Bauyrzhan Zhakanov
Machine learning
Homework II

0.1 VC - Dimension

Q 1. Exercise 2.2 (b)

No, since $m_H(N) = N + 2^{N/2}$ is an exponential function and bounded by a polynomial.

Q 2. Exercise 2.6

dataset = 600 examples
selected - subset = 200 exeimples form test $\quad d_{vc} = 1$
learning - model = 1000 hypoth.
select 400 based on training ex.

a) $\delta = 0.05$

$$E_{out_1}(g) \le E_{in}(g) + \sqrt{\frac{8}{400} \ln\left(\frac{4(801)}{0.05}\right)}$$

$$\le E_{in}(g) + 0,47$$

$$E_{out_2}(g) \le E_{test}(g) + \sqrt{\frac{8}{200} \ln\left(\frac{4(401)}{0.05}\right)}$$

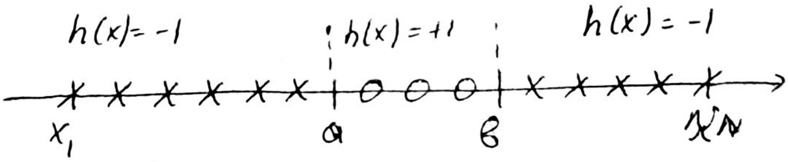
$$\le E_{test}(g) + 0.64$$

So, we could see that $E_{in} > E_{test}$ at any d_{vc} value.

b) The test set has straight finite-sample variance, but no bias. More examples in testing won't make any changes to E_{out}

Q3. (Exercise 2.3(b)) Problem 2.3 (b)

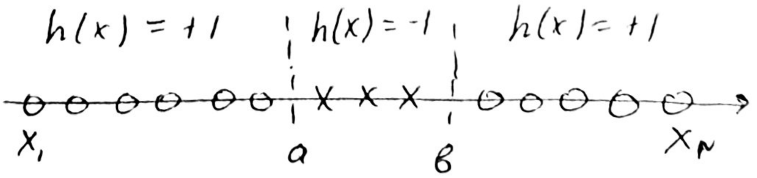
1) Positive intervals



dichotomies: $\binom{N+1}{2}$

$m_H(N) = \frac{N^2}{2} + \frac{N}{2} + 1 \qquad d_{vc} = 2.$

2) Negative Intervals



dichotomies: $\binom{N+1}{2}$

$m_H(N) = \frac{N^2}{2} + \frac{N}{2} + 1 \qquad d_{vc} = 2.$

Bauyrzhan Zhakanov

Q4. Problem 2.16

VC Dimension is $D+1$

(a) $D+1$ can be shattered by H

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

According to Sauer's Lemma $B(N,k) \leq \sum_{i=0}^{k-1} \binom{N}{i}$

statement is true for $N \leq N_0$

Given $D+1$ is a break point, the bound $\sum_{i=0}^{D+1} \binom{N}{i}$

(b) A set of dichotomies won't shatter k out N points, otherwise it would be unbounded.
Since $k = D+1$, $D+2$ cannot be shattered by H

Q5. Problem 2.18

$$H = \{h_\alpha \mid h_\alpha(x) = (-1)^{[\alpha x]}, \text{ where } \alpha \in \mathbb{R}\}$$

Consider $x_1 \dots x_n$, where $x_n = 10^n$, let $\alpha = 1$

$y_n = (-1)^{\alpha x_n}$ $\quad y_1 = (-1)^{10}$, $x_1 = 10^1$

$y_2 = (-1)^{100}$, $x_2 = 10^2 = 100$

and let $\alpha = 2$.

$y_1 = (-1)^{20}$, $x_1 = 10^1$

$y_2 = (-1)^{200}$, $x_2 = 10^2 = 100$

so H is exponential, no generalization bound is shown.

if $m_H(N) = 2^N$, then $d_{vc}(H) = \infty$.

if $\alpha = \infty$, then $d_{vc}(H) = \infty$.

