Assignment 3

Due: 13.06.2024, 23:59

Points: 15

The solutions have to be handed in via Moodle. We do not accept late submissions.

We would recommend using LaTeX for writing your submission but also accept handwritten solutions, but please note that if we can not read or understand it, we cannot grade it.

To get full points, always provide the steps in your derivation/proofs and make clear when/how you use known results, for example, from the lecture (e.g. already proven concentration inequalities).

Exercise 3.1: PAC expected risk of the learner

Let $\mathcal{H}$ be a hypothesis, and $\mathcal{A}$ be learner for $\mathcal{H}$. Assume that the loss function is non-negative and contained in $[0,1]$. Define the *expected risk of the learner* $\mathcal{A}$ as $\mathbb{E}_{S\sim\mathcal{D}^m}\left[L_{\mathcal{D}}\big(\mathcal{A}(S)\big)\right]$.

(a) Assume there is a function $\tilde{m}_{\mathcal{H}}(\epsilon)$ such that for every $\epsilon \in (0,1)$ and distribution $\mathcal{D}$,

$$\mathbb{E}_{S\sim\mathcal{D}^m}\left[L_{\mathcal{D}}\big(\mathcal{A}(S)\big)\right] \leq \inf_{h\in\mathcal{H}} L_{\mathcal{D}}(h) + \epsilon \qquad \text{for } m \geq \tilde{m}_{\mathcal{H}}(\epsilon).$$

Show that $\mathcal{H}$ is agnostic PAC learnable.

(b) Now assume the opposite: For every m, there is a distribution $\mathcal{D}$ such that

$$\mathbb{E}_{S\sim\mathcal{D}^m}\left[L_{\mathcal{D}}\big(\mathcal{A}(S)\big)\right] > \inf_{h\in\mathcal{H}} L_{\mathcal{D}}(h) + \epsilon.$$

Show that, in this case, $\mathcal{H}$ is not agnostic PAC learnable.

(3+3 = 6 points)

Exercise 3.2: Complement of a learnable class

Let $\mathcal{X} = \mathbb{R}$. Let $\mathcal{H} \subseteq \{\pm 1\}^{\mathcal{X}}$ be a PAC learnable hypothesis class, and $\overline{\mathcal{H}} = \{\pm 1\}^{\mathcal{X}} \backslash \mathcal{H}$ be the complement of $\mathcal{H}$. Prove or disprove the following statement: $\overline{\mathcal{H}}$ is PAC learnable.

(3 points)

Exercise 3.3: VC dimensions

1. Find the VC dimension of

$$\mathcal{H} = \left\{ \sum_{i=0}^{k} \mathbf{1}\left\{t_{2i} \leq x < t_{2i+1}\right\}, \quad 0 \leq t_0 < \cdots < t_{2k+1} \leq 1 \right\}$$

   for any $k \geq 1$.

2. $\mathcal{X} = \mathbb{R}$. Consider the following set of functions

$$\mathcal{H} = \{\operatorname{sign}(\sin(ax)), \quad a \in \mathbb{R}\}$$

   with $\operatorname{sign}(x) = +1$ if $x \geq 0$ and $\operatorname{sign}(x) = -1$ otherwise. Find $\operatorname{VCdim}(\mathcal{H})$.

   Hint: For your proof it is helpful to consider the set of points $x_i = 10^{-i}$. Then, for any labels $y_1, \ldots, y_n$ choose

$$a = \pi\left(1 + \sum_{i=1}^{n} \frac{(1-y_i)10^i}{2}\right)$$

(2+4 = 6 points)