30/Nov/2017

Homework # 1

1. Decision Trees:

1. (a) $\underbrace{(x_1 \wedge x_2) \vee (x_1 \wedge x_3)}$

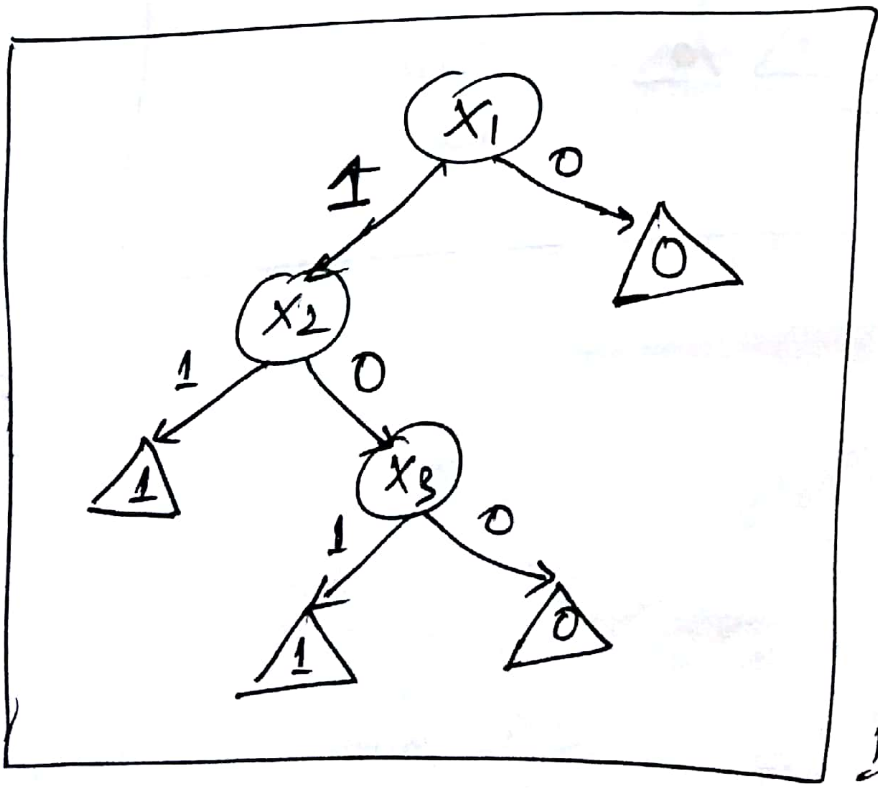
for this to be "True" both x_1 & x_2 has to be "True"

for this to be "true" both, x_1 & x_3 has to be "true"

for this to be "true" x_1 must be true and either of x_2 or x_3 should be true.

True = 1
False = 0

So,



Decision Tree.

1. (b)

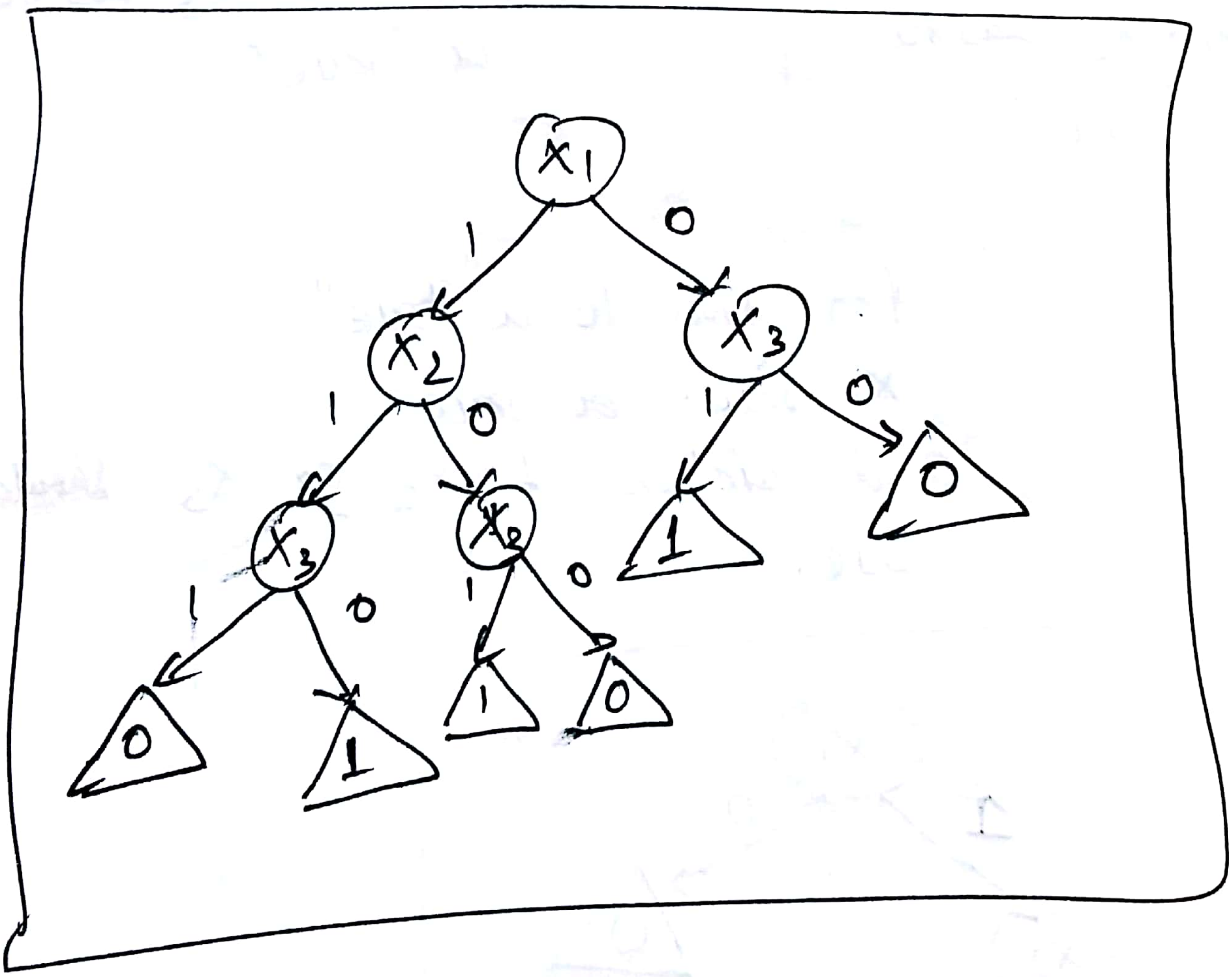
$(x_1 \wedge x_2)$ XOR (x_3)