0.2 Perceptron vs VC Dimension

Bauyrzhan Zhakanov

Q1. Ex 2.4(a)

$d_{VC} = d+1$

$$X = \begin{bmatrix} \cdots x_1^T \cdots \\ \vdots \\ \cdots x_N^T \cdots \end{bmatrix}$$ matrix of $(d+1)\times(d+1)$

$X = [X_0 \ \ldots \ X_d]$ in R^d

$X_0 = [1 \ 1 \ \ldots \ 1]$.

for any $y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \pm 1 \\ \vdots \\ \pm 1 \end{bmatrix}$, It is possible to show that perceptron may have dichotomies, so points can be shattered

~~$y = Xw$~~

~~$w = X^{-1}y$~~

$y = \text{sign}(Xw)$

$y = Xw \Rightarrow w = X^{-1}y$, so it shows any dichotomy $\begin{bmatrix} +1 \\ -1 \\ \vdots \end{bmatrix}$, so we can shatter $d+1$ or less points

Q2. Ex 2.4 (b)

for $d+2$, $x_1 \ldots x_{d+2}$

There more vectors than dimensions, so vectors are linearly dependent, $x_j = \sum_{i \neq j} a_i x_i$ x_1 are ones.

$y_i = \text{sign}(w^T x_i) = \text{sign}(a_i)$.

For this $\sum_{i \neq j} a_i w^T x_i > 0$.

We can't obtain negative, that dichotomy is not possible to get $y_j = \text{sign}(w^T x_j) = +1$

Also, according to the example $B(5,2)$. 6 dichotomies on 5 points. There is no misclassified $-1, -1$

2 points can not be shattered

$$\left\{ \begin{matrix} [1 & 1 & 1 & 1 & 1] \\ [-1 & 1 & 1 & 1 & 1] \\ [1 & -1 & 1 & 1 & 1] \\ [1 & 1 & -1 & 1 & 1] \\ [1 & 1 & 1 & -1 & 1] \\ [1 & 1 & 1 & 1 & -1] \end{matrix} \right.$$

0.3 Upper Bound

Ex 2.7 Baurzhan Zhakanov

(a) Binary function for 0 and 1

$\mathbb{P}(h(x)\neq f(x)) = \mathbb{E}[(h(x)-f(x))^2]$

Mean squared value $E_{in} = \frac{1}{N}\sum_{n=1}^{N}(h(x)-f(x))^2$

in 0 and 1.

$E_{in} = \frac{(0-1)^2+(1-1)^2+(1-0)^2}{3} = \frac{1+1}{3} = \frac{2}{3}$

the values are different than expected value $(1-0)=1$ however considering large scale numbers in mean-squared must be close to expected value

(B) Binary function for -1 and 1

$E_{in} = \frac{(-1-1)^2+(1+1)^2+(1+1)^2}{3} = \frac{4+4}{3} = \frac{8}{3}$

expected values has to be 4 if value contradicts with

Bauyrzhan Zhakanov

Q3. 2.8 (Exercise) Problem

(1) $1+N \le 2^N$. growth func.

(2) $1+N+\frac{N(N-1)}{2} = \sum_{i=0}^{2} \binom{N}{i}$ is growth func

(3) 2^N $d_v = \infty$ growth fun

(4) $2^{\lfloor\sqrt{N}\rfloor}$ growth func

(5) $2^{\lfloor N/2 \rfloor}$ bounded by 2^N can be grouth.

(6) $1+N+\frac{N(N-1)(N-2)}{6} = \sum_{i=0}^{3} \binom{N}{i}$ growth func

Bonus Problem

$$B(N,k) = \sum_{i=0}^{k-1} \binom{N}{i}$$

$$B(1,1) = \binom{N}{0} = \frac{N!}{0!\,N!} = 1$$

$$B(N_0+1,k) = \sum_{i=0}^{k-1} \binom{N_0+1}{i}$$

$$= \sum_{i=0}^{k-1} \binom{N_0}{i} + \sum_{i=0}^{k-1} \binom{N_0}{i-1}$$

$$= \sum_{i=0}^{k-1} \binom{N_0}{i} + \sum_{i=0}^{k-2} \binom{N_0}{i}$$

$$= 1 + \sum_{i=1}^{k-1} \binom{N_0}{i} + \sum_{i=1}^{k-1} \binom{N_0}{i-1}$$

//