۶.۲

۱)

Domain $= X < \infty \qquad k \leq |X|$

$H_k^X = \{h \in \{0,1\}^X : |\{x : h(x)=1\}| = k\} \qquad VCD(H_k^X) = ?$

H_k^X یعنی مجموعه توابعی که دقیقاً به k تا x، برچسب ۱ می دهند.

نشان می دهیم که $VCDim(H_k^X) = \min\{k, |X|-k\}$ است.

اگر در نظر بگیریم $C_1 = \{c_1, \dots, c_{k+1}\}$ ، $C_1 \subseteq X$ ، آنگاه هیچ تابعی وجود ندارد $h \in H_k^X$ که به همهٔ c_i ها لیبل ۱+ بدهد چون فقط می تواند k تا را لیبل ۱ بدهد.

و اگر در نظر بگیریم $C_2 = \{c_1, \dots, c_{|X|-k}\}$ ، $C_2 \subseteq X$ آنگاه هیچ تابعی وجود ندارد $h \in H_k^X$ که به همهٔ c_i ها لیبل ۰ بدهد. چون k تا لیبل ۱ داده می.

بنابراین

$$\boxed{VCDim(H_k^X) \leq \min\{k, |X|-k\}} \quad (1)$$

اگر در نظر بگیریم $C_3 = \{c_1, \dots, c_m\}$ بطوری که $m \leq \min\{k, |X|-k\}$ و $(y_1, \dots, y_m)$ برچسب آنها باشند که $y_i \in \{0,1\}^X$

در نظر می گیریم یک تابعی را $h \in H$ که به هر c_i ، برچسب آن را می دهد یعنی $h(c_i) = y_i$ در اینصورت حداکثر k عضو دارای برچسب ۱ خواهند بود

و به هر x در $X \setminus C$ برچسب ۱ می دهیم

بنابراین k نقطه وجود دارند برچسب ۱ دارند بنابراین C_3 توسط H_k^X خرد می شود بنابراین

$$(2) \quad \boxed{VCD(H_k^X) \geq \min\{k, |X|-k\}}$$

$$\xrightarrow{(1),(2)} VCDim(H_k^X) = \min\{k, |X|-k\}$$

ادامه ۶.۲

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

اگر در نظر بگیریم $C_1 = \{c_1, \dots, c_{k+1}\}$ ، $C_1 \subseteq X$ آنگاه هیچ تابع h ای $h \in H_{at}$ وجود ندارد که همه c_i ها برچسب ۱ بدهد (چون بیشتر مقادیر k تا می‌تواند لیبل ۱ بدهد) پس $VCD(H_{at}) \le k$ ①

و اگر $C_2 = \{c_1, \dots, c_k\}$ که $C_2 \subseteq X$ و $(y_1, \dots, y_k)$ برچسب‌ها آنها باشند $\forall i \in \{0,1\}^k$ در نظر بگیریم h تابعی است که $h(c_i)=1$ باشد اگر $y_i=1$ است برای باقی نقاط $h(x)=0$ باشد در اینصورت محقق است که حداکثر k عنصر لیبل ۱ دارند.

بنابراین C_2 توسط H_{at} خرد می‌شود $\Rightarrow$ $VCDim(H_{at}) \ge k$ ②

$$①,② \Longrightarrow VCDim(H_{at}) = k$$

۶.۴

از تمرین و اصول مربوط، اثبات مقصود بوده اما تش رو نفهمیدم نتوانستم حل کنم.

۶.۷

۱) چون برای هر i، یا x_i یا $\bar{x}_i = 1-x_i$ یا هیچ‌کدام می‌آید، یک حالت هم فرضیه تماماً منفی وجود دارد، بنابراین

$$|H^d_{con}| = 3^d + 1$$

۲)

$$VCD(H^d_{con}) \overset{?}{\leq} d\log 3$$

طبق سوال بالا داریم VCDim :

$$VCdim(H) \leq \log_2(|H|)$$

$$\Rightarrow VCdim(H) \leq \log_2(3^d + 1)$$

$$= \log_2 3^d + \cancel{\log_2 1}$$

$$= d\log_2 3$$

باقی سوالات مثل تمرین ۴ از فصل ۳ را اصلاً نفهمیدم نتوانستم حل کنم.

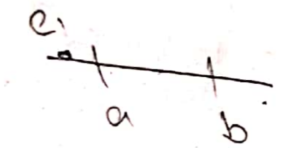
۶.۹

$$H=\{h_{a,b,s}: a\leq b,\ s\in\{-1,1\}\}$$

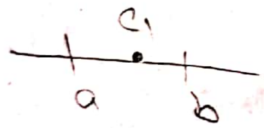
$$h_{a,b,s}(x)=\begin{cases} s & \text{if } x\in[a,b] \\ -s & \text{if } x\notin[a,b]\end{cases}$$

$VCDim(H)$?