for this to be 1, both x_1 & x_2 must be 1

for this to be 1 only one of $(x_1 \wedge x_2)$ or x_3 must be 1.

XOR

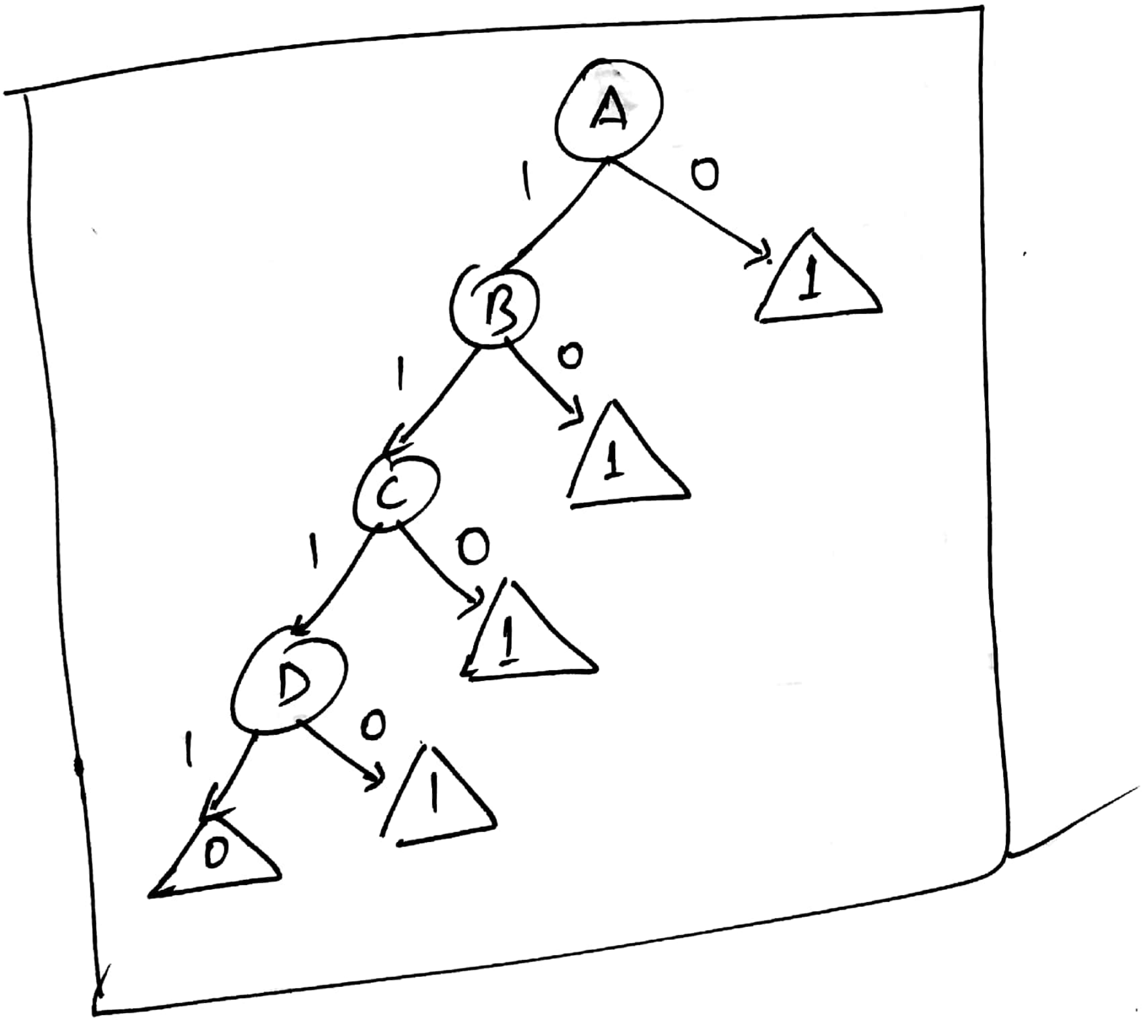
a	b	xor(a,b)
0	0	0
0	1	1
1	0	1
1	1	0



1.C

$$\underbrace{\neg A \vee \neg B \vee \neg C \vee \neg D}$$

This is 1 if any one of A or B or C or D is 0.



