٩٠٢

(1)

می خواهیم ثابت کنیم که $VCdim(H) = \min\{k, |H|-k\}$ است.

مرحله اول ثابت می کنیم برای $VCdim(H) \leq \min\{k, |H|-k\}$

$$\begin{cases} (1)\ \underline{k} & \text{if} \\ (2)\ \underline{|X|-k} & \text{if} \end{cases}$$

(۱) اگر مینیمم برابر k شود می توان یک مجموعه ی $C = \{c_1, \dots, c_{k+1}\}$ که H نتواند آن را shatter کند یعنی h وجود ندارد که ۱ را برای c_i ها قرار دهد زیرا تعداد c_i ها $k+1$ است ، پس برای هر C ، k عضوی ، $h \in H$ وجود دارد که H ، C را shatter می کند و برای یک $k+1$ عضوی shatter نمی شود.

② اگر ببینیم برای $|X|-k$ شود می‌توان یک مجموعه‌ای $\{c_1, \dots, c_{|X|-k+1}\}$ C عضوی

پیدا کرد که هیچ تابع $h \in H$ نمی‌توان پیدا کرد که C را shatter کند

یعنی هیچ h وجود ندارد که بتواند به $|X|-k+1$ تا لیبل صفر را بدهد زیرا

لیبل صفر می‌بود که این اشتباه است زیرا در این صورت $k-1$ تا لیبل 1 می‌گیرند

پس می‌توان گفت برای هر $|X|-k$ عضوی می‌توان یک تابع $h \in H$، C را shatter کند

و برای هر $|X|-k+1$ این رابطه برقرار نیست.

③ حال فقط حالت $vcdim(H) \geq \min\{k, |X|-k\}$ می‌ماند که $n = \min\{k, |X|-k\}$

آنگاه: $C = \{c_1, \dots, c_n\}$ یک زیرمجموعه از X است که لیبل $(y_1, \dots, y_n)$

می‌گیرد. آنگاه وجود دارد $h \in H$ که $h(c_i) = y_i$ که حداکثر k تا یا بیشتر وجود

دارد که لیبل 1 می‌گیرند. پس $vcdim(H) \geq n$ است پس

$vcdim(H) = \min\{k, |X|-k\}$

$Vcdim(H) \leqslant K$ ①

② اثبات می کنیم $Vcdim(H)=K$ است

$Vcdim(H) \geqslant K$ ②

① با توجه به شرط vcdim می توان گفت که $C=\{c_1,\dots,c_{k+1}\} \subseteq X$ که نمی توان $h \in H$ پیدا کرد که C را shatter کند یعنی h وجود ندارد که به تمام پارامترهای C مقدار 1 را بدهد زیرا بیشتر از K پارامتر در C وجود دارد.

② می توان گفت $C=\{c_1,\dots,c_k\} \subset X$ که می توان $h \in H$ پیشنهاد داد که $h(c_i)=1$ و $h(x)=0$ (در نقاطی غیر از C_i ها). با توجه به این در شرط vcdim می توان گفت که

$Vcdim(H)=K$

6.6

1)

الف) مشاهده می‌شود که برای هر $h \in H$ شامل دقیقاً d متغیر شود.
که می‌توانند x_i یا $\bar{x}_i$ یا داده‌ی گمشده باشد، پس 3^d حالت ممکن
تا اینجا وجود دارد و این مقدار $h(x)=0$ را تولید نمی‌کند پس به صورت
دستی اضافه می‌کنیم که: $|H^d_{con}| \leq 3^d+1$

← سایز مجموعه فرضیات

ب) حال که سایز H را می‌دانیم می‌توان یک کران بالا برای $vcdim(H)$
تعریف کرد. فرض کنید $vcdim(H)=L$ بنابراین H shatter می‌کند یک
مجموعه‌ی L عضوی از X را، که نیاز دارد H شامل حداقل 2^L تا تابع شود.

$\left\{ \begin{array}{l} True \\ False \end{array} \right.$

$$|H| \geq 2^{L} \xRightarrow[2^{L}=x]{\log_2 x = 1} \log_2 |H| \geq L$$

$\mathrm{Vcdim}(H) = L$

$$\Longrightarrow \log(|H|) \geq \mathrm{Vcdim}(H)$$

$$\Rightarrow \mathrm{Vcdim}(H^{d}_{con}) \leq \log(3^{d}+1) \leq d\log(3)$$

ج) اگر $e=\{e_j\}_{j=1}^{d}$ و $J \subseteq [d]$. اگر $J=[d]$ مجموعه نقاط ما

باشد تمام لیبل‌های مثبت. اگر $J=[\emptyset]$ باشد، تمام مجموعه نقاط با لیبل منفی

$h(e_j) = \begin{cases} 1 & J \in \\ 0 & \text{د.ص} \end{cases}$ پس این تابع می‌تواند به درستی تمام نقاط را لیبل بدهد

پس H برای C را shater می‌کند.