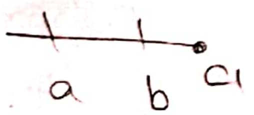
$C=\{c_1\}$



$s=+1 \Rightarrow h(c_1)=-1$
$s=-1 \Rightarrow h(c_1)=+1$



$s=+1 \Rightarrow h(c_1)=+1$
$s=-1 \Rightarrow h(c_1)=-1$

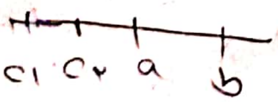


$s=+1 \Rightarrow h(c_1)=-1$
$s=-1 \Rightarrow h(c_1)=+1$

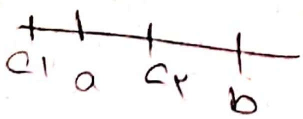
پس H یک نقطه را خرد می‌کند یعنی هم $(+1)$ را می‌دهد هم (-1) برای هر $s\in\{\pm1\}$

$C=\{c_1,c_1\}$

حال ۲ نقطه را بررسی می‌کنیم. $c_1<c_2$



$\begin{cases} s=+1 \\ s=-1\end{cases}$
$\Rightarrow h(c_1)=-1,\ h(c_2)=-1$ $(-1,-1)$
$\Rightarrow h(c_1)=+1,\ h(c_2)=+1$ $(+1,+1)$



$\begin{cases} s=+1 \\ s=-1\end{cases}$
$\Rightarrow h(c_1)=-1,\ h(c_2)=+1$ $(-1,+1)$
$\Rightarrow h(c_1)=+1,\ h(c_2)=-1$ $(+1,-1)$

فاطمه داوری یادگیری ماشین

ادامه 4.9 _ 1

$a \; c_1 \; c_2 \; b$

$$\begin{cases} S=+1 \Rightarrow h(c_1)=h(c_2)=+1 & (+1,+1) \\ S=-1 \Rightarrow h(c_1)=h(c_2)=-1 & (-1,-1) \end{cases}$$

$a \; c_1 \; b \; c_2$

$$\begin{cases} S=+1 \Rightarrow h(c_1)=+1 ,\ h(c_2)=-1 & (+1,-1) \\ S=-1 \Rightarrow h(c_1)=-1 ,\ h(c_2)=+1 & (-1,+1) \end{cases}$$

$a \; b \; c_1 \; c_2$

$$\begin{cases} S=+1 \Rightarrow h(c_1)=h(c_2)=-1 & (-1,-1) \\ S=-1 \Rightarrow h(c_1)=h(c_2)=+1 & (+1,+1) \end{cases}$$

پس H دو نقطه را نیز خورد می‌شکند یعنی هر ۴ حالت $(+1,+1)$, $(-1,-1)$, $(+1,-1)$, $(-1,+1)$ را برچسب می‌گذارد. برای هر $S \in \{\pm 1\}$

پس $VCDim \geq 2$

حال ۳ نقطه را بررسی می‌کنیم $C=\{c_1,c_2,c_3\}$ $c_1<c_2<c_3$

هر چهار حالت $c_1 \; c_2 \; c_3 \; a \; b$ و $a \; b \; c_1 \; c_2 \; c_3$ و $c_1 \; c_2 \; a \; b \; c_3$ و $c_1 \; a \; b \; c_2 \; c_3$ دارای پاسخ یکسانی است

$$\begin{cases} \text{if } S=+1 \Rightarrow h(c_1)=h(c_2)=h(c_3)=-1 & (-1,-1,-1) \\ \text{if } S=-1 \Rightarrow h(c_1)=h(c_2)=h(c_3)=+1 & (+1,+1,+1) \end{cases}$$

ادامه ۹.۴ ۲

حالت ۱

$c_1 \quad c_2 \quad a \quad c_3 \quad b$

$$\begin{cases} \text{if } s=+1 \Rightarrow h(c_1)=h(c_2)=-1,\ h(c_3)=+1 & (-1,-1,+1) \\ \text{if } s=-1 \Rightarrow h(c_1)=h(c_2)=+1,\ h(c_3)=-1 & (+1,+1,-1) \end{cases}$$

حالت ۲

$c_1 \quad a \quad c_2 \quad c_3 \quad b$

$$\begin{cases} \text{if } s=+1 \Rightarrow h(c_1)=-1,\ h(c_2)=h(c_3)=+1 & (-1,+1,+1) \\ \text{if } s=-1 \Rightarrow h(c_1)=+1,\ h(c_2)=h(c_3)=-1 & (+1,-1,-1) \end{cases}$$

