MALTA 2021–2022: Assignment

Provide your solution before the start of next lecture on October 12th. If you are unable to attend next lecture, then send your solution by email to giovanni.neglia@inria.fr.

Motivate your answers.

Ex. 1 — Show that a finite class has finite VC-dimension.

Ex. 2 — Prove that in the definition of agnostic PAC learnability, when the loss is the 0-1 loss function and the learning algorithm is required to output a hypothesis in H ($A(S) \in H$), it is possible to replace the probabilistic condition

$$\begin{aligned} &\forall \epsilon, \delta \in (0,1),\ \exists\, m_H(\epsilon,\delta) \in \mathbb{N},\ \text{such that } \forall m \geq m_H(\epsilon,\delta), \\ &\qquad L_D(A(S)) - \min_{h\in H} L_D(h) \leq \epsilon,\ \text{with prob. larger than } 1-\delta. \end{aligned}$$

with the following condition on the expected value

$$\lim_{m\to\infty} \mathop{\mathbb{E}}_{S\in D^m}\left[L_D(A(S)) - \min_{h\in H} L_D(h)\right] = 0$$

Hints: We have already mapped during the lecture a probabilistic condition to a condition on expected values. The Markov inequality may help you.

Ex. 3 — 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 H_k of binary functions over $\mathbb{R}$ which assume value 1 exactly on k points, i.e.,

$$H_k = \Big\{h : \mathbb{R} \to \{0,1\},\ \text{such that } \exists k \text{ distinct values } x_1, x_2, \dots, x_k \in \mathbb{R},$$
$$\text{such that } h(x) = 1 \text{ if } x \in \{x_1, x_2, \dots, x_k\} \text{ and } h(x) = 0 \text{ otherwise}\Big\}.$$

What is the VC-dimension of H?

Ex. 5 — Consider the class of squares in the plane

$$H = \{h_{a,b,r} : \mathbb{R}^2 \to \{0,1\},\ \text{for some } a, b \in \mathbb{R},\ \text{and } r \in \mathbb{R}^+\},$$

where

$$h_{a,b,r}(x_1, x_2) = \begin{cases} 1, & \text{if } x_1 \in [a, a+r] \text{ and } x_2 \in [b, b+r], \\ 0, & \text{otherwise.} \end{cases}$$

1. What is the VC-dimension of H?
   Hint: The following geometric result may help: given any set A of 4 points in the plane, it is always possible to split A in two disjoint sets B and C ($B \cup C = A$, $B \cap C = \emptyset$) such that $\mathcal{CH}(B) \cap \mathcal{CH}(C) \neq \emptyset$, where $\mathcal{CH}(S)$ denotes the convex hull of the set S.
2. What if r is fixed to 1?

Ex. 6 — Consider a binary classification problem over $\mathbb{R}$ and the 0-1 loss function. Let $\mathcal{P}$ denote the class of all polynomials over $\mathbb{R}$. We consider the following hypothesis class:

$$H = \{h(x) = g(p(x)),\ \text{for some } p \in \mathcal{P}\},$$

where $g : \mathbb{R} \to \{0,1\}$ and $g(x) = 1$ if and only if $x > 0$.
Is the class H PAC learnable?