(21) Decision Tree

Patient ID	Chest Pain	male	Smoking	HA
1	Yes	Yes	No	Yes
2	Yes	Yes	Yes	Yes
3	No	No	Yes	Yes
4	No	Yes	No	No
5	Yes	No	Yes	Yes
6	No	Yes	Yes	No

Iteration (1) $\text{Info}(D) = I(4, 2)$

$$= -\sum P_i \log P_i$$

$$= \frac{-4}{6} \log\left(\frac{4}{6}\right) - \frac{2}{6} \log \frac{2}{6}$$

$$= 0.38 + 0.528 = 0.908$$

* Gain (Chest Pain) $\begin{cases} \text{Yes} \\ \text{No} \end{cases}$ $\boxed{\text{Info } D - \text{Info}_{\text{chestpain}}(D)}$

$$\therefore \text{Info}_{\text{chestpain}}(D) = \frac{3}{6} I(3,0) + \frac{3}{6} I(1,2)$$

$$\Rightarrow \frac{3}{6}\left[\frac{-3}{3}\log\left(\frac{3}{3}\right) - \frac{0}{3}\log\left(\frac{0}{3}\right)\right] + \frac{3}{6}\left[\frac{-1}{3}\log\left(\frac{1}{3}\right) - \frac{2}{3}\log\left(\frac{2}{3}\right)\right]$$

$$\Rightarrow 0 + \frac{3}{6}\left[+0.528 + 0.389\right]$$

$$\frac{3}{6} \times 0.917 = 0.4585$$

Gain (Chest pain) = $\text{Info}\,D - \text{Info}_{\text{chest pain}}\,D$

$= 0.908 - 0.458$

$= \boxed{0.45}$ ← ① gain

For Male → yes, No

$\text{gain}_{\text{male}} = \text{Info}\,D - \text{Info}(\text{male})$

$\text{Info male} = \frac{4}{6} I(2,2) + \frac{2}{6} I(2,0)$

$\Rightarrow \frac{4}{6}\left[\frac{-2}{4}\log\frac{2}{4} - \frac{2}{4}\log\left(\frac{2}{4}\right)\right] + \frac{2}{6}\left[\frac{2}{2}\log\left(\frac{2}{2}\right) - \frac{0}{2}\log\left(\frac{0}{2}\right)\right]$

$\Rightarrow \frac{4}{6}(0.5 + 0.5) + \frac{2}{6}(0 - 0)$

$\Rightarrow \frac{4}{6}$

$\Rightarrow \underline{0.66}$

$\text{gain}_{\text{male}} = 0.908 - 0.66$

$= \boxed{0.248}$ ← ② gain

For Smoking → ~~yes~~, No

$\text{gain}_{\text{smoking}} = \text{Info}\,D - \text{Info}(\text{smoking})\,D$

$\text{Info}_{\text{Smoking}} \Rightarrow \frac{4}{6} I(3,1) + \frac{2}{6} I(1,1)$

$$\Rightarrow \frac{4}{6}\left[-\frac{3}{4}\log\left(\frac{3}{4}\right) - \frac{1}{4}\log\frac{1}{4}\right] + \frac{2}{6}\left[-\frac{1}{2}\log\left(\frac{1}{2}\right) - \frac{1}{2}\log\left(\frac{1}{2}\right)\right]$$

$$= \frac{4^2}{63}\left(0.311 + 0.5\right) + \frac{21}{63}\left(+0.5 + 0.5\right)$$

$$0.540 \quad +\frac{1}{3}$$

$$= \underline{0.873}$$

$$\text{gain}_{\text{Smoking}} = 0.908 - 0.873$$

$$= \boxed{0.035} \quad (3)\ \text{gain}$$

Gain (Chest Pain) = 0.45 maximum Info Gain

Gain (male) = 0.248

Gain (smoking) = 0.035

PID	Male	Smoking	HA
1			
2			
3			

Chest pain

yes → NO

PId	Male	Smoking	HA
1	Yes	NO	Yes ✓
2	Yes	Yes	Yes ✓
5	NO ~~Yes~~	Yes	Yes ✓

Leaf Node

PId	male	Smoking	HA
3	NO	Yes	Yes
4	Yes	NO	NO
6	Yes	Yes	NO

can be split.
based on male & smoking

Info D = $I(1, 2) \Rightarrow -\Sigma P_i \log P_i$

$$\Rightarrow -\frac{1}{3}\log\left(\frac{1}{3}\right) - \frac{2}{3}\log\left(\frac{2}{3}\right) \Rightarrow 0.528 + 0.389$$

$$\Rightarrow \boxed{0.917}$$

for male → yes
→ no

$$gain_{(male)}(D) = Info(D) - Info_{(male)}$$

$$Info_{(male)} = \frac{2}{3} I(0, 2) + \frac{1}{3} I(1, 0)$$

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

$$0 + 0$$

$$= 0$$

for smoking → yes, no

$$gain_{smoking} = Info(D)$$

$$\therefore Info_{gain} = 0.917 - 0$$

$$= 0.917 \quad (1)$$

for Smoking → yes, no.

