

Big Data Analytics - HW4

Part 1 :-

Generalization Performance

1.

x	\hat{f}	$P(x, \hat{f})$
0	0	0.4
0	1	0.2
1	0	0.1
1	1	0.3

$$\text{Performance} = (\hat{y} - y)^2 \quad \text{where} \quad y = f(x=x) \quad \begin{cases} 0 & x=0 \\ 1 & x=1 \end{cases}$$

Now, we know:-

$$P(x, \hat{f}) = P(\hat{f} = \hat{f} \mid X=x)$$

So,

x	\hat{f}	$P(x, \hat{f})$	f
0	0	$P(\hat{f}=0 \mid X=0) = 0.4$	0
0	1	$P(\hat{f}=1 \mid X=0) = 0.2$	0
1	0	$P(\hat{f}=0 \mid X=1) = 0.1$	1
1	1	$P(\hat{f}=1 \mid X=1) = 0.3$	1

$$\text{So, } P(\hat{f}(x=0) \mid (x=0)) = \begin{cases} \hat{f}=0 & 0.4/0.6 = 2/3 \\ \hat{f}=1 & 0.2/0.6 = 1/3 \end{cases}$$

$$\therefore E[\hat{f}(x=0) \mid (x=0)] = \sum_{\hat{f}=0}^1 \hat{f} \cdot P(\hat{f} \mid x=x)$$

$$= (0) \times \left(\frac{2}{3}\right) + (1) \times \left(\frac{1}{3}\right) = \frac{1}{3}$$