Homework 3

Runmin Lu

September 19, 2021

Exercise 1.13

(a)

2 cases where h makes an error:

$h = f, f \neq y$

or

$h \neq f, f = y$

$$\begin{aligned} P(h = f, f \neq y) + P(h \neq f, f = y) &= (1 - \mu)(1 - \lambda) + \mu\lambda \\ &= 1 - \mu - \lambda + \mu\lambda + \mu\lambda \\ &= \boxed{1 - \mu - \lambda + 2\mu\lambda} \end{aligned}$$

(b)

For some value of λ, the μ in the expression above will cancel out.

$$\begin{aligned} 1 - \mu - \lambda + 2\mu\lambda &= 1 - \lambda + \mu(2\lambda - 1) \\ 2\lambda - 1 &= 0 \\ \lambda &= \boxed{\frac{1}{2}} \end{aligned}$$

Exercise 2.1

1

$k = 2$ because for 2 points, you cannot have the left being $+1$ and the right being -1.

$m_H(2) = 2 + 1 = 3 < 2^2 = 4$

2

$k = 3$ because you cannot have $+1$'s on two ends and -1's in the middle.

$m_H(3) = \binom{4}{2} + 1 = 6 + 1 = 7 < 2^3 = 8$

3

Break point does not exist.

Exercise 2.2

(a)

(i)

$$\begin{aligned}
\text{LHS: } m_H(N) &= N + 1 \\
\text{RHS: } \sum_{i=0}^{1} \binom{N}{i} &= \binom{N}{0} + \binom{N}{1} \\
&= 1 + N \\
\text{LHS} &\leq \text{RHS}
\end{aligned}$$

(ii)

$$\begin{aligned}
\text{LHS: } m_H(N) &= \frac{1}{2}N^2 + \frac{1}{2}N + 1 \\
\text{RHS: } \sum_{i=0}^{2} \binom{N}{i} &= \binom{N}{0} + \binom{N}{1} + \binom{N}{2} \\
&= 1 + N + \frac{N(N-1)}{2} \\
&= 1 + N + \frac{N^2}{2} - \frac{N}{2} \\
&= 1 + \frac{N}{2} + \frac{N^2}{2} \\
\text{LHS} &\leq \text{RHS}
\end{aligned}$$

(iii)

Break point does not exist.

(b)

No.

Proof.

Assume there exists such a hypothesis set. Then

$$
\begin{aligned}
m_H(1) &= 1 + 2^0 = 2 = 2^1 \\
m_H(2) &= 2 + 2^1 = 4 = 2^2 \\
m_H(3) &= 3 + 2^1 = 5 < 2^3
\end{aligned}
$$

We found a break point $k = 3$.

$$
\begin{aligned}
m_H(N) &= \sum_{i=0}^{2} \binom{N}{i} \\
&= 1 + \frac{N}{2} + \frac{N^2}{2} \\
&\in O(N^2)
\end{aligned}
$$

However, we're given that $m_H(N) = N + 2^{\lfloor N/2 \rfloor} \in \Omega(2^{N/2})$. Contradiction.

Therefore no such hypothesis set exists. $\square$

Exercise 2.3

0.1 (i)

$$
\begin{aligned}
d_{VC} &= k - 1 \\
&= 2 - 1 \\
&= \boxed{1}
\end{aligned}
$$

0.2 (ii)

$$
\begin{aligned}
d_{VC} &= k - 1 \\
&= 3 - 1 \\
&= \boxed{2}
\end{aligned}
$$

(iii)

$$m_H(N) = 2^N$$
$$d_{VC} = \boxed{\infty}$$

Exercise 2.6

(a)

$E_{test}(g)$ has the higher error bar because $\varepsilon \in O(\sqrt{\frac{\ln |H|}{N}})$ where N is on the denominator. In this case the sample size for testing is smaller, which results in bigger ε.

(b)

We want $E_{in}(g) \approx 0$, which means that we want to make the error bar for E_{in} as small as possible. Therefore, we want to reserve more data used in selecting g rather than testing.

Problem 1.11

Let $e(g(x_i), y_i)$ denote the point wise error represented in the matrix.
For the supermarket case:

$$e(g(x_i), y_i) = \begin{cases} 0 & g(x_i) = y_i \\ 1 & g(x_i) = +1, y_i = -1 \\ 10 & g(x_i) = -1, y_i = +1 \end{cases}$$

supermarket For the CIA case:

$$e(g(x_i), y_i) = \begin{cases} 0 & g(x_i) = y_i \\ 1 & g(x_i) = -1, y_i = +1 \\ 1000 & g(x_i) = +1, y_i = -1 \end{cases}$$

In general, for both cases with a sample size of N:

$$E_{in}(g) = \frac{1}{N} \sum_{i=1}^{N} e(g(x_i), y_i)$$

Problem 1.12

(a)

$$
\begin{aligned}
E_{in}(h) &= \sum_{n=1}^{N}(h - y_n)^2 \\
&= \sum_{n=1}^{N}(h^2 - 2hy_n + y_n^2) \\
&= Nh^2 - 2h\sum_{n=1}^{N} y_n + \sum_{n=1}^{N} y_n^2 \\
E'_{in}(h) &= 2Nh - 2\sum_{n=1}^{N} y_n = 0 \\
h &= \frac{1}{N}\sum_{n=1}^{N} y_n
\end{aligned}
$$

(b)

Define a cutoff point M where $\forall n \leq M : h \geq y_n$ and $\forall n > M : h < y_n$

$$
\begin{aligned}
E_{in}(h) &= \sum_{n=1}^{N} |h - y_n| \\
&= \sum_{n=1}^{M}(h - y_n) + \sum_{n=M+1}^{N}(y_n - h) \\
&= Mh - \sum_{n=1}^{M} y_n + \sum_{n=M+1}^{N} y_n - (N - M)h \\
&= (2M - N)h - \sum_{n=1}^{M} y_n + \sum_{n=M+1}^{N} y_n \\
E'_{in}(h) &= 2M - N = 0 \\
M &= \frac{N}{2} \implies h \text{ is median}
\end{aligned}
$$

(c)

h_{mean} will increase because the mean is affected by every data point.

h_{med} won't change because the median is not affected by any outlier.