$$gain_{Smoking} = Info(D) - Info(smoking)$$

$$Info_{(smoking)} = \frac{2}{3} I(1, 1) + \frac{1}{3} I(0, 1)$$

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

$$= \frac{2}{3}(0 + 0.5) + \frac{1}{3}(0)$$

$$\frac{2}{3} = 0.666$$

$$\therefore Gain_{Smoking} = Info(D) - Info_{smoking}$$

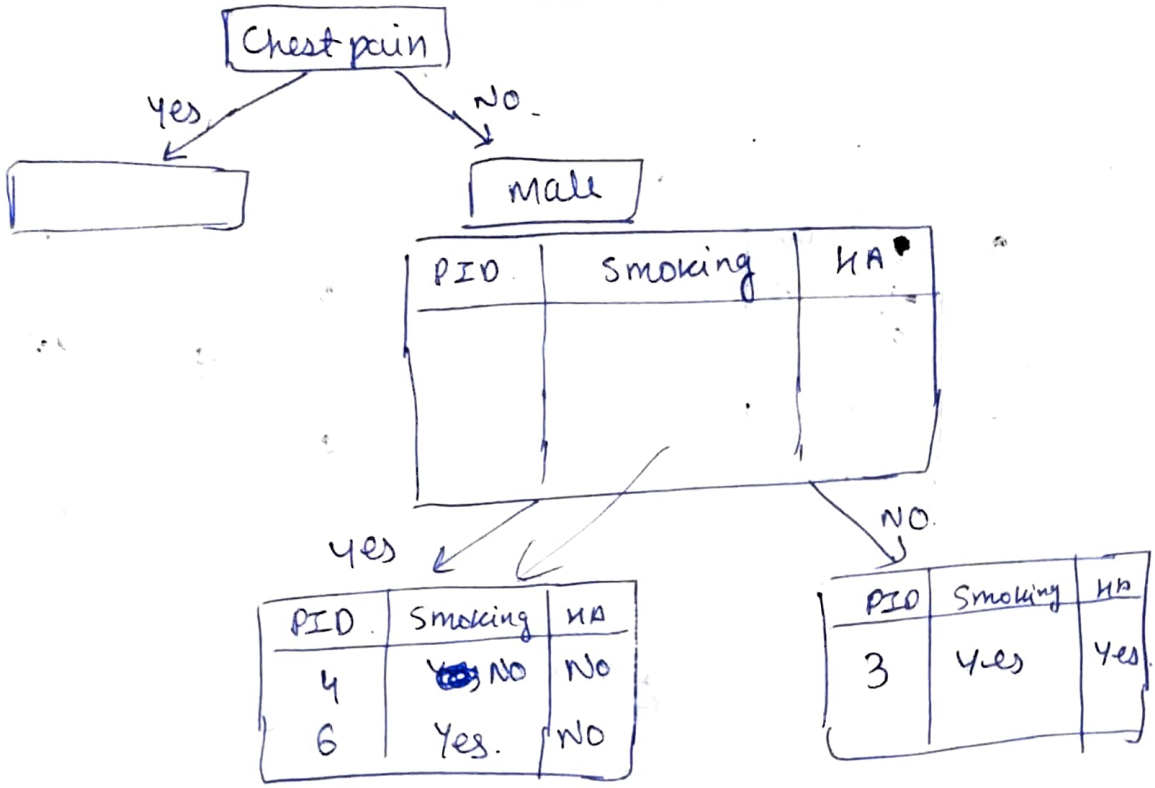
$$= 0.917 - 0.666$$

$$= 0.251 \quad (2)$$

$\therefore$ all gains

$gain_{male} = (0.917)$ Maximum

$gain_{smoking} = (0.251)$



Chest pain → No

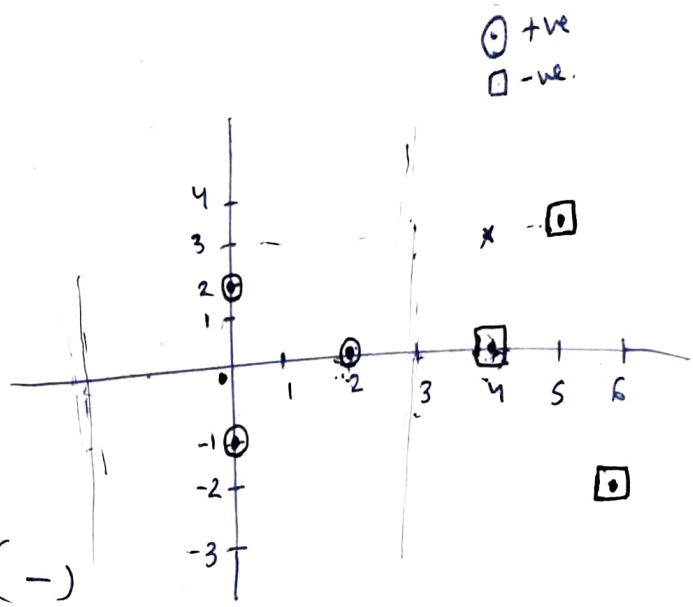
male → Yes

Smoking Yes.

∴ Heart Atk = NO

SVM

X	Y	T
0	2	+ve
0	-1	+ve
2	0	+ve
4	0	-ve
5	3	-ve
6	-2	-ve



Support vectors ⇒ $S_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ (+), $S_2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$ (−)

augmented Bias = 1 input value.

$\hat{S_1} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ $\hat{S_2} = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$

Equations of Support vector $\hat{S_1}$, $\hat{S_2}$

$$\alpha_1 \hat{S_1} \times \hat{S_1} + \alpha_2 \hat{S_2} \times \hat{S_1} = +1 \quad \text{—①}$$

$$\alpha_1 \hat{S_1} \times \hat{S_2} + \alpha_2 \hat{S_2} \times \hat{S_2} = -1 \quad \text{—②}$$

Solve ①&②

$$\alpha_1 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = +1 \quad \text{—①}$$

$$\alpha_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} = -1 \quad \text{—②}$$

$$\alpha_1 (4+0+1) + \alpha_2 (8+0+1) = 1 \rightarrow$$

$$\alpha_2 (8+0+1) + \alpha_2 (16+0+1) = -1$$

$$5\alpha_1 + 9\alpha_2 = 1 \quad \text{—①}$$

$$9\alpha_1 + 17\alpha_2 = -1 \quad \text{—②}$$

$$\alpha_1 = 6.5$$

$$\alpha_2 = -3.5$$

Weight vector = w

$$w = \sum_i \alpha_i s_i$$

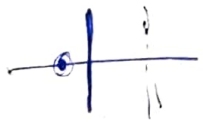
$$= \alpha_1 \hat{s}_1 + \alpha_2 \hat{s}_2$$

$$6.5 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + -3.5 \cdot \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$$

$$= 13 + 0 + 6.5 - 14 - 0 - 3.5$$

$$= 2$$

$$\begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \begin{pmatrix} 13 \\ 0 \\ 6.5 \end{pmatrix} + \begin{pmatrix} -14 \\ 0 \\ -3.5 \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \to x \\ 0 \to y \\ 3 \to \text{Bias} \end{pmatrix}$$



$y = wx + b$

$y = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} x + 3.$

$b = 3$

$b + 3 = 0$

$x = -3$

b

$x = 1, \; y = 0 \rightarrow$ IInd y axis

$x = 0 \quad y = 1 \rightarrow$ IInd to x axis

$x = 1 \quad y = 1 \rightarrow$ 45° between x & y

intersept → best line

$b = 3$

$b - 3 = 0$

$x = 3$

parallel to y axis taking x intersept.

