Date .... Subject ....

تمرین های فصل ۶

2) a

Let $C \subseteq X$, $|C| = k+1$

$\nexists h \in H$ st $\forall x \in C$ $h(x) = 1$

Let $C \subseteq X$ s.t $|C| = |x| - k + 1$

$\Rightarrow \nexists h \in H$ s.t $\forall x \in C$ $h(x) = 0$

So: $VCdim(H) \leq \min\{k, |x| - k\}$

Let $C \subseteq X$ s.t $|C| = m \leq \min\{k, |x| - k\}$

Let $(y_1, \dots, y_m) \in \{0,1\}^m$ & $S = \sum_{i=1}^{m} y_i$

$E \subseteq X - C$

$\forall x_i \in C$, $h(x) = I_E \Rightarrow C$ is shattered by H

$VCdim(H) \geq \min\{k, |x| - k\} \Rightarrow VCdim(H) = \min\{k, |x| - k\}$

Scanned with CamScanner

Let $C \subseteq X$ s.t $|C| = k+1$ b)

$\Rightarrow \nexists h \in H$ s.t $\forall x \in C$ $h(x) = 1 \Rightarrow VCdim(H) \leq k$

Let $C = \{x_1, \dots x_m\} \subseteq X$ s.t $|C| = m \leq k$

از طرفی دیگر داریم:

$(y_1, \dots, y_m) \in \{0,1\}^m$

$h \in H$ s.t $\forall x_i \in C$ $h(x_i) = y_i$ & $h(x) = 0$

$\Rightarrow VCdim(H) \geq k \Rightarrow VCdim(H) = k$

4) $X = R^d$

Let $d \geq 2$ & $H = \{ I_{\|x\|_2 \leq r} : r \geq 0 \}$

اگر $x \neq (0, 0, \dots, 0)$ باشد آنگاه، $\{x\}$ shatter می‌شود

اگر $\|x_1\|_2 \leq \|x_2\|_2 \Rightarrow y_1 = 0$ & $y_2 = 1$

آنگاه، به‌وسیله هیچ فرضی در H به‌دست نمی‌آید

Date .... Subject ....

Let $A = \{e_1, e_2\} \Rightarrow H_A = \{(0,1),(1,1)\}$

$\Rightarrow \{B \subseteq A : H \text{ shatters } B\} = \{\phi, \{e_1\}, \{e_2\}\}$

$\Rightarrow \sum_{i=0}^{d} \binom{|A|}{i} = 2$

Let H be the class of axis-aligned rectangles in R^2
we have seen that the vc dimention of H is 4
Let $A = \{x_1, x_2, x_3\}$ when $x_1 = (0,0)$, $x_2 = (1,0)$, $x_3 = (1,1)$
all the labels except $(1,0,1)$ are obtained
$\Rightarrow |H_A| = 7$, $|\{B \subseteq A : H \text{ shatters } B\}| = 7$
and $\sum_{i=0}^{d} \binom{|A|}{i} = 8$

let $d \geq 2$ and consider the class H

$H = \{\operatorname{sign}\langle w, x \rangle : w \in R^d\}$ of Homogeneous halfspaces

we will prove in theorem in chap 9 vc dim of this class

vc dim$(H) \geq 2$

Date .... Subject ....

$$\begin{cases} X_1 = e_1 \\ X_2 = e_2 \\ X_3 = (1,1,0,\dots,0) \end{cases}$$

پس $\{e_1, e_2, e_3\}$ shatter نمی‌شود چون

چون label ها بنیز $(1,1,-1)$, $(-1,-1,1)$ ممکن نیستند

$|H_A| = 4$, $|\{B \subseteq A : H \text{ shatters } B\}| = 2$, $\sum_{i=0}^{d} \binom{|A|}{i} = 2$

let $d = 1$ and consider $H = \{ I_{\{x \geq t\}} , t \in \mathbb{R} \}$

every singleton is shattered by H and every set of size

at least 2 is not shattered by H. choose any finite set $A \subseteq \mathbb{R}$

and each of three terms in sauer's inequality, equal $|A|+1$.

9) $H = \{ h_{a,b,s} : a < b , s \in \{-1, 1\} \}$ where

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

let $c = \{1, 2, 3\}$

1	2	3	a	b	s
−	−	−	۰,۵	۴,۵	−1
−	−	+	۲,۵	۴,۵	1
−	+	−	۲,۵	۲,۵	1
−	+	+	۰,۵ / ۱,۵	۴,۵	1
+	−	−	۱,۵	۱,۵	1
+	−	+			1
+	+	−	۰,۵	۲,۵	1
+	+	+	۰,۵	۲,۵	−1
				۴,۵	1

Date .... Subject ....

