16. VC dimension

Friday, February 9, 2024 5:01 PM

Def The VC ($\leftarrow$ Vapnik + Chervonenkis '71) dimension of $\mathcal{H}$, $VCdim(\mathcal{H})$, is the maximal size of a set that can be shattered by $\mathcal{H}$, ie. $VCdim(\mathcal{H}) = \max_{m \in \mathbb{N}} m$ st. $\underbrace{\tau_{\mathcal{H}}(m) = 2^m}_{\text{if } |Y|=2}$

via our no-free-lunch corollary 6.4, for $Y = \{\pm 1\}$, if $VCdim(\mathcal{H}) = \infty$, then $\mathcal{H}$ isn't PAC learnable

$d = VCdim(\mathcal{H})$ means $\exists$ at least one set of size d that is shattered, and $\nexists$ any sets of size $d+1$ that are shattered. } not as tight as Rademacher complexity

Ex continue our thresholding on $X = \mathbb{R}$ $\mathcal{H} = \{h_a : a \in \mathbb{R}\}$

$\tau_{\mathcal{H}}(1) = 2^1$

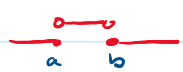
$\tau_{\mathcal{H}}(2) = 3 < 2^2$ so $VCdim(\mathcal{H}) = 1$ (= # parameters in $\mathcal{H}$)

$\uparrow$ often but not always true

Exercise (in-class), §6.3.2 in [SS]

$\mathcal{H} = \{h_{a,b} : a,b \in \mathbb{R}\}$ where $h_{a,b}(x) = \mathbb{1}_{x \in (a,b)} = \begin{cases} 1 & x \in (a,b) \\ 0 & \text{else} \end{cases}$

$X = \mathbb{R}$, $Y = \{0,1\}$

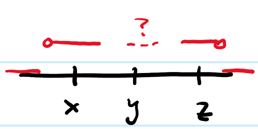


Find $VCdim(\mathcal{H})$

- $VCdim(\mathcal{H}) \geq 2$ since can shatter the set $C = \{3,4\}$

Not all sets of size 2 are shattered, eg. $C = \{3,3\}$ isn't. That's ok

- $VCdim(\mathcal{H}) < 3$ since it can't shatter any set of 3 pts, $C = \{x,y,z\}$



how can we get label $\begin{bmatrix}1\\0\\1\end{bmatrix}$?

let "x" be smallest ...

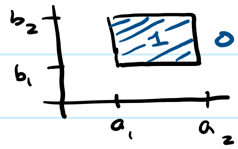
- So $VCdim(\mathcal{H}) = 2$ (again, = # parameters)

16a. VC dimension (p. 2)

Saturday, February 10, 2024 7:04 AM

Example Axis-Aligned rectangles in $\mathbb{R}^2$ (§6.3.3 in [SS])

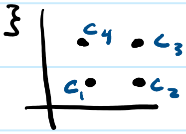
$X = \mathbb{R}^2$, $Y = \{0,1\}$, $\mathcal{H} = \{ h_{a_1,a_2,b_1,b_2} : a_1,a_2,b_1,b_2 \in \mathbb{R},\ a_1 \le a_2,\ b_1 \le b_2 \}$



Let's start w/ $m=4$

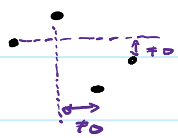
- Not all sets w/ 4 elements are shattered

eg. $\{(1,1), (1,2), (2,2), (2,1)\}$ (c_1, c_2, c_3, c_4)



If c_2, c_4 have same label, so must c_1

... but $\exists$ a set w/ 4 elements that is shattered



Try it! Draw boxes around all $2^4 = 16$ possible combinations

eg. Can we get $(0,0,0,0)$? yes.

Can we get $(1,0,0,0), (0,1,0,0), \dots, (0,0,0,1)$? yes

Can we get all pairs $(1,1,0,0), (1,0,1,0), \dots$? yes

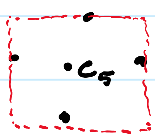
...