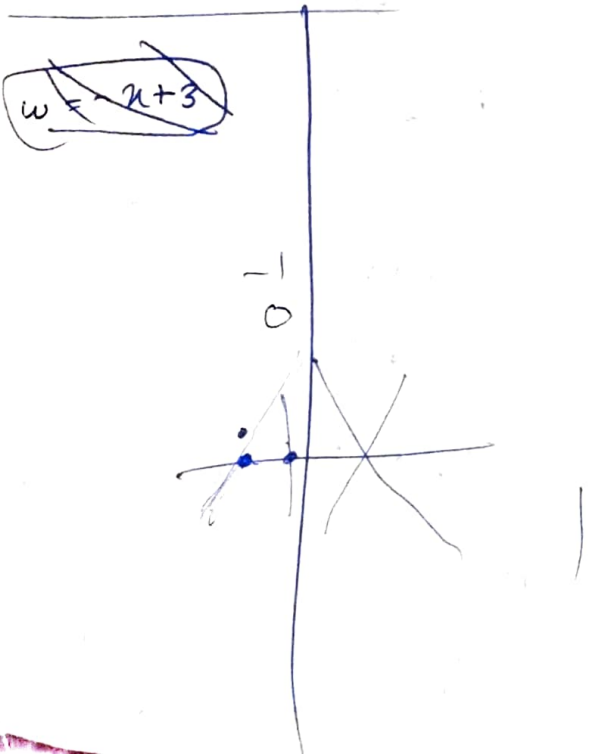
ex.

if $b = 2$

$x - 2 = 0$

$x = -2$

$w = -x + 3$



KNN

22 find the predict target class for the

(fever = NO cough = normal fatigue = Yes) (K=2)

Fever	cough	Fatigue	Disease
Yes	High	No	Yes
NO	High	Yes	Yes
Yes	Normal	Yes	Yes
NO	Normal	Yes	No
Yes	Normal	NO	NO

Fever → Yes (1), NO (0)

cough → High (2), Normal (1)

Fatigue → Yes (1), NO (0)

f=0 c=1 f=1

Euclidian dist, (d)

$$d = \sqrt{|x_0 - x_{a1}|^2 + |x_{02} - x_{a1}|^2}$$

	Fever	cough	Fatigue	Disease
1)	1	2	0	1
2)	0	2	1	1
3)	1	1	1	1
4)	0	1	1	0
5)	1	1	0	0

Rank

$d_1 = \sqrt{(0-1)^2+(1-2)^2+(1-0)^2} = \sqrt{3} = 1.732$ [5]

$d_2 = \sqrt{(0-0)^2+(1-2)^2+(1-1)^2} = 1$ [2]

$d_3 = \sqrt{(0-1)^2+(1-1)^2+(1-1)^2} = 1$ [3]

$d_4 = \sqrt{(0-0)^2+(1-1)^2+(1-1)^2} = 0$ [1]

$d_5 = \sqrt{(0-1)^2+(1-1)^2+(1-0)^2} = \sqrt{2} = 1.414$ [4]

Since. K=2, we select. d_4, d_2. ∴ sample 4, 2

	F	C	F	
4)	0	1	1	0
2)	0	2	1	1

voting 1 NO, 1 Yes

Can choose any one according to que.

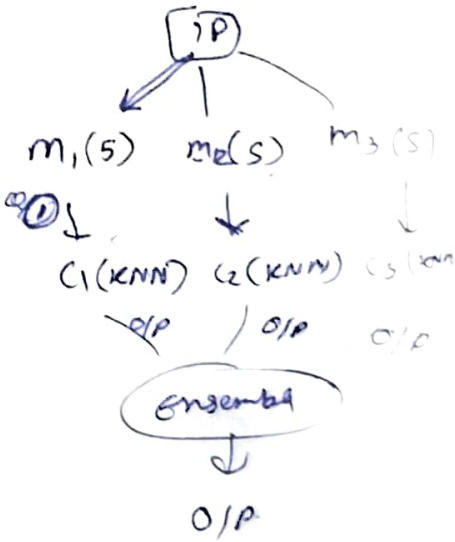
24 (KNN with bagging) $k = 2$

New Test Sample $X1 = 4.0$ $X2 = 4.5$ $X3 = 2.2$ find target.

Bootstrap Sample1 : (1, 2, 5, 6, 7)
Bootstrap sample2 : (3, 4, 5, 8, 10)
Bootstrap Sample 3 : (1, 3, 6, 7, 9)

Sample	F X1	F X2	F X3	Target
1	2.0	3.1	1.5	0
2	3.2	4.3	2.1	1
3	1.5	2.8	1.2	0
4	4.1	3.9	2.5	1
5	5.3	6.7	3.2	1
6	3.8	5.1	2.8	1
7	2.7	3.5	1.8	0
8	4.5	4.9	2.7	1
9	1.9	3.0	1.4	0
10	5.1	6.2	3.0	1

Bagging.
(Bootstrap aggregation)



$m_1(5)$ Bootstrap Sample1 (1, 2, 5, 6, 7)

1	2.0	3.1	1.5	0
2	3.2	4.3	2.1	1
5	5.3	6.7	3.2	1
6	3.8	5.1	2.8	1
7	2.7	3.5	1.8	0

Knn. $k=2$ $X_1 = 4.0$, $X_2 = 4.5$, $X_3 = 2.2$

		Rank
$d_1 = \sqrt{(4-2)^2 + (4.5-3.1)^2 + (2.2-1.5)^2}$	$= 2.539$	4
$d_2 = \sqrt{(4-3.2)^2 + (4.5-4.3)^2 + (2.2-2.1)^2}$	$= 0.830$	1 ✓
$d_5 = \sqrt{(4-5.3)^2 + (4.5-6.7)^2 + (2.2-3.2)^2}$	$= 2.744$	5
$d_6 = \sqrt{(4-3.8)^2 + (4.5-5.1)^2 + (2.2-2.8)^2}$	$= 0.871$	2 ✓
$d_7 = \sqrt{(4-2.7)^2 + (4.5-3.5)^2 + (2.2-1.8)^2}$	$= 1.688$	3

d_2, d_6 ∴ Sample 2, 6.