② features:

1. Superior Technolog $\in \{Y, N\}$.
2. Environment $\in \{Y, N\}$
3. Human $\in \{DC, L, H\}$
4. Distance $\in \{1, 2, 3, 4 \text{ lyrs.}\}$

$\boxed{2.(a)}$ Total no. of possible combinations of the input features :

$$m = 2 \times 2 \times 3 \times 4 = 48$$

$\therefore$ o/p $\in \{Y, N\}$

$$\boxed{\therefore \left(\text{Possible size of Hypothesis Space}\right) = 2^{48}}$$

2.(b) $H(S) = -\sum_{i=\{Y,N\}} p_i \log_2(p_i)$

$$= -\left[p_{\text{Invade=Yes}} \cdot \log_2(\text{invade=Yes}) + p_{\text{invade=No}} \cdot \log_2(\text{invade=No}) \right]$$

$$= -\left[\frac{5}{9} \log_2(5/9) + \frac{4}{9} \log_2(4/9) \right]$$

$\boxed{H(S) = 0.991}$ Entropy of the labels.

$S : [5^+, 4^-]$
$H(S) = 0.991$

2.(c):

$S:(5^+, 4^-)$; $H(S) = 0.991$

Techno.

Y → $[1^+, 2^-]$

N → $[4^+, 2^-]$

$$\boxed{H(\text{Tech}=Y) = -\left[\frac{1}{3} \log(1/3) + \frac{2}{3} \log(2/3) \right] = 0.918}$$

$$H(\text{Tech} = \text{No}) = -\left(\frac{4}{6}\log(4/6) + \frac{2}{6}\log(2/6)\right)$$

$$= -\left[\frac{2}{3}\log(2/3) + \frac{1}{3}\log(1/3)\right]$$

$$= 0.918$$

$$\therefore IG(S, \text{Tech}) = H(S) - \sum \frac{|S_v|}{|S|} \cdot H(S_v)$$

$$= H(S) - \left[\frac{3}{9} \times 0.918 + \frac{6}{9} \times 0.918\right]$$

$$= H(S) - 0.918$$

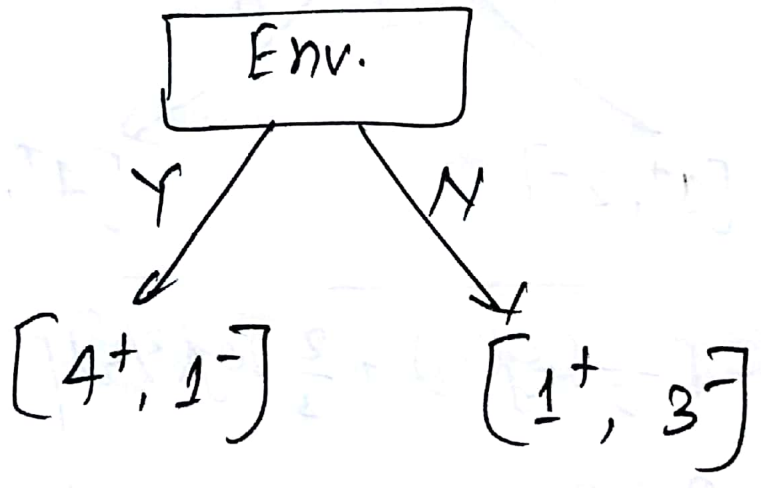
$$= 0.991 - 0.918$$

$$\boxed{IG(S, \text{Tech}) = 0.073}$$

✓

*

$S: [5^+, 4^-]$; $H(S) = 0.991$



4^+ 1^-

1^+ 3

$H(\text{Env.} = Y) = -\left(\frac{4}{5}\log\frac{4}{5} + \frac{1}{5}\log(1/5)\right)$

$= 0.722$

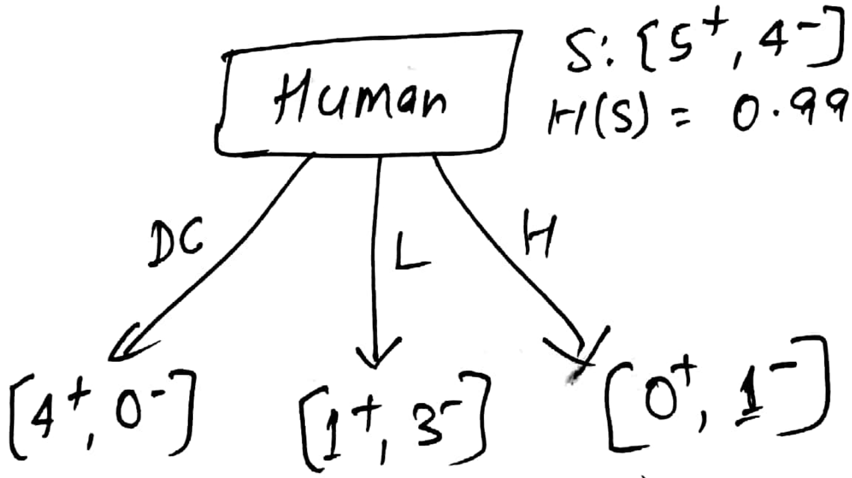
$H(\text{Env.} = N) = -\left(\frac{1}{4}\log\frac{1}{4} + \frac{3}{4}\log(3/4)\right)$

$= 0.811$

$IG(S, \text{Env.}) = H(S) - \left(\frac{5}{9}\times 0.722 + \frac{4}{9}\times 0.811\right)$