$\Rightarrow Vcdim(H) \geq 3$

let $C = \{x_1, x_2, x_3, x_4\}$

$y_1 = y_3 = -1 \quad , \quad y_2 = y_4 = 1$

که این نتوانه هیچ فرضی تولید نمی شود پس $Vcdim(H) < 4$

یعنی هیچ وقت ما ثابت کردیم برای ۳ عضوی H وجود دارد ولی برای

۴ عضوی مثالی زدیم که این شرایط برقرار نبود پس

$Vcdim(H) = 3$

10) a) فرض کنید $m < d$ و C shatter بشود به اندازه d

ما فرض می‌کنیم $X = C$ چون همیشه نزدیکترین انتخاب می‌کنیم که شامل C باشد

پس $\min_{h \in H} L_D(h) = 0$, but $E[L_D(A(S))] \geq \frac{k-1}{2k}$

$k = \frac{d}{m}$

$\Rightarrow E(L_D(A(S))) = \frac{d-m}{2d}$

Date .... Subject ....

H PAC learnable $\Rightarrow$ $VCdim(H) < \infty$ (b

assume $VCdim(H) = \infty$. Let A be a learning algorithm

we show that A fails to (PAC) learn H.

choose $\varepsilon = \frac{1}{16}$, $\delta = \frac{1}{14}$

there exists a shattered set of size $d = 4m$

(according to No free lunch theorim)

exist D $\min_{h \in H} L_D(h) = 0$ but $E(L_D(A(S))) \geq \frac{1}{\varepsilon}$

so with probability at lest $\delta < \frac{1}{7}$

$L_D(A(S)) - \min_{h \in H} L_D(h) = L_D(A(S)) \geq \frac{1}{\wedge} > \varepsilon$

$L_D(A(S)) > \varepsilon$

~~S_D is~~ PAC

Date Subject

11) we may assume VC dim $(H_i) = d \geq 3$

$H = \bigcup_{i=1}^{r} H_i$. $k \in [d]$, $\tau_H(k) = 2^k$

$k < 4d\log(2d) + 2\log r$. (purpose)

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

$d \geq 3$ we apply Sauer's lemma:

$\tau_H(k) \leq r m^d$

$\Rightarrow k < d\log m + \log r$

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

|

Date .... Subject ....

b) $VCdim(H_1) = VCdim(H_2) = d$

Let $H = H_1 \cup H_2$. Let k be a positive integer such that

$k \geq 2d+2$ we show that $\tau_H(k) < 2^k$, By Sauer's lemma

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

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

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

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

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

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

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

Scanned with CamScanner

Date .... Subject ....

9

1) Define a vector of auxiliary variables $s = (s_1, \dots, s_m)$

minimizing the empirical risk is equivalent to minimizing

: linear objective $\sum_{i=1}^{m} s_i$

$(\forall i \in [m])$ $w^T x_i - s_i \leq y_i$, $-w^T x_i - s_i \leq -y_i$

It is left to translate the above into matrix form.

Let $A \in \mathbb{R}^{2m \times (m+d)}$ be matrix $A = [X \; -I_m \; ; \; -X \; -I_m]$

Let $v \in \mathbb{R}^{d+m}$ be the vector of variables $(w_1, \dots, w_d, s_1, \dots, s_m)$

define $b \in \mathbb{R}^{2m}$ to be the vector $b = (y_1, \dots, y_m, -y_1, \dots, -y_m)^T$

$c \in \mathbb{R}^{d+m}$ be the vector $c = (0_d \, ; \, 1_m)$

It follows that the optimization problem of minimization

ER : $\min \; c^T v$ s.t $Av \leq b$

Date .... Subject ....

3) let $d = m$, $x_i = e_i$ $\text{sign}(0) = -1$

for $i = 1, 2, \dots, d$, $y_i = 1$ be label of x_i

$w^{(t)}$ weight vector which is maintained by perceptron

$w_i = \sum_{j<i} e_j$, $\langle w^{(i)}, x_i \rangle = 0$

also note vector $w^* = (1, \dots, 1)$ satisfies the requirments

listed in the question.

6) In this question we will denote the class of

halfspace in $\mathbb{R}^{d+1}$ by L_{d+1}

(a) Assume that $A = \{x_1, \dots, x_m\} \subseteq \mathbb{R}^d$ is shattered by B_d

Then $\forall y = (y_1, \dots, y_m) \in \{-1, 1\}^d$ there exist $B_{\mu,r} \in B$ s.t

for every i $B_{\mu,r}(x_i) = y_i$

Scanned with CamScanner

Hence, for the above μ and r the following identity holds for every $i \in [m]$:

$$\operatorname{sign}\left((2\mu; -1)^T (x_i; \|x_i\|^2) - \|\mu\|^2 + r^2 \right) = y_i$$

where ; denotes vector concatenation.