حالت ۳

$a \quad c_1 \quad c_2 \quad c_3 \quad b$

$$\begin{cases} \text{if } s=+1 \Rightarrow h(c_1)=h(c_2)=h(c_3)=+1 & (+1,+1,+1) \\ \text{if } s=-1 \Rightarrow h(c_1)=h(c_2)=h(c_3)=-1 & (-1,-1,-1) \end{cases}$$

حالت ۴

$a \quad c_1 \quad c_2 \quad b \quad c_3$

$$\begin{cases} \text{if } s=+1 \Rightarrow h(c_1)=h(c_2)=+1,\ h(c_3)=-1 & (+1,+1,-1) \\ \text{if } s=-1 \Rightarrow h(c_1)=h(c_2)=-1,\ h(c_3)=+1 & (-1,-1,+1) \end{cases}$$

حالت ۵

$a \quad c_1 \quad b \quad c_2 \quad c_3$

$$\begin{cases} \text{if } s=+1 \Rightarrow h(c_1)=+1,\ h(c_2)=h(c_3)=-1 & (+1,-1,-1) \\ \text{if } s=-1 \Rightarrow h(c_1)=-1,\ h(c_3)=h(c_2)=+1 & (-1,+1,+1) \end{cases}$$

یادگیری ماشین

فاطمه داوری

<u>ادامه ۹.۹</u>

۳

حالت

$c_1 \quad a \quad c_2 \quad b \quad c_3$

$$\begin{cases} \text{if } s=+1 \Rightarrow h(c_1)=h(c_3)=-1,\ h(c_2)=1 & (-1,+1,-1) \\ \text{if } s=-1 \Rightarrow h(c_1)=h(c_3)=+1,\ h(c_2)=-1 & (+1,-1,+1) \end{cases}$$

بنابراین اگر $C=\{c_1,c_2,c_3\}$ ، H ، C را خرد می‌کند. یعنی تمام حالات $(+1,+1,+1)$ ، $(-1,-1,-1)$ ، $(+1,+1,-1)$ ، $(+1,-1,+1)$ ، $(-1,+1,+1)$ ، $(-1,-1,+1)$ ، $(-1,+1,-1)$ ، $(+1,-1,-1)$ را برمی‌گرداند.

پس $VCDim(H) \geq ۳$

اما اگر $C=\{c_1,c_2,c_3,c_4\}$ باشد حالت $(-1,+1,-1,+1)$ را نمی‌توان از هیچ تابعی برگرداند

پس H ، این C را خرد نمی‌کند.

$\Rightarrow VCDim(H)=۳$

۶.۱۰

(۱

$H \quad X \longrightarrow \{0,1\}$

if $\quad VCDim(H) \geq d \quad \forall d$

برای هر توزیع D روی $X \times \{0,1\}$ برای هر سایز نمونه m می‌خواهیم نشان دهیم:

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

اگر در نظر بگیریم $C=\{c_1,\dots,c_d\}$, $C \subset H$, C را خرد کند. طبق راه‌حل‌های ارائه شده تمرین سه

$\exists\ f: X \longrightarrow \{0,1\} \ni L_D(f)=0$

$$\underset{S\sim D^m}{E}[L_D(A(S))] \geq \frac{1}{2} - \frac{1}{2K} \qquad |C| \geq Km$$

چون $|C|=d \Leftarrow$ پس $K \leq \frac{d}{m}$

$$\underset{S\sim D^m}{E}[L_D(A(S))] \geq \frac{1}{2} - \frac{m}{2d}$$

$$\Rightarrow \underset{S\sim D^m}{E}[L_D(A(S))] \geq 0 + \frac{1}{2} - \frac{m}{2d} = 0 + \frac{d-m}{2d}$$

با توجه به برقراری شرط realizability یعنی $\exists f: L_D(f)=0$ پس می‌توان نوشت

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

فاطمه دلاوری   یادگیری ماشین

ادامه ۹.۱۰

۲) اید ثابت کنیم برای هر H که PAC learnable است، $VCDim(H) < \infty$ است.

می توانیم از برهان خلف استفاده کنیم.