				Target
2	3.2	4.3	2.1	1
6	3.8	5.1	2.8	1

O/P = 1, 1

∴ By max voting O/P = 1

O/P $m_1 \rightarrow$ 1

$m_2(5)$ Bootstrap sample (3, 4, 5, 8, 10)

3	1.5	2.8	1.2	0
4	4.1	3.9	2.5	1
5	5.3	6.7	3.2	1
8	4.5	4.9	2.7	1
10	5.1	6.2	3.0	1

Knn $k=2$

$X_1 = 4.0$ $X_2 = 4.5$ $X_3 = 2.2$

		Rank
$d_3 = \sqrt{(4-1.5)^2 + (4.5-2.8)^2 + (2.2-1.2)^2}$	$= 3.184$	5
$d_4 = \sqrt{(4-4.1)^2 + (4.5-3.9)^2 + (2.2-2.5)^2}$	$= 0.678$	1 ✓
$d_5 = \sqrt{(4-5.3)^2 + (4.5-6.7)^2 + (2.2-3.2)^2}$	$= 2.744$	4
$d_8 = \sqrt{(4-4.5)^2 + (4.5-4.9)^2 + (2.2-2.7)^2}$	$= 0.812$	2 ✓
$d_{10} = \sqrt{(4-5.1)^2 + (4.5-6.2)^2 + (2.2-3)^2}$	$= 2.177$	3

d_4, d_8 ∴ sample 4, 8

		O/P
4	4.1	1
8	4.5	1

∴ O/P = 1, 1

By max voting O/P = 1

$m_2 \rightarrow$ 1

$m_3(S)$ Bootstrap Sample (1, 3, 6, 7, 9)

1	2.0	3.1	1.5	0
3	1.5	2.8	1.2	0
6	3.8	5.1	2.8	1 →
7	2.7	3.5	1.8	0 →
9	1.9	3.0	1.4	0

knn (k=2)

$X_1 = 4$

$X_2 = 4.5$

$X_3 = 2.2$

Rank

$\therefore d_1 = \sqrt{\quad} = 2.539$

$d_3 = \sqrt{\quad} = 3.184$

$d_6 = \sqrt{\quad} = 0.871$ [1] ← d_6

$d_7 \sqrt{\quad} = 1.688$ [2] ← d_7

$d_9 = \sqrt{(4-1.9)^2 + (4.5-3)^2 + (2.2-1.4)^2} = 2.701$

	O/P
6	1
7	0

⇒ any (1) we consider 0.

$\therefore$ m_1 O/P → 1

m_2 O/P → 1

m_3 O/P → 0

} Ensemble we get max voting

(1)

$\therefore$ for $X_1 = 4.0$ Target = (1)

$X_2 = 4.5$

$X_3 = 2.2$

To find impurity in Dataset.

Gini Index $= 1 - \sum P(x)^2$

	Weather	Accompanied parents	Money	Time Pass
1	Sunny	Yes	Rich	Cinema
2	Sunny	NO	Rich	Tennis
3	windy	Yes	Rich	Cinema
4	Rainy	Yes	Poor	Cinema
5	Rainy	NO	Rich	House
6	Rainy	Yes	Poor	Cinema
7	windy	NO	Poor	Cinema
8	windy	NO	Rich	Shopping
9	windy	Yes	Rich	Cinema
10	Sunny	NO	Rich	Tennis

Total 10 Cinema $= \frac{6}{10}$ House $= \frac{1}{10}$

Tennis $= \frac{2}{10}$ Shopping $= \frac{1}{10}$

Parent: yes, NO

$$1 - \left[\left(\frac{6}{10}\right)^2 + \left(\frac{2}{10}\right)^2 + \left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)^2\right]$$

$= 0.58$

Weather

- $\left(\frac{3}{10}\right)$ Sunny: cinema $= \frac{1}{3}$, Tennis $= \frac{2}{3}$
- Rainy $\left(\frac{3}{10}\right)$: cinema $= \frac{2}{3}$, House $= \frac{1}{3}$
- windy $\left(\frac{4}{10}\right)$: cinema $\frac{3}{4}$, shopping $\frac{1}{4}$

$1-\left[\left(\frac{1}{3}\right)^2+\left(\frac{2}{3}\right)^2\right]$ | $1-\left[\left(\frac{2}{3}\right)^2+\left(\frac{1}{3}\right)^2\right]$ | $1-\left[\left(\frac{3}{4}\right)^2+\left(\frac{1}{4}\right)^2\right]$

$= 0.44$ | $= 0.444$ | $= 0.375.$

Weighted avg $= \frac{3}{10}(0.444) + \frac{3}{10}(0.444) * \frac{4}{10}(0.375)$

$= \underline{0.416}$

Overall gini Index = 0.5

weather = **0.416**

Accompanied parent

- Yes $\left(\frac{5}{10}\right)$
- No $\left(\frac{5}{10}\right)$

Cinema. $= \frac{5}{5}$ | • cinema $= \frac{1}{5}$ • Tennis $= \frac{2}{5}$ • House $= \frac{1}{5}$ • Shopping $= \frac{1}{5}$

$\therefore 1-[1]$ | $1-\left[\left(\frac{1}{5}\right)^2+\left(\frac{2}{5}\right)^2+\left(\frac{1}{5}\right)^2+\left(\frac{1}{5}\right)^2\right]$

$= 0$ | $= 0.72$

weighted avg $= \frac{5}{10}\times 0 + \frac{5}{10}(0.72)$

Parent. = **0.36** (lowest)

Money.

