PARTIAL SOLUTIONS FOR EXERCISE 2 SHEET 3

Exercise 2: Let $\mathcal{X}$ be a finite domain set and let $k \leq |\mathcal{X}|$. Find the VC dimension of

$$\mathcal{H} := \{h : \mathcal{X} \to \{0,1\} \mid |\{x : h(x) = 1\}| = k\},$$

i.e. the hypothesis class consisting of functions that take the value 1 on exactly k elements.

Claim. $\mathrm{VC}(\mathcal{H}) = \min\{k, |X| - k\}$.

Proof. Let $d = \min\{k, |X| - k\}$.

We argue that $\mathrm{VC}(\mathcal{H}) \leq d$. Suppose that $C = \{x_1, \ldots, x_n\} \subseteq \mathcal{X}$ and $n > d$. There are two cases, one of which is $n > k$. Consider the function f such that $f(x_i) = 1$ for all $1 \leq i \leq n$. Then we have that $f \notin \mathcal{H}_C$ because it is not possible for f to be a restriction of a function $g \in \mathcal{H}$ since g can only have a value of 1 at k-many points. The other case is where $n > |\mathcal{X}| - k$. Then consider the function f' such that $f(x_i) = 0$ for all $1 \leq i \leq n$. Again, $f' \notin \mathcal{H}_C$. This is because any function $g \in \mathcal{H}$ must take the value 0 at exactly $(|\mathcal{X}| - k)$-many points.

Now we argue that $\mathrm{VC}(\mathcal{H}) \geq d$. Suppose $\{x_1, \ldots, x_d\}$. (Without loss of generality, since we can always shatter proper subsets of sets that we can shatter.) We can show that any such set can be shattered. Let $f \in \mathcal{H}_C$. Let $A \subseteq [d]$ be the set of i such that $f(x_i) = 1$ and let $B \subseteq [d]$ be the set of 0 such that $f(x_i) = 0$. Let $n = k - |A|$ (possibly zero). Since we have

$$n \leq d \leq |\mathcal{X}| - k \leq |\mathcal{X}| - d,$$

there are $y_1, \ldots, y_n \in \mathcal{X} \setminus C$. Let g be the binary function with domain $\mathcal{X}$ such that $g(x_i) = 1$ for all $i \in A$ and $g(y_i) = 1$ for $i \in [n]$ and $g(z) = 0$ for all other elements of $\mathcal{X}$. Then in particular we have that g takes the value 1 on exactly k points, so $g \in \mathcal{H}$. We also have that $g \restriction C = f$, so $f \in \mathcal{H}_C$. $\square$

Exercise 3,1: What is the VC dimension of the hypothesis class $\mathcal{H}$ of axis-aligned squares?

We argue that $\mathrm{VCdim}(\mathcal{H}) = 3$.

Proof. It is fairly straightforward to show that there is a set of three points that can be shattered—an equilateral triangle with its base parallel to the x-axis will do—so we will focus on the argument that four or more points cannot be shattered.

Let $C = \{r_1 = (x_1, y_1), r_2 = (x_2, y_2), r_3 = (x_3, y_3), r_4 = (x_4, y_4)\}$ be a set of points in $\mathbb{R}^2$. Let M_V be the maximal difference of the y-values and let M_H be the maximal difference of the x-values. Without loss of generality, assume that $M_V \leq M_H$ and that $M_H = |x_2 - x_1|$. This is equivalent to saying that x_1 is the minimal x value and x_2 is the maximal x value or vice versa. Again, without loss of generality, assume that $y_1 \leq y_2$ and $y_3 \leq y_4$. (There are always two points with a maximal x- or y-distance apart, so we can assign those points the first two spots in the enumeration, then compare their y values, and then we can compare the y-values of the other two points.)

Now we have three cases to consider.

Case 1. $y_3 < y_1$ and $y_2 < y_4$. Then consider the assignment f such that $f(r_1) = f(r_2) = 1$ and $f(r_3) = f(r_4) = 0$. If a square contains r_1 and r_2, then it must have height at least M_H. This means that if r_3 is excluded from such a square, then r_4 must necessarily be included since $|y_4 - y_3| \leq M_H$ and we are requiring that the points r_1 and r_2 be included. Therefore the assignment cannot be realized and $f \notin \mathcal{H}_C$.

Case 2. $y_1 \leq y_3$. Then consider an assignment f such that $f(r_1) = f(r_2) = f(r_4) = 1$ and $f(r_3) = 0$. Then since $x_1 \leq x_3 \leq x_4$ (by our choice of x_1 and x_2 originally) and $y_1 \leq y_3 \leq y_4$, it follows that any square containing r_1, r_2, and r_4, must necessarily contain r_3 as well. Hence $f \notin \mathcal{H}_C$.

Case 3. $y_4 \leq y_2$. This is analogous to the previous case, but with the roles of r_3 and r_4 reversed.

□