Denis Kim

Homework 2

Exercise 2.2 (b)

- There is no such hypothesis set.

According to theorem 2.4 if $m_H(N) < 2^N$ then there exist k for which $m_H(N) < 2^k$, so $m_H(N) \le \sum_{i=0}^{k-1} \binom{N}{i}$ for all N

Exercise 2.6

(a) error bar $= \sqrt{\frac{1}{2N} \ln \frac{2M}{\delta}}$

error bar $(E_{in}) = \sqrt{\frac{1}{2\cdot 400} \cdot \frac{2\cdot 1000}{0.05}} = 0,115$

error bar $(E_{out}) = \sqrt{\frac{1}{2\cdot 200} \cdot \frac{2\cdot 1}{0,05}} = 0,096.$

- $E_{in}(g)$ estimate has the higher error bar

(b) One of the reason why we shouldn't reserve more examples for testing is that we will have fewer examples for trainging. That will lead to decrease a chance of choosing a good hypothesis

Problem 2.3 (b)

+1 | a | −1 | b | +1   or   −1 | a | +1 | b | −1

- if we assume that a can be equal to b and we can have either no −1 or no +1

N	dichotomies $H(x_1, \dots, x_N)$	$m_H(N)$
1	{o, x}	2
2	{oo, xx, xo, ox}	4
3	{ooo, xxx, xox, oxo, oox, xoo, xxo, oxx}	8
4 break point	{oooo, xxxx, oxxo ...} except {0101, 1010}	$14 < 2^N$

$m_H(N) \le 2^N$

$m_H(N) = 2^N$

$d_{vc}(H)$ = maximum non-break point = 3

It can be seen from table above that 4 is a break point for H, while 3 is not.

Problem 2.16 {

(a)

$$H = \left\{ h_c \;\middle|\; h_c(x) = \text{sign}\left(\sum_{i=0}^{D} c_i x^i\right) \right\}$$

d_{vc} is the largest value of N for which $m_H(N) = 2^N$

Let's consider matrix $X^{d+1,d+1}$, vector c^{D+1} and vector y^{D+1}

$$\begin{bmatrix} x_0^0 & \cdots & x_0^D \\ x_1^0 & \cdots & x_1^D \\ \vdots & \cdots & \vdots \\ x_D^0 & \cdots & x_D^D \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ \vdots \\ c_D \end{bmatrix} = \begin{bmatrix} y_0 \\ \vdots \\ y_D \end{bmatrix}$$, where y vector $\in \{-1, +1\}^{D+1}$

$h_c(x) = \text{sign}\left(\sum_{i=0}^{D} c_i x_k^i\right) = y_k$, $k = 0, \dots, D$

therefore $(D+1)$ points can be shattered by H

(b) Any $D+2$ vectors of length $D+1$ will be linearly dependent therefore $D+2$ points cannot be shattered by H

Problem 2.18

$$H = \left\{ h_\alpha \;|\; h_\alpha(x) = (-1)^{\lfloor \alpha x \rfloor}, \text{ where } \alpha \in R \right\}$$

d_{vc} is infinite if $m_H(N) = 2^N$ for all N

Consider $x_1, \dots, x_N$ where $x_n = 10^N$

$y = (y_1, \dots, y_N) \in \{-1, +1\}^N$

Lets $\alpha = \frac{1}{10^k}$ if $y = -1$

$\alpha = \frac{2}{10^k}$ if $y = +1$

$\Rightarrow h_\alpha(x_k) = (-1)^{\lfloor \alpha 10^k \rfloor} = y_k$ for $k = 1, \dots, N$

$H(x_1, \dots, x_N) = \{-1, +1\}^N \Rightarrow m_H(N) = 2^N$

therefore $d_{vc} = \infty$

Exercise 2.4 (a)

(a) $X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^d \\ \vdots & \vdots & \cdots & \vdots \\ 1 & (d+1)^1 & \cdots & (d+1)^d \end{bmatrix}$ $\quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{d+1} \end{bmatrix}$

$X \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{d+1} \end{bmatrix} = y$ → therefore $d+1$ points can be shattered

X is invertible matrix with dimensions $(d+1) \times (d+1)$

Exercise 2.4 (b)