- Rich $\left(\frac{7}{10}\right)$ ← cinema ⇒ $\frac{3}{7}$, house ⇒ $\frac{1}{7}$, Tennis ⇒ $\frac{2}{7}$, Shopping ⇒ $\frac{1}{7}$
- Poor $\left(\frac{3}{10}\right)$ ⇒ cinema $= \frac{3}{3}$

Poor: $1-[1]$ | Rich: $1-\left[\left(\frac{3}{7}\right)^2+\left(\frac{1}{7}\right)^2+\left(\frac{2}{7}\right)^2+\left(\frac{1}{7}\right)^2\right]$

$= 0.694$

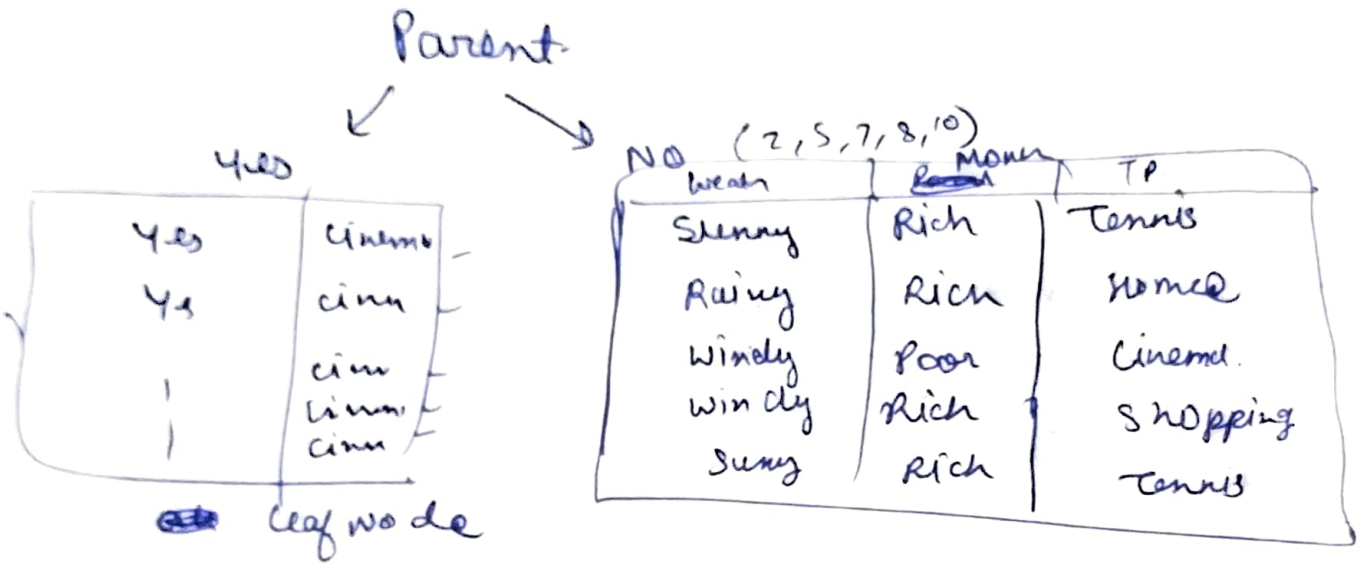
weighted Avg $= \frac{3}{10}(0) + \frac{7}{10}(0.694)$

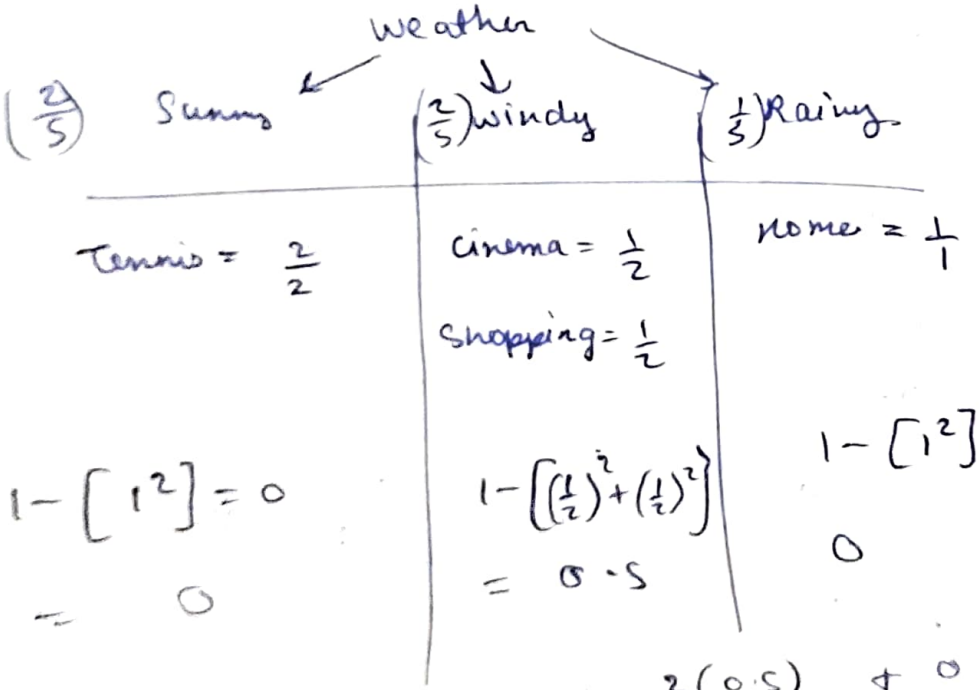
$= \boxed{0.485}$

Gini weather = 0.418

Gini parent = 0.360 (lowest) ∴ Root

Gini money = 0.485





Parent Money-

(4/5) Rich — (1/5) Poor.

- tennis = 2/4
- Hotel = 1/4
- shopping = 1/4

cinema 1/1

$1-\left[\left(\frac{2}{4}\right)^2+\left(\frac{1}{4}\right)^2+\left(\frac{1}{4}\right)^2\right]$

$= 0.625$

$1-\left[\frac{1}{1}^2\right]$

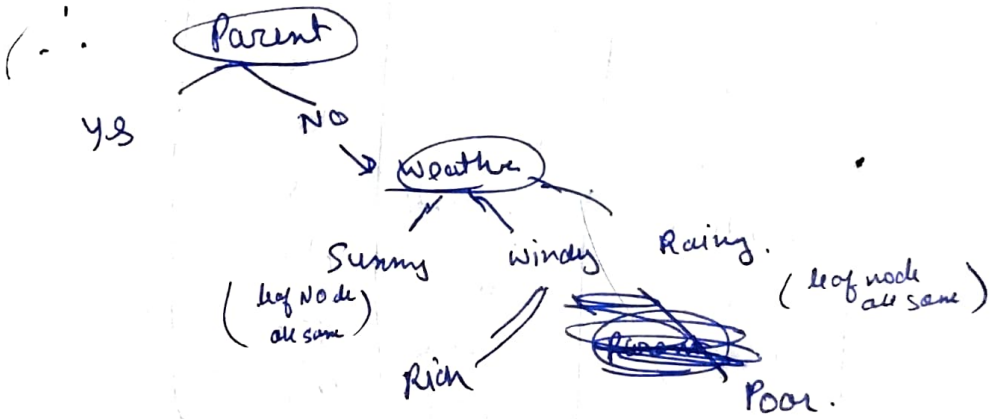
$= 0$

$\therefore WA = \frac{4}{5}(0.625) + \frac{1}{5} \times 0$

$= 0.5$

gini weather = 0.2 (lowest)

gini parent = 0.5



(Random Forest)

Person will buy a Product based on

Age (Young, Middle, Old)

Salary (Low, Med, High)

Online Ads clicked (Few, Moderate, Many)

Target → Yes, NO.

