NAVJOT KAUR
301404765

CMPT-310
ASSIGNMENT-4

ANSWER-1 :-

GIVEN : Table of holiday trips
TABLE ATTRIBUTES : Country, Season, Type, weeks
TARGET (GOAL) : much (positive)
(FUN) little (negative)

a. DECISION TREE WITH INFORMATION GAINS :

FORMULA :

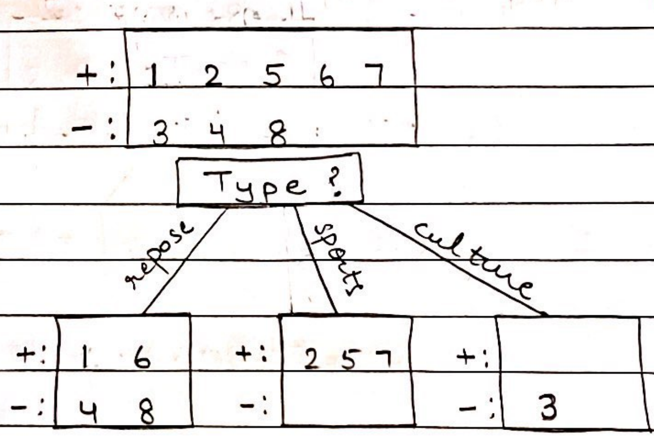
$$Gain(A) = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - Remainder(A)$$

where

$$Remainder(A) = \sum_{i=1}^{k} \frac{p_i + n_i}{p+n} I\left(\frac{p_i}{p_i+n_i}, \frac{n_i}{p_i+n_i}\right)$$

where p = no. of positive examples
n = no. of negative examples

- Top layer :



$$\text{Gain (type)} = 1 - \left[\frac{4}{8} I\left(\frac{2}{4}, \frac{2}{4}\right) + \frac{3}{8} I(1, 0) + \frac{1}{8} I(0, 1)\right]$$

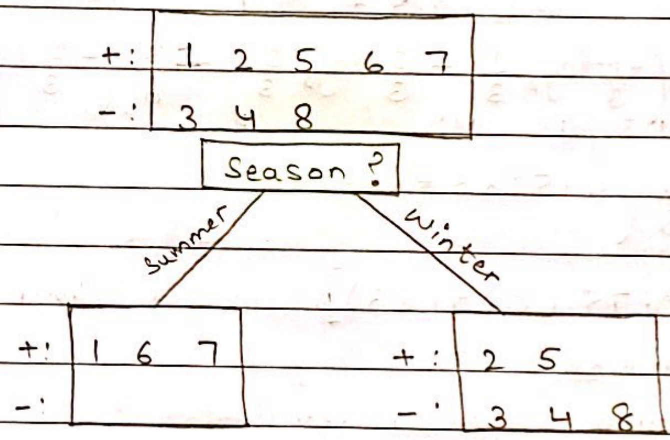
$$= 1 - \left[\frac{4}{8}\left(-\frac{2}{4}\log_2 \frac{2}{4} - \frac{2}{4}\log_2 \frac{2}{4}\right) + \frac{3}{8}(0) + \frac{1}{8}(0)\right]$$

$$= 1 - \left[\frac{1}{2}\left(-\frac{1}{2}\log_2 \frac{1}{2} - \frac{1}{2}\log_2 \frac{1}{2}\right) + 0 + 0\right]$$

$$= 1 - [0.5(0.5 + 0.5)]$$

$$= 1 - 0.5$$

$$= 0.5$$



$$\text{Gain (season)} = 1 - \left[\frac{3}{8} I(1,0) + \frac{5}{8} I\left(\frac{2}{5}, \frac{3}{5}\right)\right]$$

$$= 1 - \left[\frac{3}{8}(0) + \frac{5}{8}\left(-\frac{2}{5}\log_2 \frac{2}{5} - \frac{3}{5}\log_2 \frac{3}{5}\right)\right]$$

