Quiz 2

Q.1

$\longrightarrow$ Partitioning the data by $Feature_1 \leq 5.5$

Left node - { A, B, G }

Right node - { C, D, E, F, H }

Gini for left node:

.25 $p(X) = \frac{2}{3}$, $p(Z) = \frac{1}{3}$, $p(Y) = 0$

.25 $G_L = 1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 - (0)^2 = \frac{4}{9} \approx 0.444$

Gini for Right node:

.25 $p(X) = 0$, $p(Z) = \frac{2}{5}$, $p(Y) = \frac{3}{5}$

.25 $G_R = 1 - (0)^2 - \left(\frac{2}{5}\right)^2 - \left(\frac{3}{5}\right)^2 = \frac{12}{25} \approx 0.48$

Total Gini index:

weight for left node = 3/8

.5

weight for Right node = 5/8

$G_{Total} = \left(\frac{3}{8}\right) \times \left(\frac{4}{9}\right) + \left(\frac{5}{8}\right) \times \left(\frac{12}{25}\right) = \frac{7}{15} \approx 0.4667$

.5

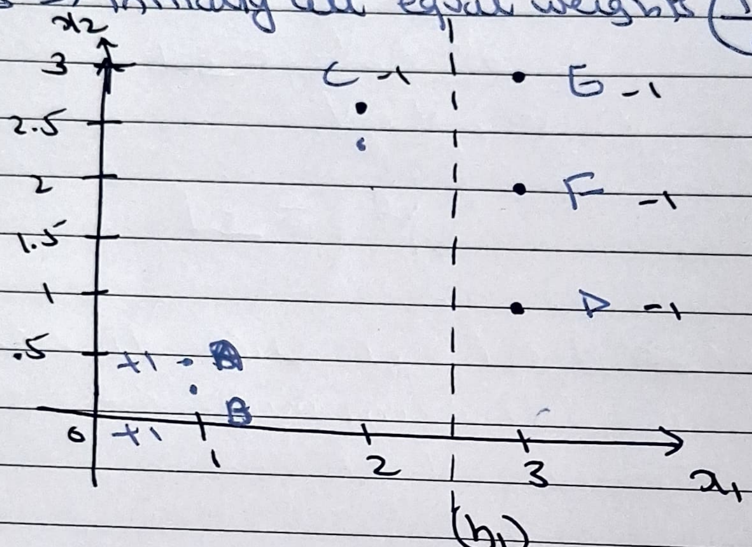
Quiz2 Q2

Given data:

	x	y	after w_0	after w_1	w_2
A	(1, 0.5)	+1	1/6	1/6	1/6
B	(1, .3)	+1	1/6	1/6	1/6
C	(2, 2.8)	−1	1/6	5/6	5/6
D	(3, 1)	−1	1/6	1/6	3/2
E	(3, 3)	−1	1/6	1/6	1/6
F	(3, 2)	−1	1/6	1/6	1/6

(a)

6 points ⇒ initially all equal weights $\left(\frac{1}{6}\right)$



Given: Initial split at $x_1 = 2.5$

$$h_1(x) = y = \begin{cases} +1 & x_1 \le 2.5 \\ -1 & x_1 > 2.5 \end{cases}$$

⇒ points E F D are correctly classified (−1)

⇒ points A B are correctly classified (+1)

point C is misclassified.

we calculate loss L_1 → indicator function

$$L_1 = \frac{\sum_{i=1}^{n} w_i \, I(h_1(x_i) \neq y_i)}{\sum_{i=1}^{n} w_i}$$

.5 $\Rightarrow L_1 = \frac{1/6}{1} = \frac{1}{6}$ [only point C is misclassified]

CLASSTIME Pg. No.
Date / /

next we calculate α_1

$$\alpha_1 = \frac{1}{2} \ln\left(\frac{1-L_1}{L_1}\right)$$

$$= \frac{1}{2} \ln\left(\frac{1-1/6}{1/6}\right)$$

$$\boxed{\alpha_1 = \frac{1}{2} \ln 5}$$ .5

Weight updation for misclassified samples:

$$w_c^{(1)} \leftarrow w_c^{(0)} \times e^{2\alpha_1} \qquad w_c^{(i)} \rightarrow \text{weight c at iteration } i$$

$$w_c^{(1)} = \left(\frac{1}{6}\right) e^{2(1/2 \ln 5)}$$

$$w_c^{(1)} = \frac{5}{6}$$ .5

others remain same

we have to find $h_2(x)$

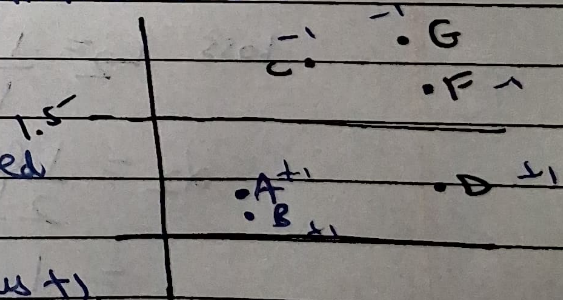
we make the cut with minimum loss:

~~$\cdot x_2 = 1.5$~~

$\boxed{\text{misclassified sample is D}}$

E, F, C are correctly classified as -1 (by majority)

A, B are correctly classified as $+1$.