ID	Age	Salary	Clicks	Buy
1	Young	Low	Few	NO
2	Young	High	Many	Yes
3	Mid	Medium	Moderate	Yes
4	Old	Low	Few	NO
5	Old	Medium	Many	Yes.
6	Mid	High	Few	NO
7	Young	Medium	Moderate	Yes
8	Mid	Low	Few	NO
9	Old	High	Many	Yes.
10	Young	Low	Many	Yes

Age='old' Salary='High' Clicks='few' Buy=?

Tree1 sample {1, 3, 4, 5, 7, 9}

	Age	Salary	Click	
1	Young	Low	Few	NO
3	Mid	Medium	Moderate	Yes.
4	Old	Low	Few	NO
5	Old	Medium	Many	Yes.
7	Young	Medium	Moderate	Yes.
9	Old	High	Many	Yes

Total (6) Yes = $\frac{4}{6}$ NO = $\frac{2}{6}$

$$1 - \left[\left(\frac{4}{6}\right)^2 + \left(\frac{2}{6}\right)^2\right]$$

$$\Rightarrow 0.444$$

Age.

Young $(\frac{2}{6})$	mid $(\frac{1}{6})$	old $(\frac{3}{6})$
• Yes = $\frac{1}{2}$	• Yes = $\frac{1}{1}$	• Yes = $\frac{2}{3}$
• NO = $\frac{1}{2}$	• NO = $\frac{0}{1}$	• NO = $\frac{1}{3}$
$1 - \left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2\right]$	$1 - \left[\left(\frac{1}{1}\right)^2 + 0\right]$	$1 - \left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right]$
= 0.5	= 0	= 0.444

$$WA = \frac{2}{6}(0.5) + \frac{1}{6}(0) + \frac{3}{6}(0.44) = 0.386$$

Salary

low $(\frac{2}{6})$	mid $(\frac{3}{6})$	high $(\frac{1}{6})$
• Yes = 0	• Yes = $\frac{3}{3}$	• Yes = $\frac{1}{1}$
• NO = $\frac{2}{2}$	• NO = $\frac{0}{3}$	• NO = $\frac{0}{1}$
$1 - \left[\left(\frac{2}{2}\right)^2\right]$	$1 - \left[\left(\frac{3}{3}\right)^2 + 0\right]$	$1 - [1^2]$
= 0	= 0	= 0

WA = 0

Click

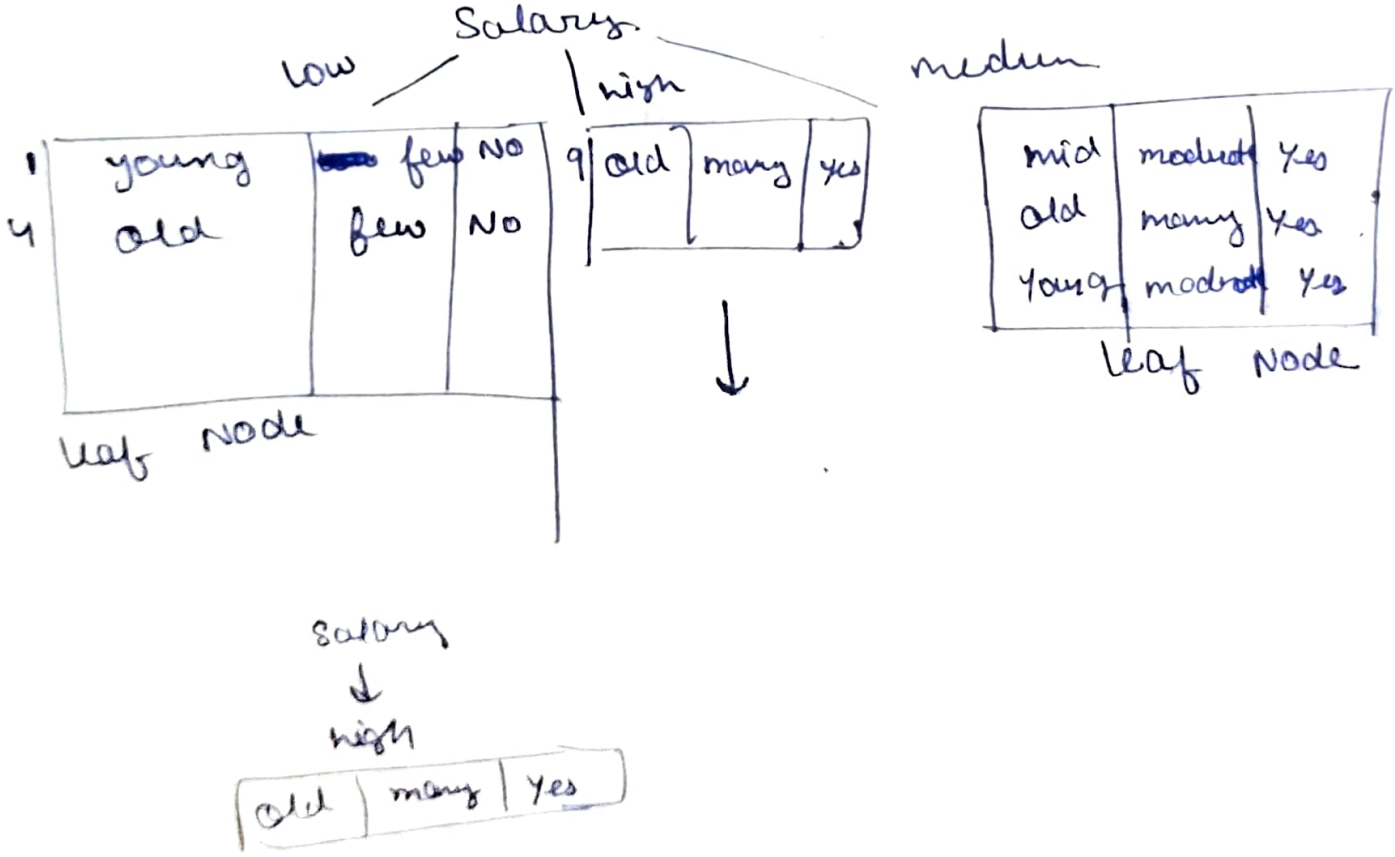
$\frac{2}{6}$ Few	$\frac{2}{6}$ Moderate	$\frac{2}{6}$ Many
• Yes → $\frac{0}{2}$	• Yes → $\frac{2}{2}$	• Yes $\left(\frac{2}{2}\right)$
• NO → $\left(\frac{2}{2}\right)$	• NO → $\frac{0}{2}$	• NO $\frac{0}{2}$
$1-\left[1^2\right]$	$1-\left[\left(\frac{2}{2}\right)^2\right]$	$1-\left[\left(\frac{2}{2}\right)^2\right]$
$= 0$	$= 0$	$= 0$

