Q.1) We take arbitary number of random dimensions of the data for each split. Let this be m.

i.e $m \longrightarrow$ random no. of dimensions to be picked for any given split.

Let $d \longrightarrow$ dimension of input i.e $x \in \mathbb{R}^d$

For bagging : $m = d$

0.5

For random forest : $m = \sqrt{d}$.

Q2. Out of bag error is the mean error for predicting the label ~~for~~ a training example such that for any example, only those trees are used which do not contain it.

$$\text{OOB error} = \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(y_i - \hat{y}_{OOB_i})$$

where $\mathcal{L}$ is the loss function,

y_i is the true label for example i

N is the no. of examples

$\hat{y}_{OOB_i}$ is the out of bag prediction for any example i, calculated as.

$$\hat{y}_{OOB_i} = \frac{1}{B^*} \sum_{j=1}^{B^*} \hat{f}_i(x_i)$$

where $\hat{f}_i$ are trees which don't have

1 (x_i, y_i) in their bootstrap sample.

Q3. $\hat{y} = w_0 \ln(w_1 e^x)$

$= w_0(\ln w_1 + \ln e^x)$

$= w_0 \ln w_1 + w_0 x$

0.5 $= ax + b$, $a = w_0$, $b = w_0 \ln w_1$.

We have samples (x_i, y_i) as $\{(0,1), (1,3)\}$.
Fit $\hat{y} = ax+b$ with OLS on:

1. $\vec{x}_1 = (0,1)$; $y_1 = 1$
2. $\vec{x}_2 = (1,1)$; $y_2 = 3$

[Add extra dimension for bias term]

Fit with OLS: $X = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $Y = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$\bar{w} = (X^TX)^{-1} X^TY$

$= \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$= \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ 0.5

$a = 2$, $b = 1$

$w_0 = 2$, $w_0 \ln w_1 = 1$

$\ln w_1 = \frac{1}{2}$

$\boxed{w_0 = 2, \ w_1 = \sqrt{e}}$ 0.5

__/__/__

Alternate:

Since we have 2 examples and 2 variables, we can exactly fit the variables ~~w_0 and w_1~~ slope and intercept. Let $\hat{y} = ax + b$.

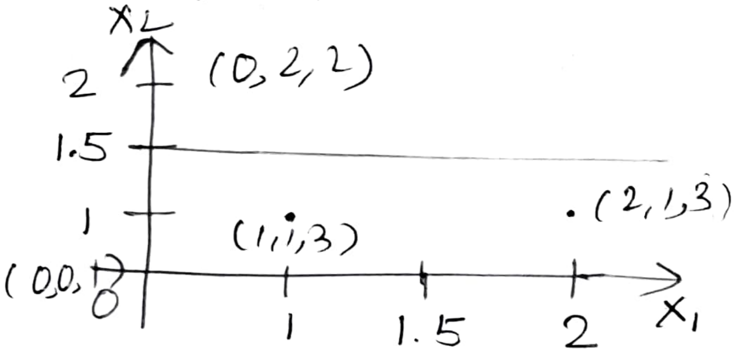
$$① \quad 3 = 1 \cdot a + b$$

$$② \quad 1 = 0 \cdot a + b$$

We get: $b = 1, \quad a = 2$. We can then solve for w_0, w_1 as before.

Q4) $D = \{(0,0,1); (1,1,3); (2,1,3); (0,2,2)\}$

Case 1: Split at $(0, 1.5)$



$(0,2,2)$ on one side

$(0,0,1)$; $(1,1,3)$; $(2,1,3)$ on other side

$c_1 = 2, \quad c_2 = \frac{1+3+3}{2} = \frac{7}{3}$.

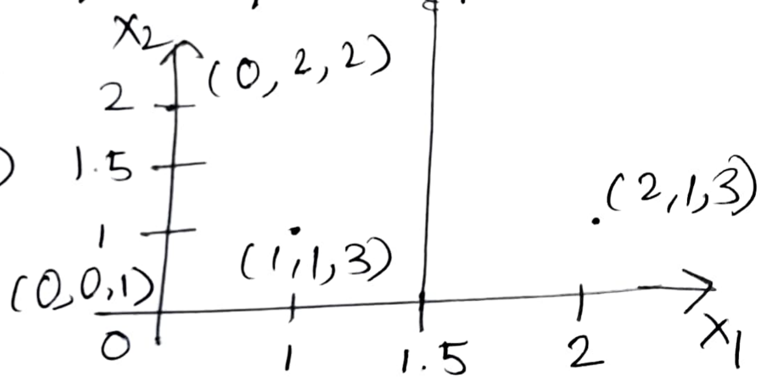
$$SSR = (2-2)^2 + (7/3 - 1)^2 + (3 - 7/3)^2 + (3 - 7/3)^2$$

$$= \left(\frac{4}{3}\right)^2 + 2\left(\frac{2}{3}\right)^2$$

$$= \frac{16}{9} + 2 \cdot \frac{4}{9} = \frac{16}{9} + \frac{8}{9} = \frac{24}{9} = 2.667$$

