Subject : هلیا جوان‌مرد Year : Month : Date :

«Exercises chapter 6»

1\. $VCdim(H_{=k}) = \min\{k, |X|-k\}$ 6_2 :

باید نشان دهیم $VCdim(H_{=k}) \le \min\{k, |X|-k\}$. اگر سایز مجموعه C را $k < |C| \le X$
یا $k+1$ در نظر بگیریم نابرابری طبق مجموعه فرضیات، تابع h ای وجود ندارد که به ازای تمام
x ها در C، label 1 برگرداند. همچنین اگر اندازه‌ی مجموعه C را $|X|-k+1$ در نظر
بگیریم تابع h به طوری که به ازای تمام $x \in C$، label 0 برگرداند وجود ندارد.
در نتیجه $VCdim(H_{=k}) \le \min\{k, |X|-k\}$. اکنون اگر نشان دهیم
$VCdim(H_{=k}) \ge \min\{k, |X|-k\}$، می‌توانیم نتیجه بگیریم $VCdim(H_{=k}) = \min\{k, |X|-k\}$

برای نشان دادن صحت دوم:

$C \subseteq X \quad C = \{x_1, \dots, x_m\} \quad m \le \min\{k, |X|-k\}$

و $\{y_1, \dots, y_m\}$ مجموعه label های C است.

در این صورت اگر $h \in H$ را تابعی در نظر بگیریم که برای $x_i \in C$، $y_i$ را و
به ازای x هایی که در $X - C$ قرار دارند و label 1 دارند، 1 را برگرداند
می‌توانیم نتیجه بگیریم که C توسط H shatter می‌شود، $VCdim(H) \ge \min\{k, |X|-k\}$
(چون حداکثر k تا داریم و $m \le \min\{k, |X|-k\}$ است)

Parsian

Subject : Year : Month : Date :

2. $H_{at\text{-}most\text{-}k} = \{h \in \{0,1\}^X : |\{x : h(x)=1\}| \le k \text{ or } |\{x : h(x)=0\}| \le k\}$

VC dim = k

اگر $C \subseteq X$ ، $k+1$ عضو داشته باشد، بنابراین هیچ h ای در H وجود ندارد

که به ازای تمام $x \in C$ ، label ۱ را برگرداند. بنابراین $vcdim \le k$

اگر $C = \{x_1, \dots, x_m\} \subseteq X$ ، $|C| = m \le k$ ، $(y_1, \dots, y_m)$

برای labelها باشد، یک $h \in H$ وجود دارد، برای همه‌ی $x_i \in C$ ، $h(x_i) = y_i$ ،

برای x های عضو $X - C$ ، $h(x) = 0$ . بنابراین C توسط H ، shatter می‌شود

و $vcdim \ge k$

$vcdim \le k,\ vcdim \ge k \longrightarrow vcdim = k$

6_4

حالت اول، $|H_A| < |\{B \subseteq A : H \text{ shatters } B\}| = \sum_{i=0}^{d} \binom{|A|}{i}$

مجموعه فرضیات دایره‌های هم مرکز را در نظر می‌گیریم. $H = \{1_{\{\|x\| \le r\}} : r \ge 0\}$

$x \in \mathbb{R}^d$ , $d \ge 2$ . می‌دانیم $vcdim\ H = 1$ زیرا برای

دو نقطه، هیچ تابعی وجود ندارد اگر $\|x_1\| < \|x_2\|$ ، $(0,1)$ را

برگرداند. در نتیجه دو نقطه توسط H ، shatter نمی‌شود.

Parsian

Subject : Year : Month : Date :

$A = \{(0,1),(1,0)\}$

$a = |H_A| = |\{(0,0),(1,1)\}| = 2$

$b = |\{B \subseteq A : H \text{ shatters } B\}| = |\{\{(0,1)\},\{(1,0)\},\phi\}| = 3$

$c = \sum_{i=0}^{1} \binom{2}{i} = \binom{2}{0} + \binom{2}{1} = 1 + 2 = 3$

$2 < 3 = 3$ ☑

حالت دوم:

$a = b = c$

H: دایره‌های هم‌مرکز

$A = \{(1,1)\}$

$|H_A| = |\{0,1\}| = 2$

$|\{B \subseteq A : H \text{ shatters } B\}| = |\{\{(1,1)\},\phi\}| = 2$

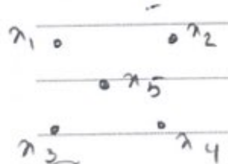
$\sum_{i=0}^{1} \binom{1}{i} = \binom{1}{0} + \binom{1}{1} = 1 + 1 = 2$

$2 = 2 = 2$ ☑

حالت سوم:

$a = b < c$

مجموعه فرضیات axis-aligned rectangles را در نظر می‌گیریم. می‌دانیم




$\text{vcdim } H = 4$. (اگر $x_1, \dots, x_5$ به شکل مقابل باشند:

هیچ تابعی وجود ندارد که table $(1,1,1,1,0)$ را ایجاد کند

Parsian

$A = \{\underbrace{(0,0)}_{x_1}, \underbrace{(0,1)}_{x_2}, \underbrace{(0,2)}_{x_3}\}$

$|H_A| = |\{(0,0,0), (1,0,0), (0,0,1), (1,1,0), (0,1,1)$

$, (0,1,0), (1,1,1)\}| = 7$

$|\{B \subseteq A : H \text{ shatters } B\}| = |\{\phi, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1, x_2\},$

$\{x_2, x_3\}, \{x_1, x_3\}\}| = 7$

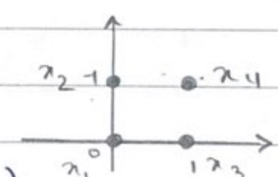
$\sum_{i=0}^{4} \binom{3}{i} = \binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3} = 1 + 3 + 3 + 1 = 8$

$7 = 7 < 8$ ☑

$a < b < c$ حالت چهارم:

مجموعه فرضیات : axis_Aligned rectangles

$A = \{\underbrace{(0,0)}_{x_1}, \underbrace{(0,1)}_{x_2}, \underbrace{(1,0)}_{x_3}, \underbrace{(1,1)}_{x_4}\}$




$|H_A| = |\{(1,0,0,0), (0,1,0,0), (0,0,0,1)$

$, (0,0,1,0), (1,1,0,0), (1,0,1,0), (0,0,1,1), (1,0,0,1),$

$(0,0,0,0), (1,1,1,1)\}| = 10$

$|\{B \subseteq A : H \text{ shatters } B\}| = |\{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_1, x_2\},$

$\{x_1, x_3\}, \{x_1, x_4\}, \{x_2, x_3\}, \{x_2, x_4\}, \{x_3, x_4\}, \phi\}| = 11$

Subject : Year : Month : Date :

$$\sum_{i=0}^{4}\binom{4}{i} = \binom{4}{0}+\binom{4}{1}+\binom{4}{2}+\binom{4}{3}+\binom{4}{4} = 16$$

$10 < 11 < 16$ ☑

---

۶-۶ :

۱- هر h می‌تواند با سه حالت $x_i$ ، $\bar{x}_i$ یا هیچ‌کدام از این‌ها در conjunction
مشخص شود. همچنین تابع all_negative هم وجود دارد. بنابراین چون در
training set ، d عضو وجود دارد: $|H| \geq 3^d + 1$

۲- اگر $VCdim\, H = a$ بنابراین اندازه H باید بزرگ‌تر یا مساوی $2^a$
باشد، بنابراین داریم (طبق تعریف VC dim) $|H| \geq 2^a \to \log|H| \geq a$

$\to VCdim \leq \log(|H|) \leq \log(3^d+1) \leq \log 3^d = d\log 3$

۳- می‌خواهیم نشان دهیم $H^d_{con}$ توسط $C=\{e_i : i \leq d\}$ shatter
می‌شود. پس $VCdim\, H^d_{con} \geq d$. اگر lable همه x های c ، ۱ باشد
در این صورت $h = (h_{empty})$ all_positive ؛ اگر lable همه x ها
۰ باشد، $h^-$ (all_negative). در غیر این صورت اگر a عضو c ، lable

۲ داشته باشد و $J$ مجموعه اندیس های این $a$ عضو باشد، $h$ را بصورت زیر

در نظر می‌گیریم: $h = \bigwedge_{i \in J} x_i$

چون تعداد متغیرها $d$ تا است و $2^d$ در نتیجه ثابت کردیم $C$ توسط $H$، Shatter می‌شود

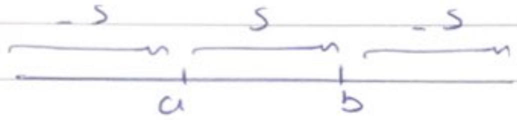
VC dim = 3 86-9

باید ثابت کنیم اولاً یک زیرمجموعه 3 عضوی از H، shatter می‌شود، دوماً تمام

زیرمجموعه‌های 4 عضوی، توسط H shatter نمی‌شوند.

اگر $C \subseteq X = \{x_1, x_2, x_3\}$ ، $x_1 < x_2 < x_3$ تمام حالت‌ها

در H وجود خواهند داشت:



$(-1,-1,-1)$ $x_1, x_2, x_3 < a$ $S = 1$

$(-1,-1,1)$ $x_1, x_2 < a$ , $a < x_3 < b$

$(-1,1,1)$ $x_1 < a$, $a < x_2, x_3 < b$

$(1,1,1)$ $a < x_1, x_2, x_3 < b$

$(1,-1,-1)$ $a < x_1 < b$ , $x_2, x_3 > b$

$(1,1,-1)$ $a < x_1, x_2 < b$ , $x_3 > b$

$(-1,1,-1)$ $x_1 < a$ , $a < x_2 < b$, $x_3 > b$

$(1,-1,1)$ $x_1 < a$, $a < x_2 < b$, $x_3 > b$ $S = -1$

اگر $C \subseteq X = \{x_1, x_2, x_3, x_4\}$ ، $x_1 < x_2 < x_3 < x_4$ هیچ تابعی

در H وجود ندارد که حالت $(-1,1,-1,1)$ یا $(1,-1,1,-1)$

را برگرداند در نتیجه هیچ مجموعه 4 عضوی توسط H shatter نمی‌شود.