$= 0.991 - \left(\frac{5}{9}\times 0.722 + \frac{4}{9}\times 0.811\right)$

$\boxed{IG(S, \text{Env.}) = 0.229}$



S: [5+, 4-]
H(S) = 0.991

$1^+ \, 3^-$

$H(\text{Human} = DC) = -\left(\frac{4}{4}\log(4/4) + \frac{0}{4}\log(0/4)\right) = 0$

$H(\text{Human} = L) = -\left(\frac{1}{4}\log\left(\frac{1}{4}\right) + \frac{3}{4}\log(3/4)\right) = 0.811$

$H(\text{Human} = H) = \frac{0}{1}\log 0 + \frac{1}{1}\log 1 = 0$

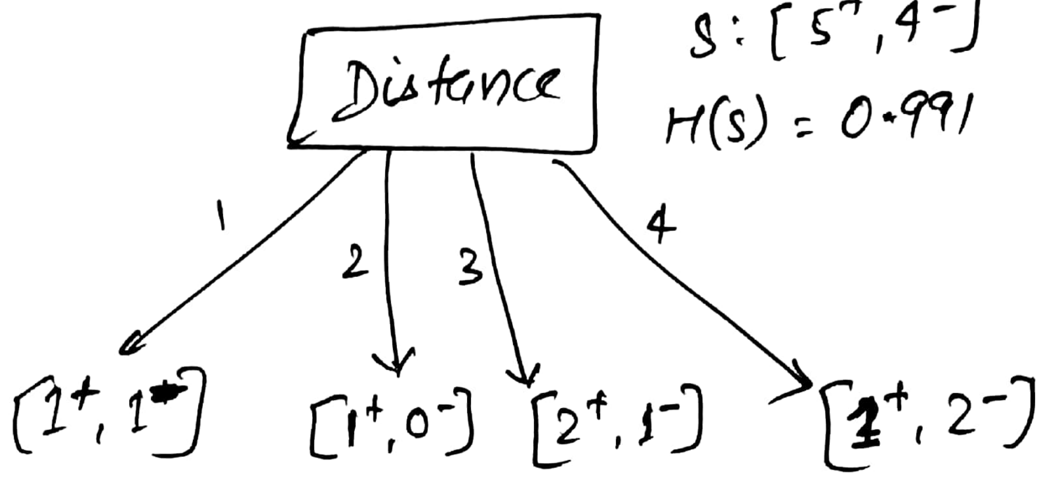
$$\therefore IG(S, \text{Human}) = H(S) - \left[\frac{4}{9}\times 0 + \frac{4}{9}\times(0.811) + \frac{1}{9}\times 0\right]$$

$$= \left(0.991 - \frac{4}{9}\times 0.811\right)$$

$$\boxed{IG(S, \text{Human}) = 0.630}$$

$S : [5^+, 4^-]$

$H(S) = 0.991$



$1^+ \; 1^-$

1^+

$2^+ \; 1$

$1^+ 2^-$

$$H(\text{Distance} = 1) = -\left(\frac{1}{2}\log \tfrac{1}{2} + \frac{1}{2}\log \tfrac{1}{2}\right) = 1$$

$$H(\text{Distance} = 2) = -\left(\frac{1}{1}\log(1) + 0\log 0\right) = 0$$

$$H(\text{Distance} = 3) = -\left(\frac{2}{3}\log(2/3) + \frac{1}{3}\log(1/3)\right) = 0.918$$

$$H(\text{Distance} = 4) = -\left(\frac{1}{3}\log(1/3) + \frac{2}{3}\log(2/3)\right) = 0.918$$

$$\therefore IG(S, \text{Distance}) = H(S) - \left(\frac{2}{9}\times 1 + 0 + \frac{3}{9}\times 0.918 + (3/9)\times 0.918\right)$$

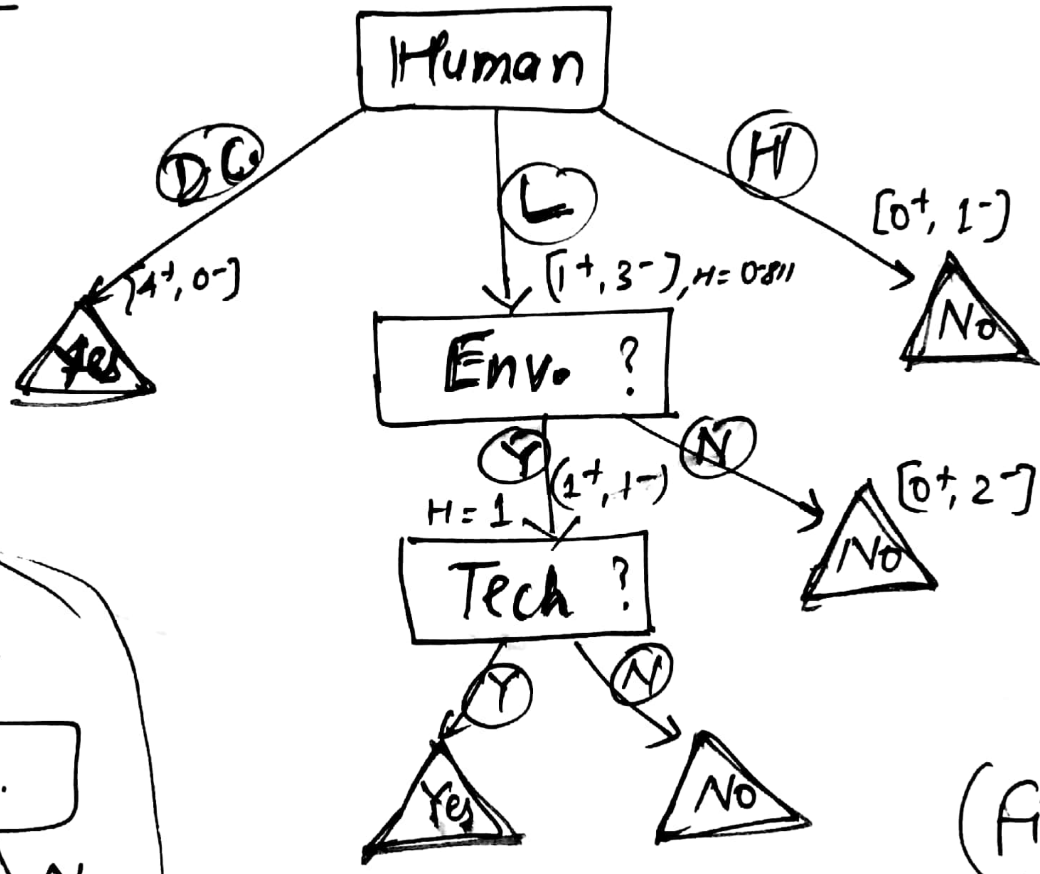
$$= \cancel{0.991} \; (\; = \boxed{0.156}$$