WA = 0

~~Salaries~~ gini (Age) = 0.386

gini (Salary) = 0 (lowest)

gini (Click) = 0.



Salary

↓

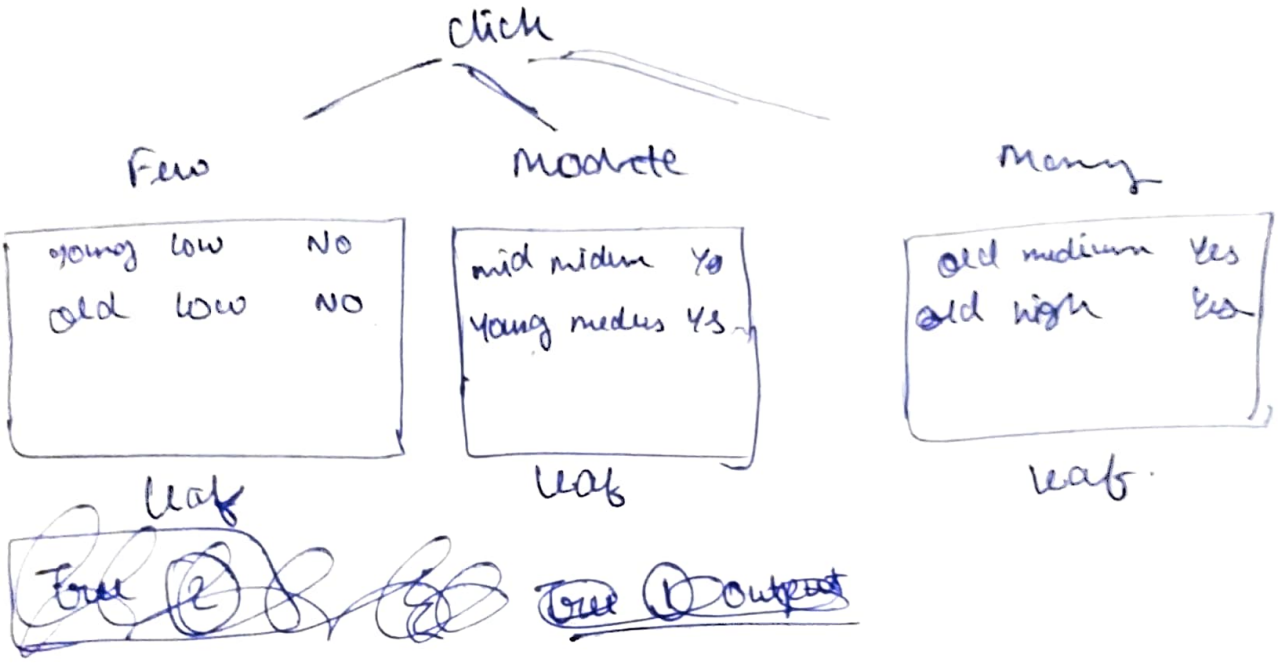
high

old	many	yes

gini (Age) = 0.386

gini (salary) = 0

gini (click) = 0



Tree 2

	Age	Salary	Click.
2	young	high	many	Yes
4	old	low	few	NO
6	middle	high	few	NO
8	middle	low	few	NO
9	old	high	many	Yes
10	young	low	many	Yes.

Gini (D) = $1 - \left[\left(\frac{3}{6}\right)^2 + \left(\frac{3}{6}\right)^2\right]$ (Yes: $\frac{3}{6}$ $\frac{1}{2}$ | NO: $\frac{3}{6}$ $\frac{1}{2}$)

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

$= \frac{1}{2} = 0.5$

gini (Age) ⇒

Age.

(2/6) young: Yes → 2/2 1, NO → 0/2, $1 - [1^2] = 0$ ⇒ $\frac{2}{6}(0)$

(2/6) old: Yes → 1/2, NO → 1/2, $1 - \left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2\right] = 0.5$ ⇒ $\frac{2}{6} \times 0.5$

(2/6) middle: Yes → 0/2, NO → 2/2 1, $1 - [1^2] = 0$ ⇒ $0 \times \frac{2}{6} = \frac{1}{6}$

gini (Age) = $\frac{1}{6}$ = 0.166

gini (Salary) = 0.44

Salary.

high $(\frac{3}{6})$	$(\frac{3}{6})$ low
yes → $\frac{2}{3}$	Yes → $\frac{1}{3}$
NO → $\frac{1}{3}$	NO → $\frac{2}{3}$
$1-\left[(\frac{2}{3})^2+(\frac{1}{3})^2\right]=0.44$	$1-\left[(\frac{1}{3})^2+(\frac{2}{3})^2\right]=0.44$

WA = $\frac{3}{6}(0.44)+\frac{3}{6}(0.44)$ =

0.22 + 0.22 = 0.44

gain (click)

