

Seminar class 3

Exercise 1 (exercise 6.7 in the book)

We have shown that for a finite hypothesis class \mathcal{H} , $\text{VCdim}(\mathcal{H}) \leq \lfloor \log(|\mathcal{H}|) \rfloor$. However, this is just an upper bound. The VC-dimension of a class can be much lower than that:

1. Find an example of a class \mathcal{H} of functions over the real interval $\mathcal{X} = [0, 1]$ such that \mathcal{H} is infinite while $\text{VCdim}(\mathcal{H}) = 1$.
2. Give an example of a finite hypothesis class \mathcal{H} over the domain $\mathcal{X} = [0, 1]$, where $\text{VCdim}(\mathcal{H}) = \lfloor \log_2(|\mathcal{H}|) \rfloor$.

Exercise 2 (exercise 6.5 in the book)

VC-dimension of axis aligned rectangles in \mathbb{R}^d : Let $\mathcal{H}_{\text{rec}}^d$ be the class of axis aligned rectangles in \mathbb{R}^d . We have already seen that $\text{VCdim}(\mathcal{H}_{\text{rec}}^2) = 4$. Prove that in general, $\text{VCdim}(\mathcal{H}_{\text{rec}}^d) = 2d$.

Exercise 3 (exercise 6.6 in the book)

VC-dimension of Boolean conjunctions: Let $\mathcal{H}_{\text{con}}^d$ be the class of Boolean conjunctions over the variables x_1, \dots, x_d ($d \geq 2$). We already know that this class is finite and thus (agnostic) PAC learnable. In this question we calculate $\text{VCdim}(\mathcal{H}_{\text{con}}^d)$.

1. Show that $|\mathcal{H}_{\text{con}}^d| \leq 3^d + 1$.
2. Conclude that $\text{VCdim}(\mathcal{H}) \leq d \log 3$.
3. Show that $\mathcal{H}_{\text{con}}^d$ shatters the set of unit vectors $\{e_i : i \leq d\}$.
4. (***) Show that $\text{VCdim}(\mathcal{H}_{\text{con}}^d) \leq d$.
Hint: Assume by contradiction that there exists a set $C = \{c_1, \dots, c_{d+1}\}$ that is shattered by $\mathcal{H}_{\text{con}}^d$. Let h_1, \dots, h_{d+1} be hypotheses in $\mathcal{H}_{\text{con}}^d$ that satisfy

$$\forall i, j \in [d+1], h_i(c_j) = \begin{cases} 0 & i = j \\ 1 & \text{otherwise} \end{cases}$$

For each $i \in [d+1]$, h_i (or more accurately, the conjunction that corresponds to h_i) contains some literal ℓ_i which is false on c_i and true on c_j for each $j \neq i$. Use the Pigeonhole principle to show that there must be a pair $i < j \leq d+1$ such that ℓ_i and ℓ_j use the same x_k and use that fact to derive a contradiction to the requirements from the conjunctions h_i, h_j .

5. Consider the class $\mathcal{H}_{\text{mcon}}^d$ of monotone Boolean conjunctions over $\{0, 1\}^d$. Monotonicity here means that the conjunctions do not contain negations. As in $\mathcal{H}_{\text{con}}^d$, the empty conjunction is interpreted as the all-positive hypothesis. We augment $\mathcal{H}_{\text{mcon}}^d$ with the all-negative hypothesis h^- . Show that $\text{VCdim}(\mathcal{H}_{\text{mcon}}^d) = d$.

Exercise 4 (exercise 6.9 in the book)

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

$$h_{a,b,s}(x) = \begin{cases} s & \text{if } x \in [a, b] \\ -s & \text{if } x \notin [a, b] \end{cases}$$

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