2.(d) Our goal is to uniformly classify the data as we increase the ~~[illegible]~~ level of DT using ID3.

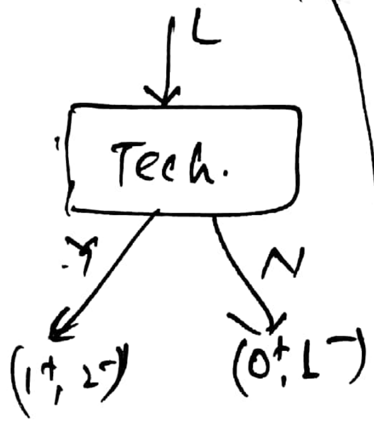
So, we choose the "attribute" which has the maximum IG.

i.e. our root node will be $\boxed{\text{Human}}$

2.(e)

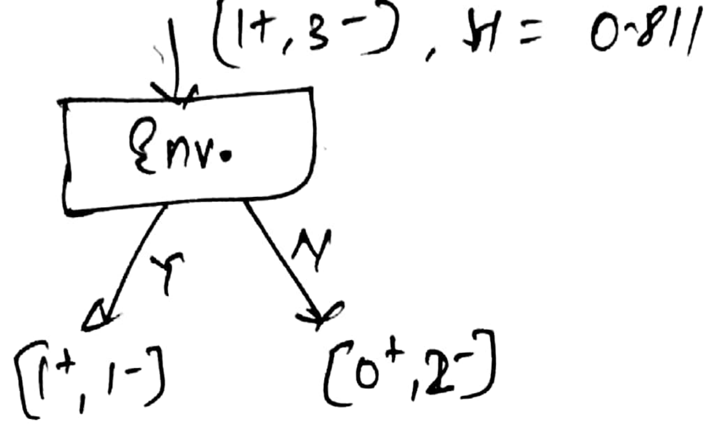


(fig: 2.e)

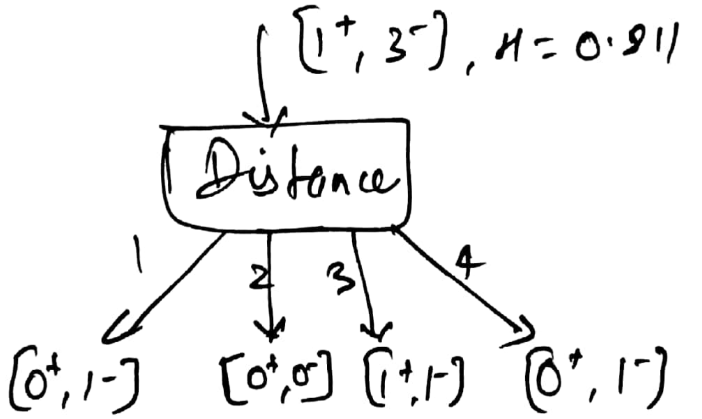


$\cancel{IG = 0.811 - \left(\frac{3}{4} \times \log\frac{3}{4}\right)}$

$$IG = 0.811 - \left[\frac{3}{4}\left(-\left(\frac{1}{3}\log\left(\frac{1}{3}\right) + \frac{2}{3}\log\left(\frac{2}{3}\right)\right)\right) + \frac{1}{4} \times 0\right] = 0.122$$

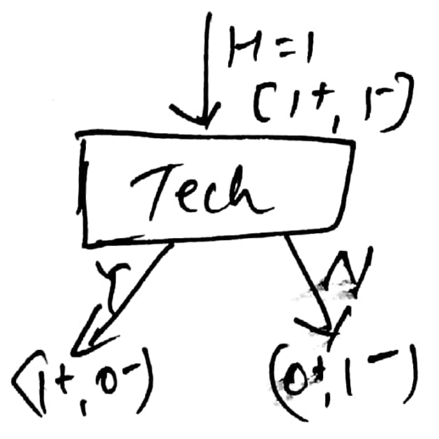


$$IG(\text{Human, Env.})$$
$$= 0.811 - \left[\frac{2}{4}\right]\left[1 + 0\right]$$
$$= 0.811 - \frac{1}{2}$$
$$= \underline{\underline{0.311}}$$