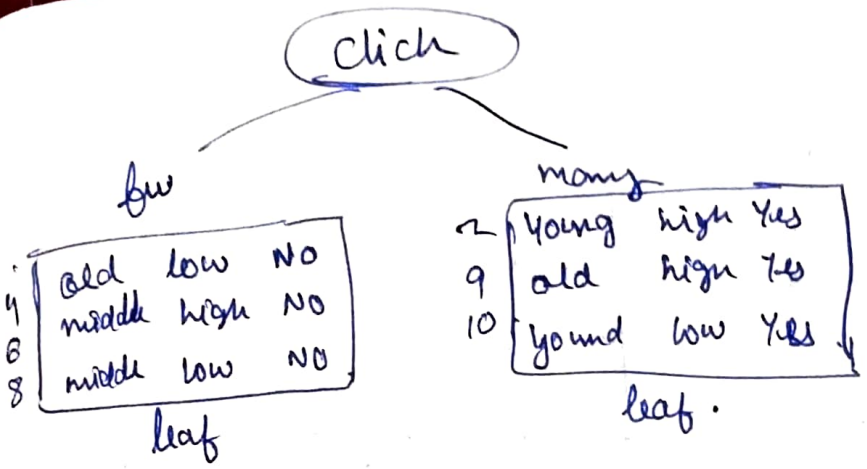
few $(\frac{3}{6})$	$(\frac{3}{6})$ many.
yes → $\frac{0}{3}$	Yes → $\frac{3}{3}$ 1
NO → $\frac{3}{3}$ 1	NO → $\frac{0}{3}$ 0
$1-[1]=0$	$1-[1]^2=0$

WA = $(\frac{3}{6})\times 0 + \frac{3}{6}\times 0 = 0$

gain click = 0. lowest.

gin (Age) = 0.166

gain (salary) = 0.44



True(3)

	Age	Salary	Click	
1	young	low	few	NO
3	middle	medium	Moderate	Yes
5	old	medium	many	Yes
6	middle	high	few	NO
7	young	medium	moderate	Yes
10	young	low	many	Yes

$$gini(D) = \frac{4}{6}(Yes) + \frac{2}{6}(NO)$$

$$1 - \left[\left(\frac{4}{6}\right)^2 + \left(\frac{2}{6}\right)^2\right]$$

$$\simeq \underline{0.444}$$

gini (Age)

$\left(\frac{3}{6}\right)$ young

Yes $\rightarrow \frac{2}{3}$

NO $\rightarrow \frac{1}{3}$

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

$\underline{0.444}$

$\left(\frac{2}{6}\right)$ middle

Yes $\rightarrow \frac{1}{2}$

NO $\rightarrow \frac{1}{2}$

$$1 - \left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2\right]$$

$= \underline{0.5}$

$\left(\frac{1}{6}\right)$ old

Yes $\rightarrow \frac{1}{1}$

NO $\rightarrow \frac{0}{1}$

$$1 - [1^2]$$

$= \underline{0}$

$$WA = \frac{3}{6}(0.44) + \frac{2}{6}(0.5) + 0. = \underline{0.386}$$

$$gini(Age) = \underline{0.386}$$

gini (salary) =

$(\frac{2}{6})$ low	$(\frac{3}{6})$ medium	$(\frac{1}{6})$ high
Yes → $\frac{1}{2}$	Yes → $\frac{3}{3}$	No → $\frac{1}{1}$
No → $\frac{1}{2}$	No → $\frac{0}{3}$	Yes → $\frac{0}{1}$
$1-\left[(\frac{1}{2})^2+(\frac{1}{2})^2\right]$	$1-[1^2]=0$	$1-[1^2]=0$
$=0.5=\frac{1}{2}$		

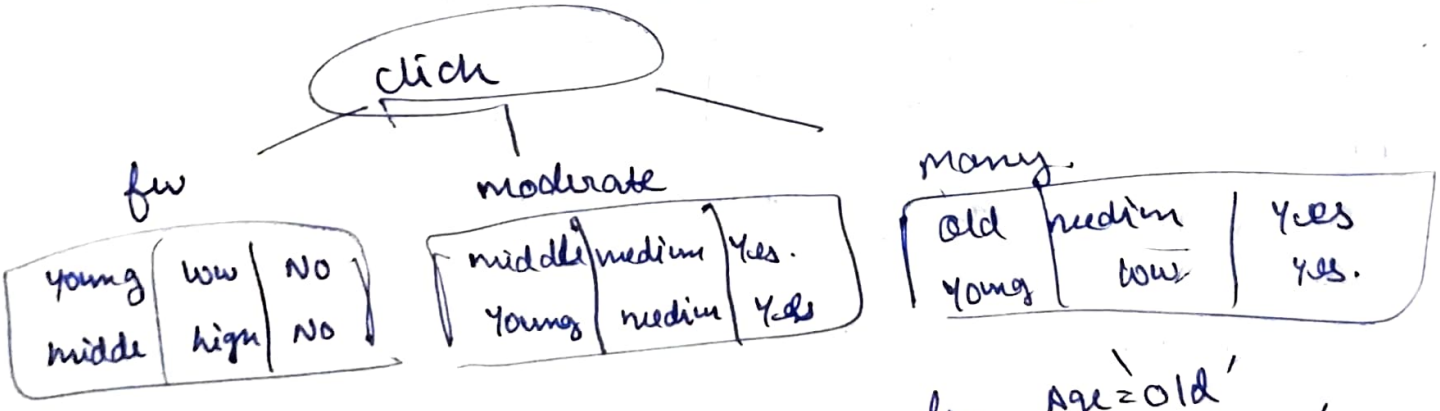
WA ⇒ $\frac{2}{6} \times \frac{1}{2} = \frac{1}{6} = 0.168$

gini(click)

$\frac{2}{6}$ few	$\frac{2}{6}$ moderate	$\frac{2}{6}$ many
Yes → $\frac{0}{2}$	Yes → $\frac{2}{2}$	Yes → $\frac{2}{2}$
No → $\frac{2}{2}$	No → $\frac{0}{2}$	No → $\frac{0}{2}$
0	0	0

WA = 0

gin (click = (lowest)



Tree ① → o/p → No

Tree ② → o/p → No

Tree ③ → o/p → No .

∴ for. Age='old'
Salary='High'
click='few'

Buy = No.