MALTA 2020–2021: Assignment

Provide your solution before the start of next lecture on October 20th. If you are unable to attend next lecture, then send your solution by email to giovanni.neglia@inria.fr.

Motivate your answers.

Ex. 1 — Consider the domain $X = \{x_1, x_2\}$ and the following distribution D over $X \times \{0, 1\}$;

$$D((x_1, 1)) = \frac{2}{3}p, \qquad D((x_1, 0)) = \frac{2}{3}(1-p),$$
$$D((x_2, 1)) = \frac{1}{3}q, \qquad D((x_2, 0)) = \frac{1}{3}(1-q),$$

where $p \in (0.5, 1]$ and $q \in [0, 0.5)$.
Consider the usual 0–1 loss:

$$l(h, (x, y)) = \begin{cases} 1, & \text{if } h(x) \neq y, \\ 0, & \text{otherwise.} \end{cases}$$

We want to minimize the expected loss $L_D(h) = \mathbb{E}_D[l(h, (x, y))]$.

1. What is the optimal deterministic predictor? Can you prove its optimality?
2. Prove that no randomized predictor can achieve a smaller loss.

Ex. 2 — In the proof of the "no free lunch theorem," we only proved that

$$\exists D, f \text{ such that } \mathbb{E}\left[L_{D,f}(A(S))\right] \geq \frac{1}{4}.$$

Show that this is sufficient to conclude that

$$\text{Prob}\left(L_{D,f}(A(S)) \geq \frac{1}{8}\right) \geq \frac{1}{7}.$$

(Probability and expectation are computed over the possible sampling of the dataset S.)

Ex. 3 — Prove that if the class H can shatter A, then it can shatter any set $B \subset A$.
If you show that H cannot shatter any set of size n, do you need to check if it can shatter a set of size $n' > n$? Why?

Ex. 4 — Consider the class of functions

$$H = \{h_{a,b,c,d} : \mathbb{R} \to \{0,1\},\ \text{for some } a,b,c,d \in \mathbb{R},\ \text{with } a < b < c < d\},$$

where

$$h_{a,b,c,d}(x) = \begin{cases} 0, & \text{if } x < a, \\ 1, & \text{if } x \in [a,b), \\ 0, & \text{if } x \in [b,c), \\ 1, & \text{if } x \in [c,d). \\ 0, & \text{if } x \geq d. \end{cases}$$

What is the VC-dimension of H?

Ex. 5 — Consider the class of concentric disks

$$H_0 = \{h_r : \mathbb{R}^2 \to \{0,1\},\ \text{for some } r \in \mathbb{R}^+\},$$

where

$$h_r(\mathbf{x}) = \begin{cases} 1, & \text{if } \|\mathbf{x}\| \leq r, \\ 0, & \text{otherwise.} \end{cases}$$

What is the VC-dimension of H?

Ex. 6 — Consider the class HS^n of the linear classifiers over $\mathbb{R}^n$. Prove that VC-dim$(HS^n) \geq n+1$.
(For example you can try to shatter the set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{0}\}$, where $\mathbf{e}_i$ is the canonical vector with $e_{i,j} = 1$ if $j = i$ and $e_{i,j} = 0$ otherwise, and $\mathbf{0}$ is the vector with all components equal to 0.)