Telecom Paris
A. Tchamkerten APM0EL05TP

ASSIGNMENT 3

Exercise 1 (Error decomposition)**.** Let h_S be an ERM$_{\mathcal{H}}$ predictor for some function class $\mathcal{H}$. Write the prediction error $L_P(h_s) = \mathbb{E}_{Z\sim P}(\ell(Z, h_S))$ as

$$L_P(h_s) = \varepsilon_{\text{app}} + \varepsilon_{\text{est}}$$

where $\varepsilon_{\text{app}} := \min_{h\in\mathcal{H}} L_P(h)$ and $\varepsilon_{\text{est}} := L_P(h_s) - \varepsilon_{\text{app}}$. Interpret this error decomposition.

Exercise 2 (VC dimension, parity)**.** Let $\mathcal{X} = \{0,1\}^n$. Given $\mathcal{I} \in \{1, 2, \ldots, n\}$ let

$$h_{\mathcal{I}}(x) = (\sum_{i\in\mathcal{I}} x_i) \mod 2$$

denote the parity of x over the coordinates in $\mathcal{I}$. Show that the VC dimension of the set of all such functions, that is

$$\mathcal{H}_{parity} = \{h_{\mathcal{I}} : \mathcal{I} \subset \{1, 2, \ldots, n\}\},$$

is n.

Exercise 3 (VC dimension, signed intervals)**.** Consider the class of signed intervals over $\mathcal{X} = \mathbb{R}$

$$\mathcal{H} = \{h_{a,b,s} : a \leq b, s \in \{-1, 1\}\}$$

where $h_{a,b,s}(x) = s$ if $x \in [a, b]$ and $h_{a,b,s}(x) = -s$ if $x \notin [a, b]$. Show that VCdim($\mathcal{H}$)=3.

Exercise 4 (VC dimension, halfspaces)**.** A homogeneous halfspace is specified by a vector $\mathbf{w}$ in $\mathbb{R}^d$ which defines a binary function

$$\mathbf{x} \mapsto h_{\mathbf{w}}(\mathbf{x}) := \text{sign}\langle \mathbf{w}, \mathbf{x}\rangle$$

Show that the VCdimension of the class of homogeneous halfspaces in $\mathbb{R}^d$ is equal to d. Show that the VCdimension of the class of non-homogeneous halfspaces defined by

$$\mathbf{x} \mapsto h_{\mathbf{w},b}(\mathbf{x}) := \text{sign}\langle \mathbf{w}, \mathbf{x}\rangle + b$$

with $\mathbf{w}$ in $\mathbb{R}^d$ and b in $\mathbb{R}$ is $d + 1$.

Exercise 5 (VC dimension, bounds)**.** In class we established the upper bound $VCdim(\mathcal{H}) \leq \log(|\mathcal{H}|)$. Here we will show that this bound can be quite loose.

1. Find an example of a class $\mathcal{H}$ of functions on the unit interval $[0, 1]$ such that $VCdim(\mathcal{H}) < \infty$ while $|\mathcal{H}| = \infty$.

2. Find an example of a finte class $\mathcal{H}$ of functions on the unit interval $[0, 1]$ where $VCdim(\mathcal{H}) < \log(|\mathcal{H}|)$.