Assignment 3:

Komal Wavhal
Student Id -20034443

Exercise 1

1. Show the following monotonicity property of VC-dimension: For every two hypothesis classes if $\mathcal{H}' \subseteq \mathcal{H}$ then $\mathrm{VCdim}(\mathcal{H}') \leq \mathrm{VCdim}(\mathcal{H})$.

Solution

To prove the monotonicity property of VC-dimension, we need to show that if $\mathcal{H}' \subseteq \mathcal{H}$, then $\mathrm{VCdim}(\mathcal{H}') \leq \mathrm{VCdim}(\mathcal{H})$.

Proof:

1. **Definition of VC-Dimension**: The VC-dimension of a hypothesis class $\mathcal{H}$ is the size of the largest set of points that can be shattered by $\mathcal{H}$. A set of points is said to be *shattered* by $\mathcal{H}$ if, for every possible binary labeling of the set, there exists a hypothesis in $\mathcal{H}$ that realizes that labeling.
2. **Implication of $\mathcal{H}' \subseteq \mathcal{H}$**: Since every hypothesis in $\mathcal{H}'$ is also in $\mathcal{H}$, the ability of $\mathcal{H}'$ to shatter any given set of points cannot be greater than that of $\mathcal{H}$. This means that if $\mathcal{H}$ can shatter a set of size d, then $\mathcal{H}'$ may or may not be able to do so, but it cannot shatter a strictly larger set than $\mathcal{H}$.
3. **Bounding the VC-Dimension**: Since $\mathcal{H}'$ has fewer hypotheses (or possibly the same number) compared to $\mathcal{H}$, the largest shattered set by $\mathcal{H}'$ must be of size at most the largest shattered set by $\mathcal{H}$. Therefore,

$$\mathrm{VCdim}(\mathcal{H}') \leq \mathrm{VCdim}(\mathcal{H}).$$

Conclusion:

This proves the monotonicity property of VC-dimension: A larger hypothesis class has a VC-dimension at least as large as any of its subsets.

Exercise 2

2. Given some finite domain set, $\mathcal{X}$, and a number $k \leq |\mathcal{X}|$, figure out the VC-dimension of each of the following classes (and prove your claims):
   1. $\mathcal{H}_{=k}^{\mathcal{X}} = \{h \in \{0,1\}^{\mathcal{X}} : |\{x : h(x) = 1\}| = k\}$. That is, the set of all functions that assign the value 1 to exactly k elements of $\mathcal{X}$.

[1] http://gilkalai.wordpress.com/2008/09/28/extremal-combinatorics-iii-some-basic-theorems

Solution

If $|X| = n$ and we consider the class of all 0–1-valued functions on X whose support has exactly k points, then its VC dimension is

$$\min\{k,\ n-k\}.$$

Let us denote our domain by X with $|X| = n$. The hypothesis class in question is

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

Equivalently, each $h \in \mathcal{H}_{=k}$ is the indicator function of some k-element subset of X.

1. Showing $\mathrm{VCdim}(\mathcal{H}_{=k}) \ \geq \ \min\{k,\ n-k\}$

We must exhibit a subset $S \subseteq X$ of size $d = \min\{k,\ n-k\}$ that is **shattered** by $\mathcal{H}_{=k}$. To say that S is shattered means: for **every** way of labeling the points of S by 0's and 1's, there is a hypothesis $h \in \mathcal{H}_{=k}$ (that is, a k-element subset of X) whose indicator function matches that labeling on S.

1. **Choose $S \subseteq X$ of size $d = \min\{k,\ n-k\}$.**
2. Let $T \subseteq S$ be the set of points in S that we want to label as 1 (and thus $|T| = t$ for some $t \leq d$).
3. We want to find a subset $H \subseteq X$ of size k whose indicator matches the labeling on S; that is,
   - $H \cap S = T$ on the points of S labeled 1,
   - $H \cap S = \varnothing$ on the points of S labeled 0.

   Equivalently, inside S, the subset H should coincide with T. Outside S, we are free to add or not add points, as long as the total size of H is exactly k.

4. To achieve this, we need:

   - $|T| = t \;\le\; k$, so that it is possible to have at least t points in H.
   - We also need $k - t$ further points (to reach total size k) chosen **outside** S. Hence we need $|X \setminus S| = n - d \ge k - t$.

Because $d = \min\{k,\, n-k\}$, both conditions can be met for **any** $t \le d$. Concretely:

- If $d = k$, then $k \le n - k$ so $n - d \ge k$. We can always pick the needed $k - t$ points from outside S.
- If $d = n - k$, then $n - k \le k$, so $t \le n - k$ is always $\le k$.

Hence we can always choose H so that $H \cap S = T$ and $|H| = k$. This shows that any subset S of size $d = \min\{k,\, n-k\}$ is shattered by $\mathcal{H}_{=k}$. Consequently,

$$\mathrm{VCdim}(\mathcal{H}_{=k}) \;\ge\; \min\{k,\, n-k\}.$$

2. Showing $\mathrm{VCdim}(\mathcal{H}_{=k}) \;\le\; \min\{k,\, n-k\}$

We now argue that **no** subset $S \subseteq X$ of size bigger than $\min\{k,\, n-k\}$ can be shattered.

