Homework 7

ESE 4020/5420

Ungraded

Problem 1. Let $\mathcal{H}$ be the class of functions of the form $h_{a,b} : \mathbb{R} \mapsto \{0, 1\}$, where

$$h_{a,b}(x) = \mathbb{1}_{\{x \in [a,b]\}}$$

for all $a, b \in \mathbb{R}$. Our goal will be to show that $\text{VCdim}(\mathcal{H}) = 2$.

(a) Let $C = \{c_1, \ldots, c_k\} \subset \mathbb{R}$. Recall that the restriction of $\mathcal{H}$ to C, denoted by $\mathcal{H}_C$, is the set of all binary k-tuples that can be derived from evaluating the functions in $\mathcal{H}$ on the set C. That is,

$$\mathcal{H}_C = \{(h(c_1), \ldots, h(c_k)) : h \in \mathcal{H}\}$$

Compute $\mathcal{H}_C$ for $C = \{1, 2\}$, and $C = \{1, 2, 3\}$.

(b) For any set C with $|C| = 2$, show that $|\mathcal{H}_C| = 4$. Also, can we say that for any set C with $|C| = 3$ we have $|\mathcal{H}_C| < 8$?

(c) Recall that a function class $\mathcal{H}$ shatters a set C if $|\mathcal{H}_C| = 2^{|C|}$. Given the specific choice of $\mathcal{H}$ as above, does $\mathcal{H}$ shatter any set C of size 2? Is there a set C of size 3 that is shattered by $\mathcal{H}$?

(d) Recall that the VC dimension of $\mathcal{H}$ is the maximal size of a set C that can be shattered by $\mathcal{H}$. What is the VC dimension of the class $\mathcal{H}$ considered in the question?

Problem 2. Consider two hypothesis classes $\mathcal{H}_1$ and $\mathcal{H}_2$. Show that if $\mathcal{H}_1 \subseteq \mathcal{H}_2$, we have $\text{VCdim}(\mathcal{H}_1) \leq \text{VCdim}(\mathcal{H}_2)$

Problem 4. Let $\mathcal{H}$ be the class of signed intervals, that is, $\mathcal{H} = \{h_{a,b,s} : a \leq b, s \in \{0, 1\}\}$ where

$$h_{a,b,0}(x) = \begin{cases} 0 & \text{if } x \in [a, b] \\ 1 & \text{if } x \notin [a, b] \end{cases} \quad \text{and} \quad h_{a,b,1}(x) = \begin{cases} 1 & \text{if } x \in [a, b] \\ 0 & \text{if } x \notin [a, b] \end{cases}$$

Calculate $\text{VCdim}(\mathcal{H})$.

Problem 5. *Short answer:*

(a) In your own words, define what it means for a hypothesis class to be *PAC learnable.*

(b) Let $\mathcal{H}$ be the class of *all* functions from $\mathbb{R}$ to $\{0,1\}$. What is the VC dimension of $\mathcal{H}$?

(c) For a finite function class $\mathcal{H}$ with m functions $h_1, \ldots, h_m$, such that $h_i : \mathcal{X} \to \{0,1\}$, explain why $\text{VCdim}(\mathcal{H}) \leq \log_2 m$. Are there cases where the VC dimension is exactly $\log_2 m$?

(d) For a given hypothesis class, how can we prove that $VCdimension \geq k$? How can we prove that $VCdimension = k$?

Problem 6. Consider the following hypothesis class

$$\mathcal{H} = \{h_{w,b} \mid w, b \in \mathbb{R}\},$$

in which

$$h_{w,b}(x) = \begin{cases} 1 & wx + b \geq 0 \\ 0 & wx + b < 0 \end{cases}$$

Show that the VC dimension of $\mathcal{H}$ is 2.