Subject :

Year : Month : Date :

8 6_10

1- $VCdim \geq d$ ، یک مجموعه C با d عضو در نظر می‌گیریم که توسط H، shatter

می‌شود. $C \subseteq X$ ، بنابراین می‌توانیم $X = C$ در نظر بگیریم. در این صورت

طبق تمرین 3 در فصل 5 : اگر $|X| \geq km$ : (طبق No_Free_lunch)

یک f وجود دارد که $L_D(f) = 0$ ، $E_{S\sim D^m}[L_D(A(S))] \geq \frac{k-1}{2k}$

می‌دانیم $m < d$ ، $|X| = d \leftarrow d = km \leftarrow k = \frac{d}{m}$

بنابراین اگر به جای k، $\frac{d}{m}$ قرار دهیم داریم:

$$E_{S\sim D^m}[L_D(A(S))] \geq \frac{d-m}{2d} + \min_{h\in H} L_D(h)$$

2- حال می‌دانیم $VCdim(H) = \infty$ ، اگر در نظر بگیریم یک training set m تایی

می‌توانیم این الگوریتم را learn کنیم، چون $VCdim = \infty$ است. یک مجموعه (C)

$am$ $(a > 2)$ تایی وجود دارد که توسط H، shatter می‌شود بنابراین چون

سایز training set از نصف C کمتر است طبق th No_Free_Lunch، H

PAC learnable نیست. پس می‌توانیم نتیجه بگیریم اگر H، PAC learnable باشد

$VCdim < \infty$ .

Parsian

۱- فرض می‌کنیم برای هر $i \in [r]$ ، $\mathrm{VCdim}\, H_i \geq d \geq 3$

$$H = \bigcup_{i=1}^{r} H_i$$

اگر $k < d$ : $\tau_H(k) = 2^k$

$$\tau_H(k) = \max_{C \subset X, |C| = k} |H_C| = \max \left|\bigcup_{i=1}^{r} H_i\right| \underset{\text{union bound}}{\leq} \max \left|\sum_{i=1}^{r} H_i\right|$$

$$\leq \sum_{i=1}^{r} \max |H_i| = \sum_{i=1}^{r} \tau_{H_i}(k)$$

$$\tau_H(k) \leq \sum_{i=1}^{r} \tau_{H_i}(k) \leq \sum_{i=1}^{r} \left(\frac{ek}{d}\right)^d \leq \sum_{i=1}^{r} ek^d \leq \sum_{i=1}^{r} k^d = rk^d$$

$$2^k \leq rk^d \longrightarrow k \leq \log r + d \log k$$

Lemma A.2 : $x \geq 4\log(2a) + 2b \to x \geq a\log(x) + b$

$$\longrightarrow k \leq 4d\log(2d) + 2\log r$$

$$\longrightarrow \mathrm{VCdim}\, H \leq k \leq 4d\log(2d) + 2\log r$$

۲- کافی است نشان دهیم هیچ مجموعه‌ی $2d+2$ عضوی توسط $H$، shatter نمی‌شود. $H = H_1 \cup H_2$ ، $\mathrm{VCdim}\, H_1 , \mathrm{VCdim}\, H_2 = d$

اگر نشان دهیم برای $k \geq 2d+2$ ، تعداد اعضای $H$ $(\tau_H(k))$ از $2^k$ کمتر است یعنی می‌توانیم، هیچ مجموعه‌ی بزرگ‌تر یا مساوی $2d+2$ توسط $H$ shatter نمی‌شود

Parsian

Subject :   Year :   Month :   Date :

$$T_H(k) \leq T_{H_1}(k) + T_{H_2}(k)$$ (در قسمت قبل ثابت کردیم)

$$T_H(k) \leq \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=0}^{d} \binom{k}{i} = \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=0}^{d} \binom{k}{k-i}$$

$$= \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=k-d}^{k} \binom{k}{i} \leq \sum_{i=0}^{d} \binom{k}{i} + \sum_{\substack{i=d+2 \\ k \leq 2d+2}}^{k} \binom{k}{i}$$

$$< \sum_{i=0}^{d} \binom{k}{i} + \sum_{i=d+1}^{k} \binom{k}{i} = \sum_{i=0}^{k} \binom{k}{i} = 2^k$$