.5

Case 2: Split at $(1.5, 0)$



$(0,2,2)$; $(1,1,3)$; $(0,0,1)$ on one side

$(2,1,3)$ on the other side

$c_1 = \frac{2+3+1}{3} = \frac{6}{3} = 2, \quad c_2 = 3$

$$SSR = (2-2)^2 + (3-2)^2 + (1-2)^2 + (3-3)^2$$

$$= 0 + 1 + 1 = 2.$$

.5

$\therefore$ $SSR_{case\ 2} < SSR_{case\ 1}$

$\Rightarrow$ split at $(1.5, 0)$ leads to a better decision stump

Prediction of ~~a~~ decision stump at $(1.5, 0)$

are $= \begin{cases} 3 & \text{if } x_1 \geq 1.5 \\ 2 & \text{if } x_1 < 1.5 \end{cases}$

.5

Q5) $D = \{(x,y)\} = \{(1,1), (2.5,1), (3.5,1), (5,-1), (6.5,-1)\}$

$$WGI = \sum_{k=1}^{K} p'_{mk}(1-p'_{mk})$$

where $$p'_{mk} = \frac{\sum_{i=1}^{N} w_i I(x_i = k)}{\sum_{i=1}^{N} w_i}$$

Initialise : $w_i = \frac{1}{N} = \frac{1}{5}$

Iteration 1

Decision stump at cut 2

$$p'_{m(+1)} = \frac{1(\frac{1}{5})}{5(\frac{1}{5})} = \frac{1}{5}$$

$$p'_{m(-1)} = \frac{4(\frac{1}{5})}{5(\frac{1}{5})} = \frac{4}{5}$$

.5 $$WGI = \frac{1}{5}\left(1-\frac{1}{5}\right) + \frac{4}{5}\left(1-\frac{4}{5}\right) = \frac{8}{25}$$

Decision stump at cut 3

$$p'_{m(+1)} = \frac{2(\frac{1}{5})}{5(\frac{1}{5})} = \frac{2}{5}$$

$$p'_{m(-1)} = \frac{3(\frac{1}{5})}{5(\frac{1}{5})} = \frac{3}{5}$$

.5 $$WGI = \frac{2}{5}\left(1-\frac{2}{5}\right) + \frac{3}{5}\left(1-\frac{3}{5}\right) = \frac{12}{25}$$

Choose decision stump at cut 2 as it has lower WGI.

$$d_1 = \log\left(\frac{1 - WGI}{WGI}\right)$$ .5

$$= \log\left(\frac{17}{8}\right)$$

$$h_1(x) = \begin{cases} +1, & x < 2 \\ -1, & x > 2 \end{cases}$$

$$w_2 \leftarrow w_2 e^{\log(\frac{17}{8})}$$
$$w_3 \leftarrow w_3 e^{\log(\frac{17}{8})}$$

$w_1 = \frac{1}{5}, w_2 = \frac{17}{40}, w_3 = \frac{17}{40}, w_4 = \frac{1}{5}, w_5 = \frac{1}{5}$ .5

Iteration 2

Decision stump at cut 2

even is there is a calculation mistake, you can award full marks if the approach is fine in iteration 2

$$P_{m(+1)} = \frac{(\frac{1}{5})}{3(\frac{1}{5}) + 2(\frac{17}{40})} = \frac{4}{29}$$

.5

$$P_{m(-1)} = \frac{2(\frac{17}{40}) + 2(\frac{1}{5})}{3(\frac{1}{5}) + 2(\frac{17}{40})} = \frac{25}{29}$$

$$WGI = \frac{4}{29}\left(1 - \frac{4}{29}\right) + \frac{25}{29}\left(1 - \frac{25}{29}\right) = \frac{200}{841}$$

Decision stump at cut 3

$$P_{m(+1)} = \frac{\frac{1}{5} + \frac{17}{40}}{3(\frac{1}{5}) + 2(\frac{17}{40})} = \frac{25}{58}$$

$$P_{m(-1)} = \frac{\frac{17}{40} + 2(\frac{1}{5})}{3(\frac{1}{5}) + 2(\frac{17}{40})} = \frac{33}{58}$$

$$WGI = \frac{25}{58}\left(1 - \frac{25}{58}\right) + \frac{33}{58}\left(1 - \frac{33}{58}\right) = \frac{825}{1682}$$

Choose decision stump at cut 2 as it has lower WGI.

$$\alpha_2 = \log\left(\frac{1-\frac{200}{841}}{\frac{200}{841}}\right) = \log\left(\frac{641}{200}\right)$$

$$h_2(x) = \begin{cases} +1, & x < 2 \\ -1, & x > 2 \end{cases}$$

$$w_2 \leftarrow w_2 \, e^{\log\left(\frac{641}{200}\right)}$$

$$w_3 \leftarrow w_3 \, e^{\log\left(\frac{641}{200}\right)}$$

.5 $w_1 = \frac{1}{5}$, $w_2 = \frac{641}{1000}$, $w_3 = \frac{641}{1000}$, $w_4 = \frac{1}{5}$, $w_5 = \frac{1}{5}$

(b) Prediction of sample $x = 4$

$$h(x) = \text{sign}\left(\alpha_1 h_1(x) + \alpha_2 h_2(x)\right)$$ .5

$$h(4) = \text{sign}\left(\log\left(\frac{17}{8}\right)(-1) + \log\left(\frac{641}{200}\right)(-1)\right)$$

$$= -1$$

due to calculation mistake one may not get the correct answer but if they have right approach in part (a) and have correctly given the boosted tree formula, you can award full marks

Hence, prediction for sample $x = 4$ is -1.