د) می‌خواهیم نشان دهیم که $VCdim(H) \le d$ است. فرض کنید $VCdim(H) > d$ باشد پس H باید یک مجموعه‌ای مانند C را با $d+1$ نقطه shatter کند. $C=\{c_1, \dots, c_{d+1}\}$ که $|H_L| = 2^{d+1}$ است. با استفاده از راهنمایی $h_1, \dots, h_{d+1}$ و $L_1, \dots, L_{d+1}$ را تعریف کرده‌ایم. با توجه به اصل لانه کبوتر، حداقل یک متغیر وجود دارد که دوبار اتفاق بیفتد.

با همین فرض می‌گوییم L_1 و L_2 با یک متغیر در رابطه باشند. اول فرض کنید که $L_1 = L_2$ باشد پس اگر اینجا برای c_1 لیبل T بدهد L_2 نیز به همین صورت عمل کند، که با فرض ما در تناقض است. حال $L_1 \neq L_2$ فرض کنید. در این حالت $h_1(c_2)$ false و $L_2(c_2)$، True است. که باز هم تناقض بوجود می‌آید.

$$L_i = L_j \Rightarrow h_i(c_1) = 0 \wedge h_j(c_1) = 0$$

$$L_i \neq L_j \Rightarrow h_i(c_k) = \neg h_j(c_k)$$

هر دو مورد فرض ما را رد می‌کند پس فرض باطل و حکم $VCdim(H) \le d$ است ثابت می‌شود.

ه) همانگونه که در صورت سوال آمده، در H^d_{con} یک حرف به خالی به عنوان فرض تمام مثبت تفسیر می‌شود. می‌توان با تقویت H^d_{mcon} با مجموعه فرض تمام منفی $\bar{h}$ ادامه داد که:

$$|H^d_{mcon}| \leq 2^d + 1$$

پس:

$$VCdim(H) \leq \log(2^d+1) \leq d$$

حال می‌خواهیم یک مجموعه‌ی C تعریف کنیم که H^d_{mcon} می‌تواند آن را shatter کند.

پس $C=\{c_1, \dots, c_d\}$ و $c_i = \{(1,\dots,1)-e_i\}_{i\leq d}$. اگر A را یک مجموعه شاخص‌های اعضای C که لیبل منفی می‌گیرند باشد، آنگاه:

$$\forall \alpha \in A: y_\alpha = 0$$

کافی است بسازیم: $h(x) = x_{i_1} \wedge x_{i_2} \wedge \dots \wedge x_{i_m}$ ، $m \leq d$. اگر $A = \emptyset$ باشد آنگاه $h(x) = 1$

با این تعریف نتیجه می‌گیریم:

$$\boxed{h(c_i) = 0 \iff i \in A}$$

پس $VCdim(H^d_{mcon}) = d$

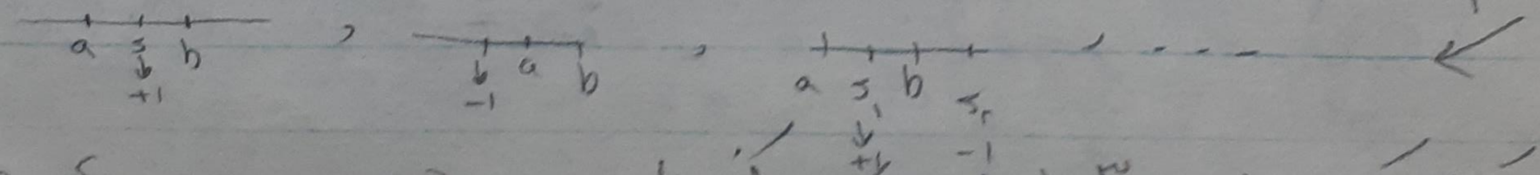
۹.۹

۱) که H shatter کند C را که $|C|=d$; $\exists C$

شرط $Vc\ dim(H)$

۲) H را shatter نمی‌کند . $\forall D, |D|=d+1$

$C=\{s_1,s_2,s_3\}$



می‌توان نتیجه گرفت که برای ① ، $Vcdim \geq ۳$ کار می‌کند حال ، $C=\{s_1,s_2,s_3,s_4\}$