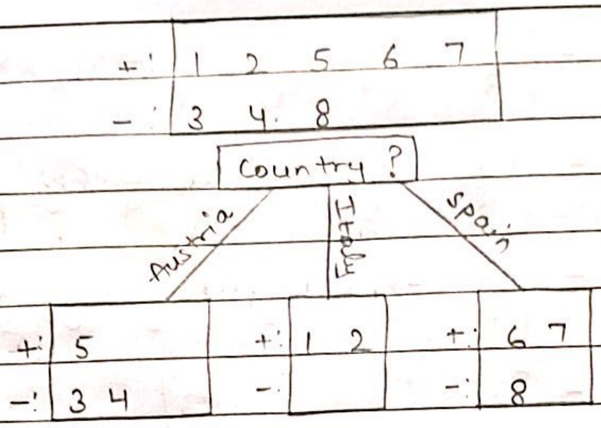
$$= 1 - \left[\frac{5}{8}(0.5288 + 0.4422)\right]$$

$$= 1 - [0.625(0.971)]$$

$$= 1 - 0.6069$$

$$= 0.3931$$

___/___/___



$$\text{Gain (country)} = 1 - \left[\frac{3}{8} I\left(\frac{1}{3}, \frac{2}{3}\right) + \frac{2}{8} I(1, 0) + \frac{3}{8} I\left(\frac{2}{3}, \frac{1}{3}\right)\right]$$

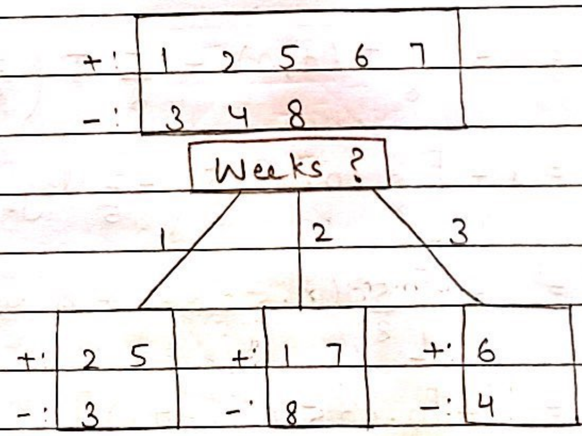
$$= 1 - \left[\frac{3}{8}\left(-\frac{1}{3}\log_2 \frac{1}{3} - \frac{2}{3}\log_2 \frac{2}{3}\right) + \frac{3}{8}\left(-\frac{2}{3}\log_2 \frac{2}{3} - \frac{1}{3}\log_2 \frac{1}{3}\right)\right]$$

$$= 1 - \left[\frac{6}{8}\left(0.5283 + 0.38997\right)\right]$$

$$= 1 - [0.75(0.9183)]$$

$$= 1 - 0.6887$$

$$= 0.3113$$



$$\text{Gain (weeks)} = 1 - \left[\frac{3}{8} I\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{3}{8} I\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{2}{8} I\left(\frac{1}{2}, \frac{1}{2}\right)\right]$$

$$= 1 - \left[0.6887 + \frac{1}{4}(1)\right] \quad \text{[from above]}$$

$$= 1 - [0.6887 + 0.25]$$

$$= 1 - 0.9387$$

$$= 0.0613$$

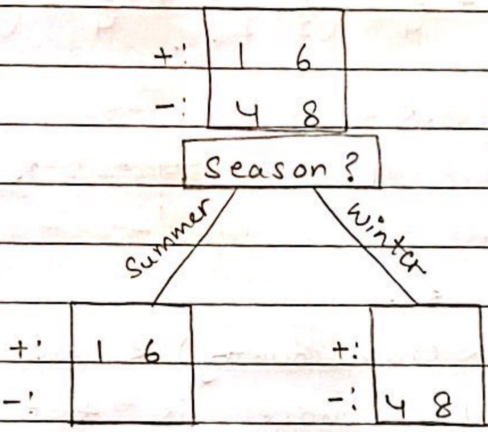
Here,

$0.0613 < 0.3113 < 0.3931 < 0.5$

i.e., Gain (weeks) < Gain (country) < Gain (season) < Gain (type)

So, we choose the top layer i.e. first layer of decision tree as (type).