So $\mathrm{VCdim}(\mathcal{H}) \ge 4$

- $m=5$, $C = \{c_1, c_2, \dots, c_5\}$ if c_i not unique, clearly can't be shattered.

Re-arrange labels so c_5 isn't the top-most, isn't bottom-most, isn't left-most, ...

(if we have any ties for, eg., top-most, then we can't shatter for same reason as example above)



i.e. $c_5 \in \mathrm{Conv}(c_1, \dots, c_4)$ ← we'll use more precise geometric things later

convex hull, intersection of all convex sets containing the set

So label $(1,1,1,1,0)$ is impossible since if $\underbrace{\text{rectangle}}_{\text{convex}}$ covers $c_1, \dots, c_4$

$\Rightarrow$ rectangle covers c_5

So

- $\mathrm{VCdim}(\mathcal{H}) = 4$ (= # parameters!)

Cautionary Example Exer. 6.8 [SS]

$X = \mathbb{R}$, $Y = \{0,1\}$, $\mathcal{H} = \{ h_\theta : \theta \in \mathbb{R} \}$ where $h_\theta(x) = \lceil \sin(\theta x) \rceil$ ← ceiling function

$\mathcal{H}$ has 1 parameter but $\mathrm{VCdim}(\mathcal{H}) \ne 1$ (what do you think it is?)

Example $|\mathcal{H}| < \infty$

$\mathcal{H}_C = \{ (h(c_1), \dots, h(c_m)) : h \in \mathcal{H} \}$ so $|\mathcal{H}_C| \le |\mathcal{H}|$

so when (for $|Y|=2$ case) $|\mathcal{H}| < 2^m$ (ie $m > \log_2(|\mathcal{H}|)$)

then no set C of size m can be shattered.

So... $\mathrm{VCdim}(\mathcal{H}) \le \log_2(|\mathcal{H}|)$

16b. VC dimension (p. 3)

Saturday, February 10, 2024 7:27 AM

Example : Hyperplanes in 2D (see Ex. 3.12 Mohri, or Thm 9.2 in [SS])

$X = \mathbb{R}^2$, $Y = \{\pm 1\}$, $\mathcal{H}_0 = \{ x \mapsto \text{sign}(\langle w, x \rangle) : w \in \mathbb{R}^2 \}$

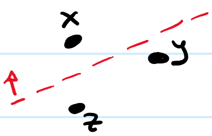
$\mathcal{H} = \{ x \mapsto \text{sign}(\langle w, x \rangle + b) : w \in \mathbb{R}^2,\ b \in \mathbb{R} \}$

It's obvious $\text{VCdim}(\mathcal{H}_0) \le \text{VCdim}(\mathcal{H})$ since $\mathcal{H}$ is more expressive

In fact (in any dimension) $\text{VCdim}(\mathcal{H}_0) = \text{VCdim}(\mathcal{H}) - 1$. (needs to be proven)

Let's focus on $\mathcal{H}$ (not $\mathcal{H}_0$).

- $m = 3$, $C = \{x, y, z\}$ let x, y, z be anything not collinear (not ...)



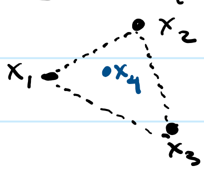
Can we shatter it ? i.e. achieve all $2^3 = 8$ "dichotomies" ?

Yes! (play around with it)

[What are hyperplanes? Divide space into 2 halves, w/ normal direction w]

- $m = 4$ $|C| = 4$. Claim no set of size 4 can be shattered

case 1: $\exists\, x_i \in \text{interior}(\text{conv}(C))$, i.e. $x_4 = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$, $\alpha_i > 0$, $\sum \alpha_i = 1$



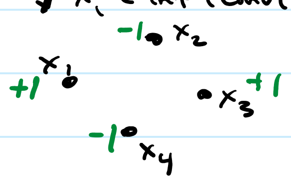
$\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$ isn't achievable

If $\langle w, x_i \rangle + b \ge 0$ for $i = 1, 2, 3$

then $\langle w, x_4 \rangle + b = \sum \alpha_i (\langle w, x_i \rangle + b) \ge 0$

so x_4 must have same label.

