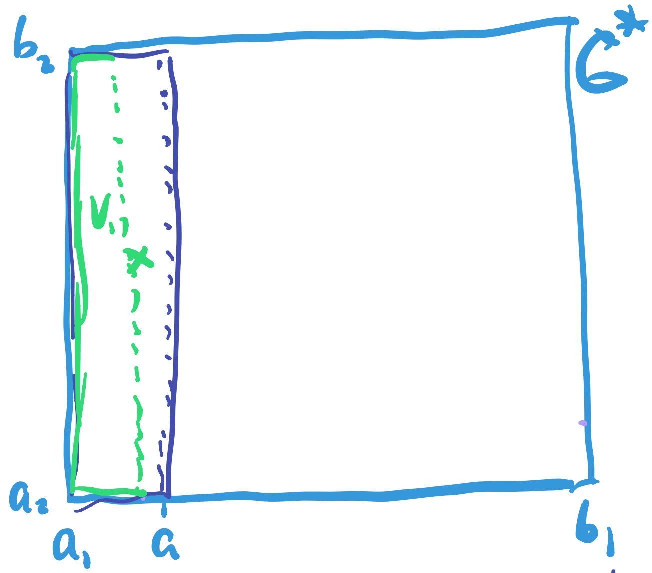


Done if $P(V_1 \cup V_2 \cup H_1 \cup H_2) \le \varepsilon$

with probability at least $1-\delta$.

Newer picture

P, C^* fixed.



Select a so $[a_1, a] \times [a_2, b_2]$ is smallest rectangle of this form with P probability $\geq \varepsilon/4$

$$P([a_1, a] \times [a_2, b_2]) \geq \frac{\varepsilon}{4} \geq P([a_1, a) \times [a_2, b_2])$$

If $4\left(1 - \frac{\varepsilon}{4}\right)^n \leq \delta$

then $P(V_1 \cup V_2 \cup H_1 \cup H_2) \leq \varepsilon$ with prob. at least $1-\delta$.

$\square$

#1

$X = [0,1]$

$\mathcal{C} = \{ [a, 1] : 0 \le a \le 1 \}$

$\mathcal{P}$ - all dist$^{\text{ns}}$ on $[0,1]$