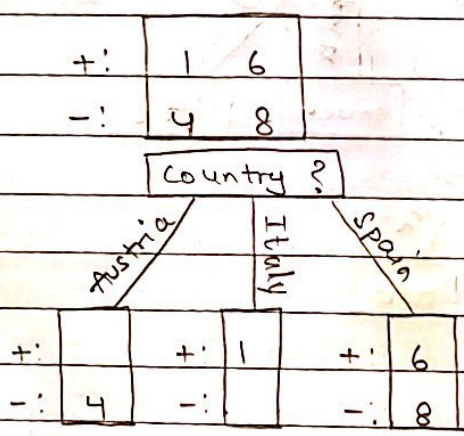
- **Second layer**



$$\text{Gain (season)} = 1 - \left[\frac{2}{4} I(1,0) + \frac{2}{4} I(1,0)\right]$$

$$= 1 - [0 + 0]$$

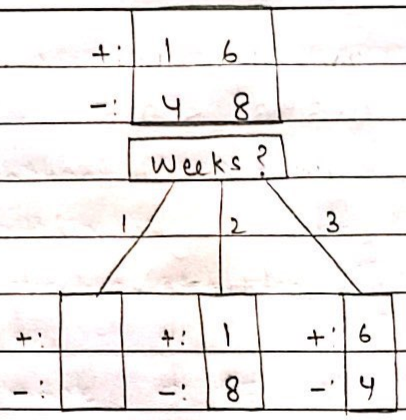
$$= 1$$



$$\text{Gain (country)} = 1 - \left[\frac{1}{4} I(0,1) + \frac{1}{4} I(1,0) + \frac{2}{4} I\left(\frac{1}{2}, \frac{1}{2}\right)\right]$$

$$= 1 - [0 + 0 + 0.5]$$

$$= 0.5$$



$$\text{Gain (weeks)} = 1 - \left[0 + \frac{2}{4} I\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{2}{4} I\left(\frac{1}{2}, \frac{1}{2}\right)\right]$$

$$= 1 - [0.5 + 0.5]$$

$$= 1 - 1$$

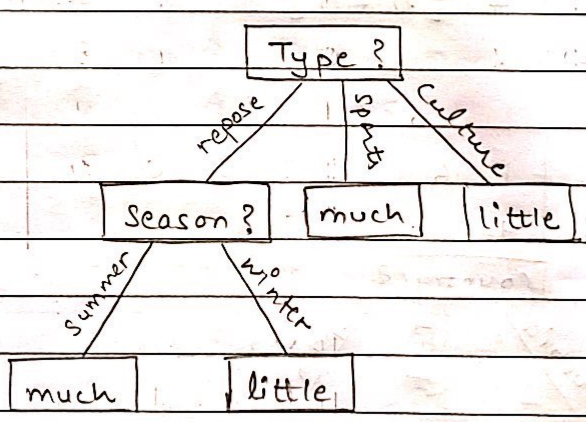
$$= 0$$

Here, $0 < 0.5 < 1$

Gain (weeks) < Gain (country) < Gain (season)

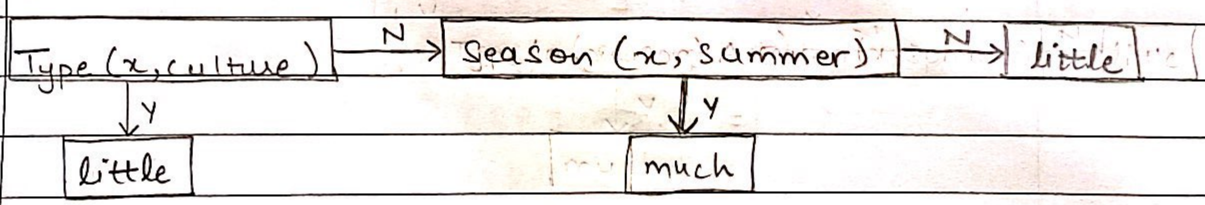
So, second layer we chose is that of season.