case 2: $\nexists\, x_i \in \text{int}(\text{conv}(C))$



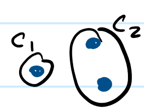
$\begin{bmatrix} +1 \\ -1 \\ +1 \\ -1 \end{bmatrix}$ isn't achievable, proof via pictures for now

- So, $\text{VCdim}(\mathcal{H}) = 3$ (in $\mathbb{R}^2$). And (not proven) $\text{VCdim}(\mathcal{H}_0) = 2$

Radon's Thm Let $C \subseteq \mathbb{R}^d$, then if $|C| = d + 2$, it can be partitioned into two sets C_1 and C_2 ($C_1 \cup C_2 = C$, $C_1 \cap C_2 = \emptyset$) s.t.

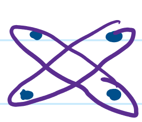
$\text{conv}(C_1) \cap \text{conv}(C_2) \ne \emptyset$

ex in 2D $|C| = 3 < d + 2$, not true:

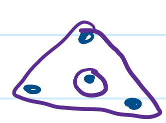


don't intersect

$|C| = 4 = d + 2$



or



Do overlap ✓

Example: Hyperplanes in $\mathbb{R}^d$ same setup, $\mathcal{H} = \{ x \mapsto \text{sign}(\langle w, x \rangle + b), b \in \mathbb{R}, w \in \mathbb{R}^d \}$

$X = \mathbb{R}^d$, $Y = \{\pm 1\}$ $\mathcal{H}_0$ similar but $b = 0$

Can analyze $\mathcal{H}^d$ as $\mathcal{H}_0^{d+1}$ or vice-versa.

- $\text{VCdim}(\mathcal{H}) \geq d+1$

  canonical basis $e_2 = (0, 1, 0, 0, \dots, 0)$

  i.e. choose $m = d+1$ points $c_1, \dots, c_d, c_{d+1}$ via $c_i = e_i$ $i \in [d]$, $c_{d+1} = 0$

  then i^{th} coordinate of w_i only affects c_i

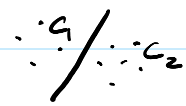
  so to label c_i ± 1, set $w_i = \pm 1$

  To label $c_{d+1} = 1$, set $\underbrace{b = \tfrac{1}{2}}_{|1/2| < |1| \text{ so it doesn't affect sign}(\dots) \text{ for others}}$, and $c_{d+1} = -1$ via $b = -\tfrac{1}{2}$

- $\text{VCdim}(\mathcal{H}) < d+2$

  Use Radon's Thm, partition any C into $C_1 \cup C_2$, $\text{conv}(C_1) \cap \text{conv}(C_2) \neq \emptyset$

  Then if we labeled all of C_1 with $+1$ and all of C_2 with -1, that means we can separate C_1 and C_2 with a hyperplane

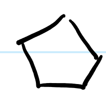
  

  ... but that hyperplane would also separate $\text{conv}(C_1)$ from $\text{conv}(C_2)$ ... yet $C_1 \cap C_2 \neq \emptyset$ so that's impossible.

  (or, prove $\text{VCdim}(\mathcal{H}_0) < d+1$ using that $d+1$ pts in $\mathbb{R}^d$ must be linearly dependent, cf §9.1 [SS])

light on details

Example Convex d-gons in $X = \mathbb{R}^2$ (polygons P with d edges)



$Y = \{0, 1\}$, $h = \mathbb{I}_P$, $h(x) = \begin{cases} 1 & x \in P \\ 0 & x \notin P \end{cases}$

Fact: $\exists$ a set of $2d+1$ points that can be shattered (in fact, put on a circle)

ex: triangles, $d = 3$ so $m = 2d+1 = 7$ cf. Mohri

Fact: No set of $m > 2d+1$ pts. can be shattered

So $\text{VCdim}(d\text{-gons}) = 2d+1$

Implication $\text{VCdim}(\text{convex indicator functions})$

$\geq \text{VCdim}(d\text{-gons}) = 2d+1 \quad \forall d \in \mathbb{N}$ ... so $\text{VCdim}(\text{convex} \dots) = \infty$