(b) Any $(d+2)$ vector in $(d+1)$ dimension is linearly dependent, that means that we have a linear combination of z_i

$$z_{d+2} = \sum_{i=1}^{d+1} k_i z_i$$

$$w^T z_{d+2} = \sum_{i=1}^{d+1} w^T k_i z_i$$

$$\text{sign}(w^T z_{d+2}) = \text{sign}\left(\sum_{i=1}^{d+1} k_i w^T z_i\right)$$

$(\text{sign}(z_1), \text{sign}(z_2), \dots \text{sign}(z_{d+1}), -\text{sign}(z_{d+2}))$ cannot be shattered

therefore $d_{vc} \leq d+1$

Exercise 2.7 (a)

if $g(x) = |h(x) - f(x)|$, then $g(x) = 0$ or $g(x) = 1$

Because

$h(x)$	0	0	1	1
$f(x)$	0	1	0	1
$g(x)$	0	1	1	0

$P[h(x) \neq f(x) = P(|h(x) - f(x)| = 1)$

then

$$E(g(x)) = \sum_{n=1}^{N} g(x) P(X = x_n) =$$

$$= \sum_{n=1}^{N} g(x) \cdot \frac{1}{N} = \frac{1}{N} \sum_{n=1}^{N} (0|1) =$$

$$= P(h(x) \neq f(x))$$

Exercise 2.7 (b)

$h(x)$	1	-1	1	-1
$f(x)$	1	1	-1	-1
$g(x)$	0	2	2	0

$$E(g(x)) = \sum^{N} g(x) P(X = x_n) =$$

$$= \frac{1}{N} \sum^{N} (2|0) =$$

$$P[h(x) \neq f(x)] = \frac{1}{4} E(g(x))$$

Problem 2.8

$$m_H(N) \leq \sum_{i=0}^{d_{vc}} \binom{N}{i} \qquad (2.9)$$

- $1+N \rightarrow$ possible ✓

if $d_{vc}=1$, then $\sum_{i=0}^{1} \binom{N}{i} = \binom{N}{0} + \binom{N}{1} = 1 + \frac{N!}{(N-1)!} = 1+N$

- $1+N+\frac{N(N-1)}{2} \rightarrow$ possible ✓

if $d_{vc}=2$, then $\sum_{i=0}^{2} \binom{N}{i} = \binom{N}{0} + \binom{N}{1} + \binom{N}{2} = 1+N+\frac{N!}{(N-2)!\,2!} = 1+N+\frac{N(N-1)}{2}$

- $2^N \rightarrow$ possible ✓, because it satisfies $m_H(N) \leq 2^N$, and $d_{vc} = \infty$, then $m_H(N) = 2^N$

- $2^{\lfloor\sqrt{N}\rfloor} \rightarrow$ NOT possible ✗

if $m_H(N) = 2^{\lfloor\sqrt{N}\rfloor}$ ~~[illegible]~~

$d_{vc} = 1$, because $2^{\sqrt{1}} = 2^1$ and $2^{\sqrt{2}} \neq 2^2$

$$m_H(N) \leq N^{d_{vc}} + 1 \qquad (2.10)$$

Lets consider $N = 25$

$m_H(N) = 2^{\lfloor\sqrt{N}\rfloor} = 2^5 = 32$

$N^{d_{vc}} + 1 = 25^1 + 1 = 26$

$32 \not\leq 26$, therefore it's NOT possible

- $2^{\lfloor N/2 \rfloor} \rightarrow$ NOT possible ✗

$d_{vc} = 0$, because $2^{\lfloor 0/2 \rfloor} = 2^{\lfloor 0 \rfloor}$ and $2^{\lfloor 1/2 \rfloor} \neq 2^{\lfloor 1 \rfloor}$

$m_H(N) \leq N^{d_{vc}} + 1$

Lets consider $N=4$

$m_H(N) = 2^{\lfloor N/2 \rfloor} = 2^2 = 4$

$N^{d_{vc}} + 1 = 1+1 = 2$

$4 \not\leq 2$, therefore it is NOT possible

- $m_H(N) = 1+N+\frac{N(N-1)(N-2)}{6} \rightarrow$ possible ✓

if $d_{vc} = 3$, then $\sum_{i=0}^{3} \binom{N}{i} = 1+N+\frac{N(N-1)(N-2)}{6}$