DECISION TREE BASED ON ABOVE CALCULATIONS:



b. DECISION LISTS:

i. The tests of DL contain as few literals as possible.



__/__/__

ii. The DL consists of as few tests as possible.

$$\begin{tikzcd}
\boxed{type(x, sports)} \arrow[r, "N"] \arrow[d, "Y"] & \boxed{type(x, repose) \wedge season(x, summer)} \arrow[r, "N"] \arrow[d, "Y"] & \boxed{little} \\
\boxed{much} & \boxed{much} &
\end{tikzcd}$$

One test is not possible so the DL is same as above with (two tests).

ANSWER 2 :-

GIVEN : Table on observations about movies

TABLE ATTRIBUTES : genre, actors, marketing, cost

GOAL : good (positive), bad (negative).

a.

example	I_D	I_c	I_s	$I_{\#}$	I_M	I_n	I_m	I_L	T
1	0	0	1	2	1	1	0	0	1
2	0	1	0	3	0	0	0	1	0
3	0	1	0	1	0	0	1	0	0
4	1	0	0	2	1	0	0	1	1
5	1	0	0	1	0	1	0	0	1
6	0	0	1	3	1	0	0	1	1
7	0	1	0	1	0	1	0	0	1
8	0	0	1	2	0	0	0	1	0

ATTRIBUTES	ENCODING
genre	distributed
actors	local
marketing	local
cost	distributed
reception	local

Scanned with CamScanner

_/_/_

b. NEURAL - NETWORK LEARNING :-

TABLE :-

O → preception output , E = T - O where T = correct output (E: Error)

Learning rate, $\alpha = 2$

$W_j := W_j + \alpha \times I_j \times E$

	example	O	E	W_D	W_c	W_s	$W_\#$	W_M	W_h	W_m	W_e
initial				+1	+1	+1	+1	+1	+1	+1	+1
	1	1	0								
	2	1	-1		-1		-5				-1
	3	0	0								
epoch 1	4	0	+1	+3			-1	+3			+1
	5	1	0								
	6	1	0								
	7	0	1		+1		+1		+3		
	8	1	-1			-1	-3				-1
	1	0	1			+1	+1	+5	+5		
	2	1	-1		-1		-5				-3
	3	0	0								
epoch 2	4	0	1	+5			-1	+7			-1
	5	1	0								
	6	1	0								
	7	1	0								
	8	1	0								
	1	1	0								
	2	0	0								
	3	0	0								
epoch 3	4	1	0								
	5	1	0								
	6	1	0								
	7	1	0								
	8	0	0								

__/__/____

ANSWER 3 :-

a. 3-ary Boolean function,

$$f_1(x_1, x_2, x_3) = (x_1 \equiv (x_2 \wedge x_3))$$

firstly,

$step_t(x) = 0$, if $x < t$

TRUTH TABLE:

$step_t(x) = 1$, if $(x \geq t)$

x_1	x_2	x_3	$x_2 \wedge x_3$	$x_1 \equiv (x_2 \wedge x_3)$
0	0	0	0	1
1	0	0	0	0
0	1	0	0	1
0	0	1	0	1
1	1	0	0	0
0	1	1	1	0
1	0	1	0	0
1	1	1	1	1

FOR HIDDEN LAYER :

As

$$x_1 \equiv (x_2 \wedge x_3) \Rightarrow x_1 \Leftrightarrow (x_2 \wedge x_3)$$

Then,

$$f_1(x_1, x_2, x_3) = (x_1 \equiv (x_2 \wedge x_3))$$
$$= (x_1 \Rightarrow (x_2 \wedge x_3))$$
$$= (x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge \neg(x_2 \wedge x_3))$$
$$= (x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge (\neg x_2 \vee \neg x_3))$$
$$= (x_1 \wedge x_2 \wedge x_3) \vee ((\neg x_1 \wedge \neg x_2) \vee (\neg x_1 \wedge \neg x_3))$$
$$= (x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge \neg x_2) \vee (\neg x_1 \wedge \neg x_3)$$

There are 3 neurons.