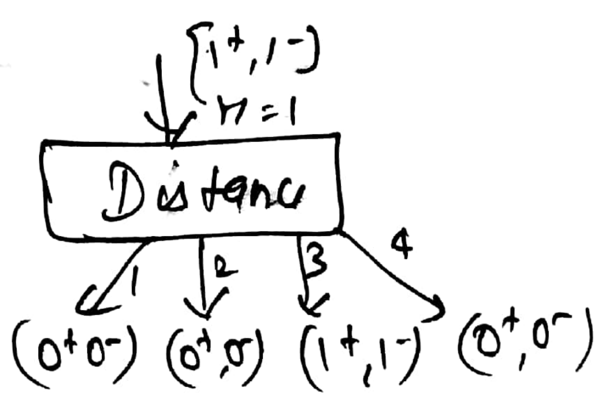


$$IG(\text{Human, Distance})$$
$$= 0.811 - \left[\left(\frac{1}{4} \times 0\right) + (0 \times 0) + \left(\frac{2}{4} \times 1\right) + \frac{1}{4} \times 0\right]$$
$$= 0.811 - \left(\frac{1}{2}\right) = \underline{\underline{0.311}}$$

We can pick either $\boxed{\text{Distance}}$ or $\boxed{\text{Env.}}$ as our next node:



$$IG = 1 - \left[\frac{1}{2} \times 0 + \frac{1}{2} \times 0\right]$$
$$= 1$$



$$IG = 1 - \left(\frac{2}{2} \times 1\right) = 0$$

2·(f)

Ex.	Tech.	Env.	Human	Distance	Invade?	Predicted Value of Invade?
(A)	Yes	Yes	Like	2	No	Yes
(B)	No	No	Hate	3	No	No
(C)	Yes	Yes	Like	4	Yes	Yes

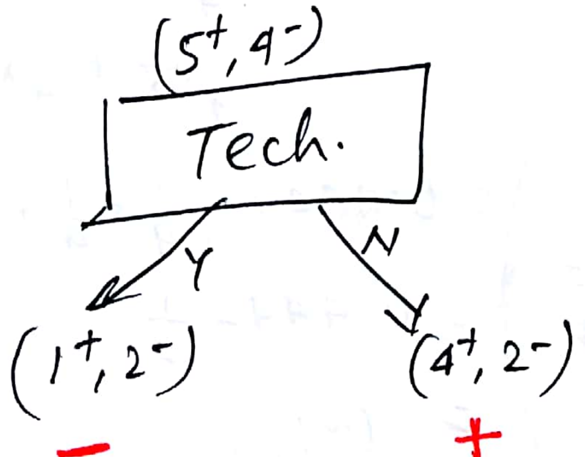
Accuracy = ?

Two values. (i.e. for test example B, C) are correctly predicted by our DT., but incorrectly predicted for ex. (A).

Thus, $\boxed{\text{Accuracy} = \frac{2}{3} = 66.67\%}$.

3·(a)

$$\text{MajorityError}(S) = 1 - \left(\frac{5^+}{9}\right) = \left(\frac{4}{9}\right) = 0.444$$



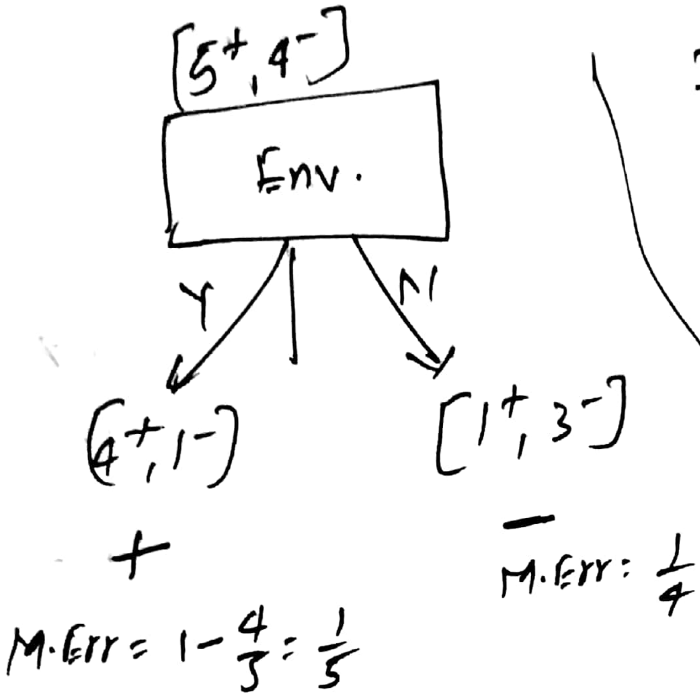
$\text{M.Err.} = 1 - \frac{2}{3} = \left(\frac{1}{3}\right)$