که لیبل $\{+1,-1,+1,-1\}$ نمی‌تواند که توسط هیچ تابع $h\in H$ بدست آید پس

$Vcdim(H) \leq ۳$ ② $\Rightarrow$ ①,② $\boxed{Vcdim(H)=۳}$

۹۰/۰

الف) اگر یک مجموعه $C \subseteq X$ وجود داشته باشد با سایز d. می‌توان یک H_C

که شامل تمام توابع که از $C \to \{0,1\}$ را شامل شود. با توجه به تمرین ۵.۳

روی $C \times \{0,1\}$ یک توزیع وجود دارد:

$\to$ True label

1) $\exists f \in H_C : L_D(f) = 0$

2) $E_{S \sim C^m}[L_D(A(S))] \geq 1/2 - 1/2K$ ؛ $K \leq \frac{|C|}{m} = \frac{d}{m}$

S: Training set

حال اگر دامنه D را به X گسترش دهیم یعنی وارد کردن احتمال ناحیه‌هایی

که در مجموعه C نیستند با احتمال ۰ و انتخاب $\boxed{K = \frac{d}{m}}$

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

پس: $E_{S \sim D^m}[L_D(A(S))] \geq 0 + \frac{d-m}{2d} = \min_{h' \in H} L_D(h') + \frac{d-m}{2d}$

ب) خلاف فرض می کنیم پس $Vc\dim(H)=\infty$ که با استفاده از قضیه ی "no free launch theory". می گوییم برای هر Training set با سایز m وجود دارد یک مجموعه ی shatter شده با سایز $d=2m$. با این نتیجه می توان گفت:

$$E_{S\sim D^m}[A(S)] \geq 0 + \frac{2m-m}{4m} = 1/4$$

با نامساوی مارکف:

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

بنابراین با انتخاب $\varepsilon \leq \frac{1}{8}$ و $\delta \leq \frac{1}{7}$ که نقض می‌کند

Pac learning را پس برای $\infty >$ Vcdim(H) می‌توان گفت که Pac learnible است

9.11

الف) $(H_1, \ldots, H_r)$

اگر $\text{Vcdim}(H_i) = d \geq 3$ یک $k \leq d$ انتخاب می‌کنیم که یک مجموعه‌ی

$C = \{c_1, \ldots, c_k\}$ که تمام توابع $h \in H_i$ آن را shatter می‌کند.

یک تابع رشد که در Hint آمده به صورت: $\tau_{H_i}(k) = 2^k$

1| مرحله اول با توجه به union bound برای تابع رشد از H_i ها داریم:

$$\tau_{\cup H_i}(k) = \max_{\substack{C \subseteq X \\ |C| = k}} |H_C| = \max \left| \{ h(c_1), \dots, h(c_k) : h \in \bigcup_{i=1}^{r} H_i \} \right|$$

$$\Rightarrow \max \left| \bigcup_{i=1}^{r} H_i^C \right| \leq \max \sum_{i=1}^{r} |H_i^C|$$

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

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

2| با استفاده از "Sauer's lemma" برای پیدا کردن یک bound برای تابع رشد:

$$\tau_{H_i}(k) < \left(\frac{ek}{d}\right)^d \leq \left(\frac{ek}{3}\right)^d \leq \boxed{k^d} \quad (2)$$

حال با توجه به (1), (2) داریم:

$$\tau_{\cup H_i}(k) \leq \sum_{i=1}^{r} \tau_{H_i}(k) < \sum_{i=1}^{r} k^d = \boxed{r k^d}$$

حال باتوجه به تابع رشد که $\tau_{\cup H_i}(k)=2^k$ بود:

$$2^k < r k^d$$

$$\Rightarrow k < \log_2 r + d \log_2 k$$

مقدار ۲ باتوجه به A.2 lemma

$$\Longrightarrow k < \log_2 r + d\log_2 k \Rightarrow k < 2\log r + 4d\log 2d$$

می دانیم k تا پارامترها توسط $\cup H_i$ ها shatter می‌شوند پس:

$$vc\dim\left(\bigcup_{i=1}^{r} H_i\right) \leqslant k < 4d\log(2d) + 2\log(r)$$