Now, $\longrightarrow$

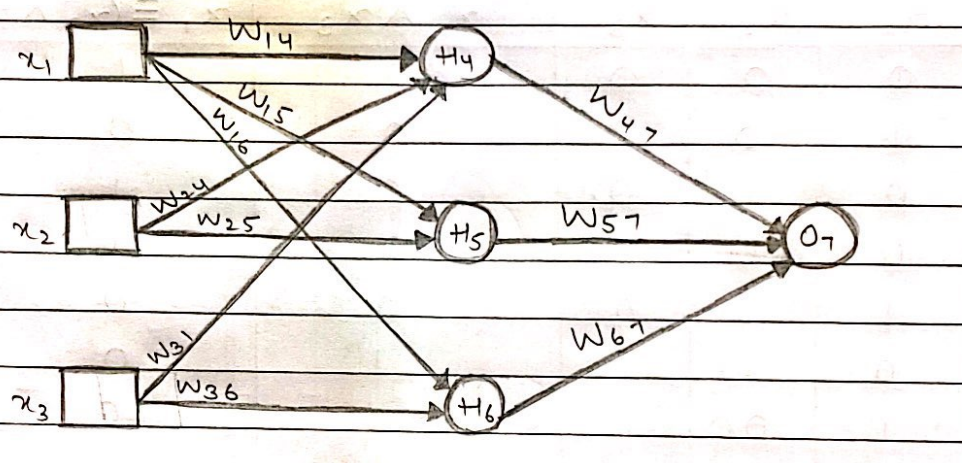
__/__/__

FEED FORWARD NEURAL NETWORK :-

ACTIVATION FUNCTION :

$$f(x) = \begin{cases} 0 , & x < 0 \\ 1 , & x \geq 0 \end{cases}$$

$$f(x_1, x_2, x_3) = (x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge \neg x_2) \vee (\neg x_1 \wedge \neg x_3)$$



Input Hidden layer Output

b. FUNCTION :

$$f_2(x_1, x_2, x_3) = (x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge \neg x_2 \wedge \neg x_3)$$

Firstly,

TRUTH TABLE :

x_1	x_2	x_3	$\neg x_1$	$\neg x_2$	$\neg x_3$	$x_1 \wedge x_2 \wedge x_3$	$\neg x_1 \wedge \neg x_2 \wedge \neg x_3$	$f_2(x_1, x_2, x_3)$
0	0	0	1	1	1	0	1	1
1	0	0	0	1	1	0	0	0
0	1	0	1	0	1	0	0	0
0	0	1	1	1	0	0	0	0
1	1	0	0	0	1	0	0	0
0	1	1	1	0	0	0	0	0
1	0	1	0	1	0	0	0	0
1	1	1	0	0	0	1	0	1

__/__/____

The Boolean function, $f_2(x_1, x_2, x_3)$ [given] cannot be represented by a perceptron using only step functions.

WHY? :

⇒ Using only step functions states that the perceptron works only when the boundaries are linearly separable.

PROOF: As discussed in lecture notes,
Assume otherwise, i.e, the function is represented by a perceptron using only step functions, then

$$f_2(w_1x_1, w_2x_2, w_3x_3) = \begin{cases} 0, & w_1x_1 + w_2x_2 + w_3x_3 < t \\ 1, & \text{otherwise} \end{cases}$$

then, from truth table we know,
$f_2(0,0,0) = 1 = f_2(1,1,1)$ and according to our assumption $f_2(1,0,0)$ (suppose) and other values should be 1 as well but truth table states otherwise. Hence, contradicting the statement.

Also, to back up the definition above, when a figure (3-D) is drawn, there is no way $f_2(0,0,0)$ and $f_2(1,1,1)$ can be separated from other values i.e, not linearly separable.

Hence, the function $f_2(x_1, x_2, x_3)$ cannot be represented by a perceptron using only step functions.