$\text{M.Err} = \frac{2}{6} = 1/3$

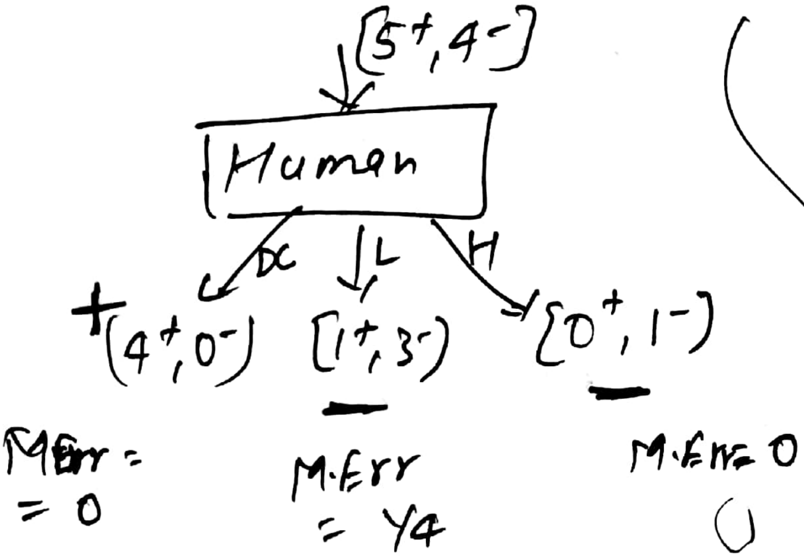
$$IG = (0.444) - \left[\frac{3}{9} \times \frac{1}{3} + \frac{6}{9} \times \frac{1}{3}\right]$$

$$= 0.444 - \left(\frac{1}{3}\right) \times 1$$

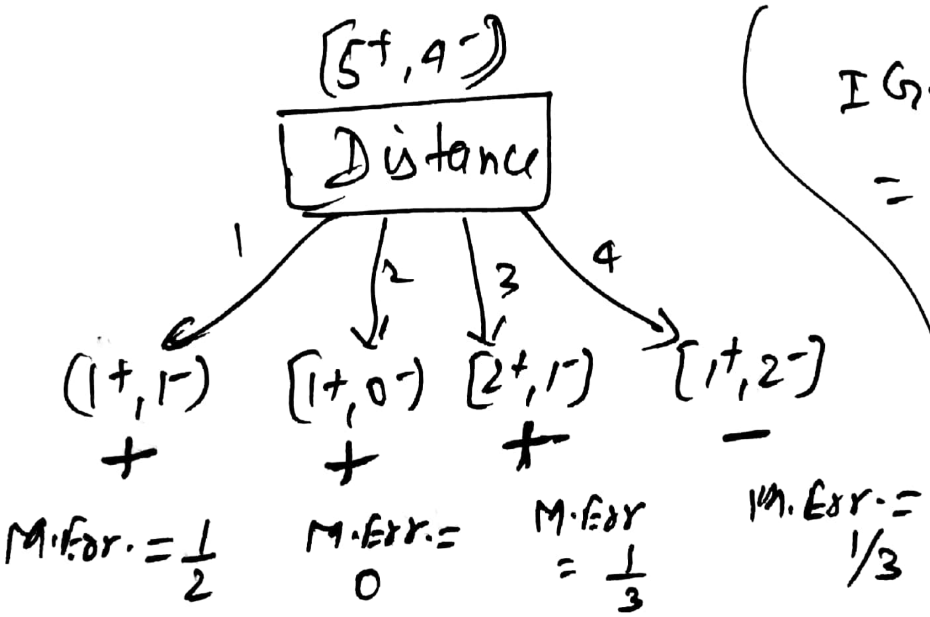
$$(IG)_{\text{Tech}} = 0.111$$



$$IG(S, Env.)$$
$$= 0.444 - \left[\frac{5}{9}\times\frac{1}{5} + \frac{4}{9}\times\frac{1}{4}\right]$$
$$= 0.444 - 2/9$$
$$= \underline{\underline{0.222}}$$



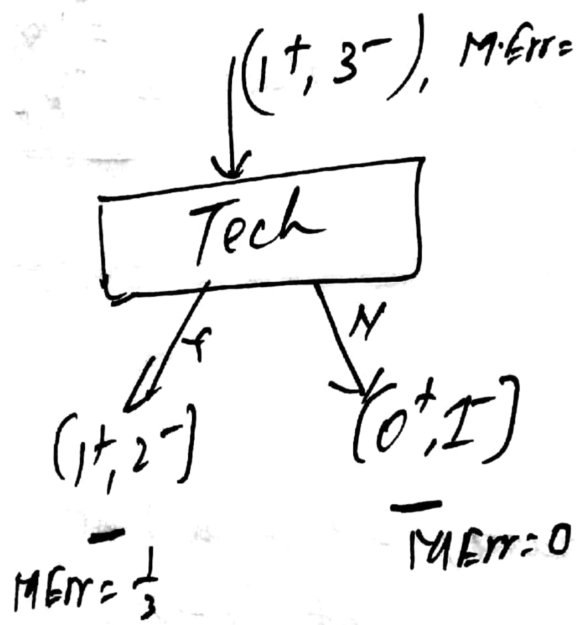
$$IG(S, Human)$$
$$= 0.444 - \left(\frac{4}{9}\times 0 + \frac{4}{9}\times\frac{1}{4} + \frac{1}{9}\times 0\right)$$
$$= 0.444 - \frac{1}{9}$$
$$= \underline{\underline{0.333}}$$



$$IG.(S, Distance)$$
$$= 0.444 - \left(\frac{2}{9}\times\frac{1}{2} + 0 + \frac{3}{9}\times\frac{1}{3} + \frac{3}{9}\times\frac{1}{3}\right)$$
$$= 0.444 - \left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9}\right)$$
$$= 0.444 - \frac{1}{3}$$
$$= \underline{\underline{0.111}}$$

1-3-(b) According to the previous results (using Maj. Err. as a measure of impurity) choosing $\boxed{\text{Human}}$ as a root node ~~decreases the~~ achieves maximum <u>Information Gain</u>.