فرض میکنیم $VCDim(H) = \infty$ است. بنابراین برای هر مجموعه داده آموزشی به سایز m یک C به سایز $d = 2m$ توسط H خرد می شود. $|C| = d = 2m$

بنابراین طبق قسمت قبل

$$E_{S \sim D^m}[L_D(A(S))] \geq \frac{d-m}{2d} = \frac{2m-m}{4m} = \frac{1}{4}$$

حال طبق نتیجه قضیه مارکوف:

$$P(Z > a) \geq \frac{E(Z)}{a}$$

$a \in (0,1)$:

$$\Rightarrow P(Z > a) \geq \frac{E(Z) - a}{1 - a}$$

اگر در نظر بگیریم $a = \frac{1}{8}$

$$P_{S \sim D^m}[L_D(A(S)) > \frac{1}{8}] \geq \frac{E(L_D(A(S))) - \frac{1}{8}}{1 - \frac{1}{8}} = \frac{\frac{1}{4} - \frac{1}{8}}{\frac{7}{8}} = \frac{1}{7}$$

و این یعنی خطای حقیقی با احتمال بیشتر از $\frac{1}{7}$ از $\frac{1}{8}$ بیشتر است. اگر در نظر بگیریم $\frac{1}{8} = \epsilon$ و $\frac{1}{7} = \delta$ که $\epsilon, \delta \in (0,1)$ هستند این در تناقض با PAC learning بودن H است. چون وقتی H، PAC learnable است، خطای حقیقی با احتمال کمتر از $\frac{1}{7} = \delta$ از $\frac{1}{8}$ بیشتر است.

در نتیجه فرض خلف باطل و حکم برقرار است یعنی $VCDim(H)$ متناهی است.

6.11

(1

$H_1, \dots, H_r$ are hypothesis classes over some fixed domain X

$d = \max_i VCDim(H_i)$ $d \geq 3$

Prove that:

$$VCDim\left(\bigcup_{i=1}^{r} H_i\right) \leq 4d\log(2d) + 2\log(r)$$

فرض می‌کنیم $C = \{c_1, \dots, c_k\}$ که C توسط $\cup H_i$ ها خرد می‌شود.

اگر $k < d$ باشد: $\tau_{\cup H}(k) = 2^k$

از طرفی

$$\tau_{\cup H}(k) = \max_{C \subset X : |C| = k} |H_C| = \max \left|\bigcup_{i=1}^{r} H_{iC}\right| \overset{\text{union bound}}{\leq} \sum_{i=1}^{r} \max |H_{iC}|$$

$$\Rightarrow \tau_{\cup H}(k) \leq \sum_{i=1}^{r} \tau_{H_i}(k)$$

طبق Sauer's lemma:

$$\tau_{H_i}(k) < \left(\frac{ek}{d}\right)^d$$

$$\Rightarrow \tau_{H_i}(k) \leq \left(\frac{ek}{3}\right)^d < k^d$$

چون $d \geq 3$

چون $\left(\frac{e}{3}\right)^d < 1$

$$\Rightarrow \tau_{\cup H}(k) \leq \sum_{i=1}^{r} k^d = rk^d \quad \Rightarrow \boxed{\tau_{\cup H}(k) \leq rk^d}$$

یادگیری ماشین   فاطمه نادری

ادامه ۶-۱۱

داریم

$\tau_{H}(k) = 2^k$

$\Rightarrow 2^k \le r k^d \Rightarrow k < d\log_2 k + \log_2 r$

طبق لم A2

$a \ge 1 \quad b > 0 \quad \text{if } x \ge 4a\log(2a) + 2b \Rightarrow x \ge a\log(x) + b$

در نتیجه نقیض آن

$x < a\log(x) + b \Rightarrow x < 4a\log(2a) + 2b$

پس: $a = d$، $b = \log_2 r$

$k < d\log k + \log r \Rightarrow \boxed{k < 4d\log_2(2d) + 2\log_2(r)}$

حال: $|C| = k$ ، C توسط $\cup H_i$ خرد می‌شود

$\boxed{VCDim(\cup H_i) \le |C| = k}$

$\Rightarrow VCDim(\cup H_i) \le k < 4d\log_2(2d) + 2\log_2(r)$

فاطمه داوری

یادگیری ماشین

ادامه ۱۱-۲

۲)

۲=۲

$$VCDim(H_1 \cup H_2) \overset{?}{\leq} 2d+1$$

نشان می‌دهیم $H_1 \cup H_2$ نمی‌تواند $2d+1$ عنصری را خورد کند

یعنی $\tau_{H_1 \cup H_2}(k) \neq 2^k$ $\quad k \geq 2d+2$

$$\tau_{H_1 \cup H_2}(k) \leq \tau_{H_1}(k) + \tau_{H_2}(k)$$

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

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

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

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

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

$$= \sum_{i=0}^{k} \binom{k}{i} = 2^k$$

$$\Rightarrow \tau_{H_1 \cup H_2}(k) < 2^k$$

$\Rightarrow$ $H_1 \cup H_2$ نمی‌تواند $2d+2$ عنصری را خورد کند

بنابراین $VCD \leq 2d+1$