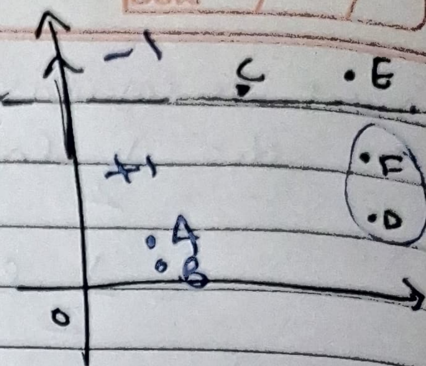


$$L_2 = \frac{\sum_{i=1}^{n} w_i \, I(y_i \neq h_2(x_i))}{\sum_{i=1}^{n} w_i}$$

$$L_2 = \frac{1/6}{5/6 + 5/6} \quad \text{(only D)} \quad = \frac{1}{10}$$

.5

if $\cdot\, x_2 = 2.5$



points C and E are correctly classified as -1.
points A and B are correctly classified as $+1$.
F and D are misclassified.
we find loss

$$L_2 = \frac{1/6 + 1/6}{1/6 + 1/6 + 5/6 + 1/6 + 1/6 + 1/6}$$

.5

$$L_2 = 2/10$$

Loss at split $x_2 = 1.5$ is ~~less~~ lesser than split at $x_2 = 2.5$, we make the split at

$$\boxed{x_2 = 1.5}$$

i.e $h_2(x) \begin{cases} +1 & x_2 \leq 1.5 \\ -1 & x_2 > 1.5 \end{cases}$

loss $L_2 = \frac{1}{10}$

mention that cut at 1.5 will be h2, assign .5 marks

$$\alpha_2 = \frac{1}{2} \ln \left(\frac{1 - L_2}{L_2} \right)$$

$$= \frac{1}{2} \ln \left(\frac{1 - 1/10}{1/10} \right)$$

$$\boxed{\alpha_2 = \frac{1}{2} \ln 9}$$

move this calculation to part b
[.5]

Weights will update for points D

$$w_D^{(2)} = 1/6 \cdot 9 = 3/2$$

(b) for the boosted classifier

$$\cancel{f_2(x) = \sum_{J=1}^{2}}$$

$$f(x) = sgn\left(\sum_{J=1}^{2} \alpha_J \cdot h_J(x)\right)$$

$$f(\bar{x}) = sgn(\alpha_1 h_1(x) + \alpha_2 h_2(x))$$

$$sgn\left[\frac{1}{2}\ln 5 \; h_1(x) + \frac{1}{2}\ln 9 \; h_2(x)\right]$$

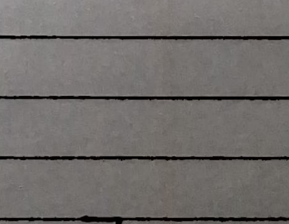
for $x = \begin{bmatrix} 1.5 \\ 4 \end{bmatrix}$ $\rightarrow h_1(x) = +1$ (cus $x_1 \leq 2.5$)

$\rightarrow h_2(x) = -1$ (cus $x_2 > 1.5$)

$$f(x) = sgn\left(\frac{1}{2}\ln 5 \cdot +1 + \frac{1}{2}\ln 9 \cdot -1\right)$$

$$sgn\left(\frac{1}{2}\ln 5 - \frac{1}{2}\ln 9\right)$$ 1

$$sgn\left(\frac{1}{2}\ln\frac{5}{9}\right)$$



-ve

$\rightarrow$ (-1)

we predict

$\underline{f(x) = -1}$ for $x = \begin{bmatrix} 1.5 \\ 4 \end{bmatrix}$

OURSTORYBEGINS

Q3.) For a given input x, the expected squared error is $E\left[(y-\hat{f}(x))^2\right]$

$y = f(x) + \eta$, $\Rightarrow E\left[(f(x) + \eta - \hat{f}(x))^2\right]$ .5

$\Rightarrow E\left[(f(x)-\hat{f}(x))^2\right] - 2E\left[\eta\left(f(x)-\hat{f}(x)\right)\right] + E\left[\eta^2\right]$

since $\eta \sim N(0, \sigma^2) \Rightarrow E\left[\eta^2\right] = \sigma^2$

.5 $\Rightarrow E\left[(f(x)-\hat{f}(x))^2\right] + \sigma^2 - 0$

adding and substituting $E\left[\hat{f}(x)\right]$

$\Rightarrow$ .5 $E\left[\left(\left(f(x) - E\left[\hat{f}(x)\right]\right) + \left(E\left[\hat{f}(x)\right] - \hat{f}(x)\right)\right)^2\right] + \sigma^2$

2ab term is missing, please edit 2E(f-Ef^)(Ef^-f^) = 2(f-Ef^)(Ef^ - Ef^). [.5]

$\Rightarrow E\left[\left(f(x) - E\left[\hat{f}(x)\right]\right)^2\right] + E\left[\left(E\left[\hat{f}(x)\right] - \hat{f}(x)\right)^2\right] - 0 + \sigma^2$

$\Rightarrow$ $Bias^2$ + Variance + ~~noise~~ σ^2

1

σ^2 also may present the inherent noise