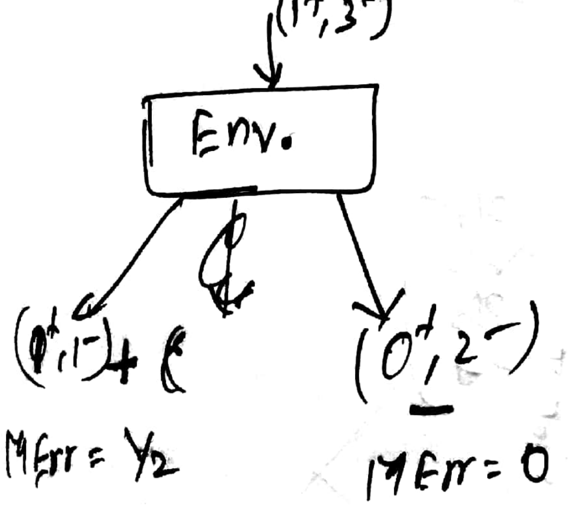
So; the root node should be $\boxed{\text{Human}}$.



IG (Human, Tech)

$= 0.250 - \left(\frac{3}{4} \times \frac{1}{3} + 0\right)$

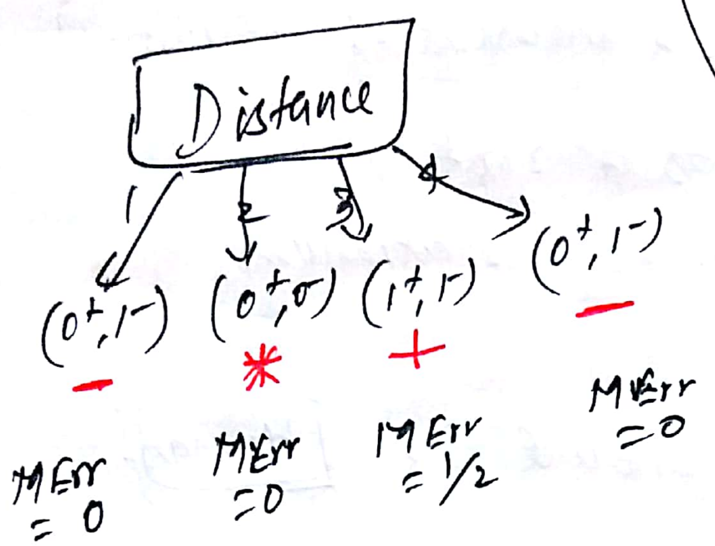
$= 0.25 - \frac{1}{4} = \underline{0}$



IG (Human, Env.)

$= 0.25 - \left(\frac{2}{4} \times \frac{1}{2} + 0\right)$

$= \underline{0}$



$IG(\text{Human, Distance})$

$= 0.25 - \left(\frac{2}{4} \times \frac{1}{2}\right) = \underline{0}$

Ans: NO ; the two measures (Entropy & Majority Error) donot lead to the same Decision Tree.

2. Linear Classifiers:

2.1 Looking at the training data; we find that one of the weight vectors that totally satisfy the data is

$$w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} \qquad \cancel{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$2x_4 - 1$

$b = 0 (-1)$

$\therefore\ y = w^T x + b \Rightarrow \boxed{y = w^T x}$

$y \in \{1, -1\}$

$w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}$; $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$; $b = -1$

$$\boxed{y = w^T x + b} \Rightarrow$$

2.2.

x_1	x_2	x_3	x_4	O	Prediction	
0	0	0	1	1	1	
0	0	1	1	1	1	
0	0	0	0	-1	-1	
1	0	1	0	1	-1	← Error
1	1	0	0	1	-1	← Error
1	1	1	1	1	1	
1	1	1	0	1	-1	← Error

x_1	x_2	x_3	x_4	O	Prediction	
0	1	0	0	-1	-1	
0	1	1	0	-1	-1	
0	1	1	1	1	1	
1	0	0	0	1	-1	← Error
1	0	0	1	1	1	
1	1	0	1	1	1	

Testing Dataset.

2.3

x_1	x_2	x_3	x_4	O
1	0	1	1	1
0	1	0	1	1
0	0	1	0	-1
0	0	0	1	1
0	0	1	1	1
0	0	0	0	-1
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1
1	1	1	0	1
0	1	0	0	-1
0	1	1	0	-1
0	1	1	1	1
1	0	0	0	1
1	0	0	1	1
1	1	0	1	1

Complete Training Dataset

0	0	1	0	-1
0	0	0	0	-1
0	1	0	0	-1
0	1	1	0	-1

Based on above observation we find that if $(\neg x_1 \wedge \neg x_4)$ is True(1) then

$o = -1$

$$w = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \; ; \; x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \; ; \; b = -1$$

i.e. $(y = -1)$ if $(\neg x_1 \wedge \neg x_4)$ is True.

i.e.

$$(y = -1) \text{ if } \left(\left((1 - x_1) + (1 - x_4) \right) \geq 2 \right)$$

$$\Rightarrow \boxed{x_1 + x_4 \leq 0}$$

$\because x_1, x_4 \in \{0, 1\}$

$\therefore (x_1 + x_4) \leq 0 \Rightarrow \boxed{(x_1 + x_4) = 0}$