- If $|S| > k$, then it is impossible to label **all** points of S by 1's, because each hypothesis h can only have k points labeled as 1 in the entire domain. Thus we cannot realize the labeling "all 1's" on S. Hence, no set S of size larger than k can be shattered.
- If $|S| > n - k$, then it is impossible to label **all** points of S by 0's. Indeed, that labeling would require the hypothesis h to have all its k "1"s **outside** S. But there are only $n - |S| < k$ points outside S, so we cannot fit all k ones outside S. Therefore the labeling "all 0's on S" cannot be realized if $|S| > n - k$.

Either way, we see that if $|S|$ exceeds $\min\{k,\, n-k\}$, there is at least one labeling of S that no function in $\mathcal{H}_{=k}$ can realize. Thus no set larger than $\min\{k,\, n-k\}$ can be shattered.

Putting these two arguments together completes the proof that

$$\mathrm{VCdim}(\mathcal{H}_{=k}) \;=\; \min\{k,\, n-k\}.$$

- A function in $\mathcal{H}_{=k}$ labels exactly k points as 1 (and the rest 0).
- To shatter a set S of size d, we need to be able to produce every possible 0–1 labeling of S.
- But any labeling of S with t ones forces the hypothesis to place those t ones inside S and the other $k - t$ ones **outside** S.
- We can only do this for **all** $t \leq d$ if $d \leq k$ (to allow t up to d) and $d \leq n - k$ (so that we can place up to k ones outside if needed).
- Hence the maximum d that can be shattered is $\min(k, n-k)$.

Exercise 9

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

Solution

The Class of Signed Intervals

We are working over the real line. Each hypothesis in

$$\mathcal{H} = \{ h_{a,b,s} : \mathbb{R} \to \{-1,+1\} \mid a \leq b, \ s \in \{-1,+1\}\}$$

is defined by an interval $[a, b] \subseteq \mathbb{R}$ and a sign $s \in \{-1, +1\}$. Concretely,

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

Equivalently, $h_{a,b,+1}$ is $+1$ on $[a, b]$ and -1 outside, while $h_{a,b,-1}$ is -1 on $[a, b]$ and $+1$ outside.

We claim that the VC dimension of this class is 3. That is:

1. There is a set of 3 points in $\mathbb{R}$ that can be **shattered** by $\mathcal{H}$.
2. No set of 4 points can be shattered by $\mathcal{H}$.

Below is the sketch of the argument.

1. Shattering 3 Points

Let $x_1 < x_2 < x_3$ be three distinct real numbers. We show that for **every** desired labeling of $\{x_1, x_2, x_3\}$ by $\{-1, +1\}$, we can pick an interval $[a, b]$ and sign s so that $h_{a,b,s}$ matches that labeling.

In essence, $h_{a,b,s}$ can change its sign **at most twice** along the real line (once at $x = a$ and once at $x = b$). But for three points, that suffices to realize all $2^3 = 8$ labelings. For example:

- To label all three as +1, pick $s = +1$ and let $[a, b]$ cover all three points.
- To label $(x_1 = +1, x_2 = -1, x_3 = +1)$, choose $s = -1$ and pick $[a, b]$ so that $x_2 \in [a, b]$ but $x_1, x_3 \notin [a, b]$. Then x_2 is labeled -1, and x_1, x_3 are labeled $+1$.
- Etc.

One can systematically verify that for every possible triple of labels in $\{-1, +1\}^3$, there is a suitable $[a, b]$ and sign s. Hence $\{x_1, x_2, x_3\}$ is shattered, and so

$$\mathrm{VCdim}(\mathcal{H}) \;\geq\; 3.$$

2. Inability to Shatter 4 Points

Consider four points $x_1 < x_2 < x_3 < x_4$. A hypothesis $h_{a,b,s}$ can have **at most two sign changes** along the real line:

- one change at $x = a$ (where the function switches from $-s$ to $+s$ if $s = +1$, or vice versa if $s = -1$),
- and one change at $x = b$.

Thus, any labeling realized by $h_{a,b,s}$ can switch signs **no more than twice** when reading from left to right. But among the $2^4 = 16$ possible ways to label $\{x_1, x_2, x_3, x_4\}$, there are labelings with **three or more sign changes**, for example:

$$(x_1 = +1,\ x_2 = -1,\ x_3 = +1,\ x_4 = -1),$$

which has three sign flips (+1 to -1, then -1 to +1, then +1 to -1). Such a labeling is impossible to realize with just a single interval $[a, b]$ and sign s. Therefore, no set of 4 points can be shattered, implying

$$\mathrm{VCdim}(\mathcal{H}) \;\leq\; 3.$$

Combining the two parts:

- We can shatter some set of 3 points,
- We cannot shatter any set of 4 points,

we obtain

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