

Question 4 – PAC, VC dimension

Section 1

Look at the following classifiers family:

$$\mathcal{H} = \{h_{a,b}: a, b \in \mathbb{R}\} \text{ where } h_{a,b}(x, y) = 1 \text{ iff } x \geq a \wedge y \geq b$$

Find the VCdim of this class with full proof.

Section 2

Consider instances X containing 4 boolean variables $\{X_1, X_2, X_3, X_4\}$ and responses Y are $(X_1 \wedge X_4) \vee (\neg X_1 \wedge X_3)$. We try to learn the function $f: X \rightarrow Y$ using a "depth 2 decision trees". A "depth-2 decision tree" is a tree with four leaves, all distance 2 from the root.

Analyze the problem and give the lower bound on the sample complexity that matches this case.

Section 3 – optional, 5 pts bonus.

In specific classification problem, let $\mathcal{X} \in \{0,1\}^d$ be our domain. Let \mathcal{H} be the classifiers family, such that:

$$\mathcal{H} = \left(\prod_{i \in R_1} x_i \right) \vee \left(\prod_{i \in R_2} x_i \right)$$

Where R_1, R_2 are subgroups of $\{1, \dots, d\}$.

Show that the upper bound of $VC(\mathcal{H})$ is $2d$.

Hint: use combinatorics and VC upper bound we learned.