

Age	Competitive	Type	Profit
old	yes	Software	Down
old	No	Software	Down
old	No	Hardware	Down
med	Yes	Software	Down
med	yes	Hardware	Down
med	No	Hardware	up
med	No	Software	up
new	yes	Software	up
new	No	Hardware	up
new	No	Software	up

P - possibility of getting

up

(Down)

P = 5 (up)

N = 5 (down)

Entropy (class) :

$$I(P, N) = -\frac{P}{P+N} \log_2 \left(\frac{P}{P+N} \right) - \frac{N}{P+N} \log_2 \left(\frac{N}{P+N} \right)$$

Entropy (attributes) :

$$= \sum_{i=1}^n \frac{P_i + N_i}{P + N} \left(I(P_i, N_i) \right)$$

Gain

$$= \text{Entropy}_{\text{class}} - \text{Entropy}_{\text{attributes}}$$

Target f

P - up
N - down

$$P = 5 \quad N = 5$$

Entropy (class)

$$= \frac{-P}{P+N} \log_2 \left(\frac{P}{P+N} \right) - \frac{N}{P+N} \log_2 \left(\frac{N}{P+N} \right)$$

$$= 1$$

① Age : $\begin{matrix} \rightarrow P \\ \rightarrow N \end{matrix}$ Target

	up	down
old	0	3
mid	2	2
new	3	0
	P	N

$$P_1 = 0 \quad N_1 = 3$$

$$P_2 = 2 \quad N_2 = 2$$

$$P_3 = 3 \quad N_3 = 0$$

Entropy (age) :

$$H_1 = 0 \quad H_2 = 1 \quad H_3 = 0$$

$$0, 3 = 0$$

$$2, 2 = 1$$

$$3, 0 = 0$$

Entropy (attribute)

$$= \sum \frac{P_i \cdot N_i}{P+N} \log_2 \left(\frac{P_i \cdot N_i}{P+N} \right)$$

$$= \frac{0+3}{5+5} \log_2 \left(\frac{0+3}{5+5} \right) + \frac{2+2}{5+5} \log_2 \left(\frac{2+2}{5+5} \right) + \frac{3+0}{5+5} \log_2 \left(\frac{3+0}{5+5} \right)$$

$$= 4/16 = 0.25$$

$$\text{Gain} = 1 - \frac{2}{5} = \frac{5-2}{5} = \frac{3}{5}$$

$$\text{Gain (Age)} = \frac{3}{5}$$

$$= 0.6$$

② Competitive : — Target

	UP	Down
yes	1	3
No	4	2
	P	N

$$P_1 = 1 \quad N_1$$

$$P_2 = 4 \quad N_2$$

Entropy (Competitive) :

$$① = -\frac{P}{P+N} \log_2 \left(\frac{P}{P+N} \right) - \frac{N}{P+N} \log_2 \left(\frac{N}{P+N} \right)$$

$$= -\frac{1}{4} \log_2 \left(\frac{1}{4} \right) - \frac{3}{4} \log_2 \left(\frac{3}{4} \right)$$

$$= -0.25 * -2.0 - 0.75 * -0.41503$$

$$① = 0.8112$$

② Entropy

$$\begin{aligned} \textcircled{2} &= -\frac{4}{6} \log_{10} \left(\frac{4}{6} \right) - \frac{2}{6} \log_{10} \left(\frac{2}{6} \right) \\ &= -0.66 * -0.599 - 0.33 * -1.599 \\ &= 0.49301 \end{aligned}$$

Entropy (attributes) :

$$\begin{aligned} &\sum \frac{P_i + N_i}{P + N} I(P_i, N_i) \\ &= \frac{1+3}{10} * 0.8112 + \frac{4+2}{10} * 0.9930 \\ &= \frac{4}{10} * 0.8112 + \frac{6}{10} * 0.9930 \\ &= 0.87828 \end{aligned}$$

$$h_{\text{gain}} = 1 - 0.87828$$

$$h_{\text{gain}} = 0.12172$$

③ Type \rightarrow Target

	up	down
Software	3 3	3 3
Hardware	2	2

Entropy (Type)

$$(1) \Rightarrow \boxed{1}$$

$$(2) \Rightarrow \boxed{1}$$

Entropy (all kinds)

$$= \frac{\sum p_i + N_i}{P + N} I(p_i, N_i)$$

$$= \frac{3+3}{10} * (1) + \frac{2+2}{10} * (1)$$

$$= \frac{6}{10} + \frac{4}{10} =$$

$$\frac{10}{10}$$

gain

$$= 1 - 1 = 0$$

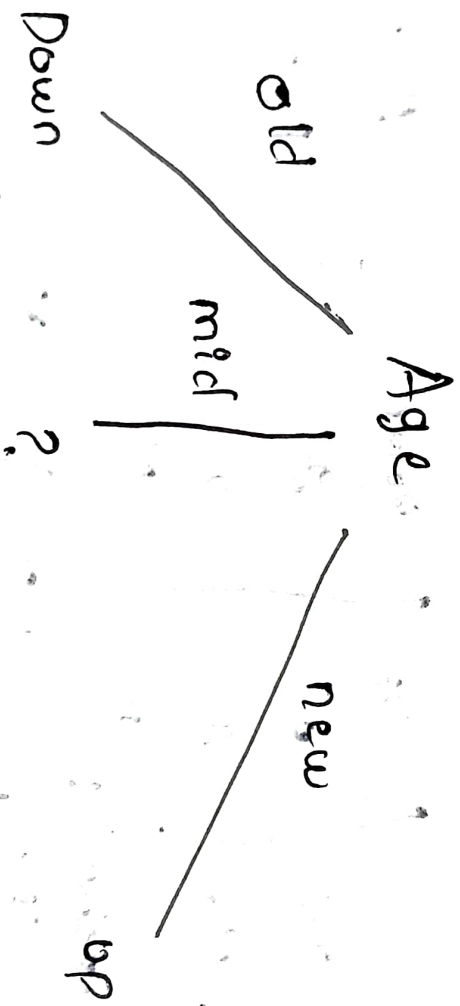
Crain

Age : 0.6

Competition : 0.12172

Type : 0

high \rightarrow 0.6
[Age]



mid	yes	Software	Down
mid	yes	Hardware	Down
mid	No	Hardware	up
mid	No	Software	up

Given cases

$$P = 2$$

$$N = 2$$

$$\text{Entropy (target)} = 1$$

① Competitive \rightarrow target

	Up	Down
yes	0	2
No	2	0

$$P_1 = 0$$

$$P_2 = 2$$

Entropy (competitive):

$$① = \frac{-P}{P+N} \log_2 \left(\frac{P}{P+N} \right) - \frac{N}{P+N} \log_2 \left(\frac{N}{P+N} \right)$$

$$① = 0$$

$$② = 0$$

Entropy (after rules):

$$= \frac{2}{4} * 0 + \frac{2}{4} * 0$$

$$= 0$$

$$\text{Gain} = 1 - 0$$

$$= 1$$

② Type \rightarrow Target

	UP	Down
Software	1	1
Hardware	1	1

$$P_1 = 1 \quad N_1 = 1$$

$$P_2 = 1 \quad N_2 = 1$$

Entropy (Type) :

$$(1) = 1$$

$$(2) = 1$$

Gain Entropy (Calculate)

$$= \frac{2^1}{4} * 1 + \frac{2^1}{4} * 1 = 1$$

$$\text{Gain} = 1 - 1$$

$$= 0$$

Crash ②

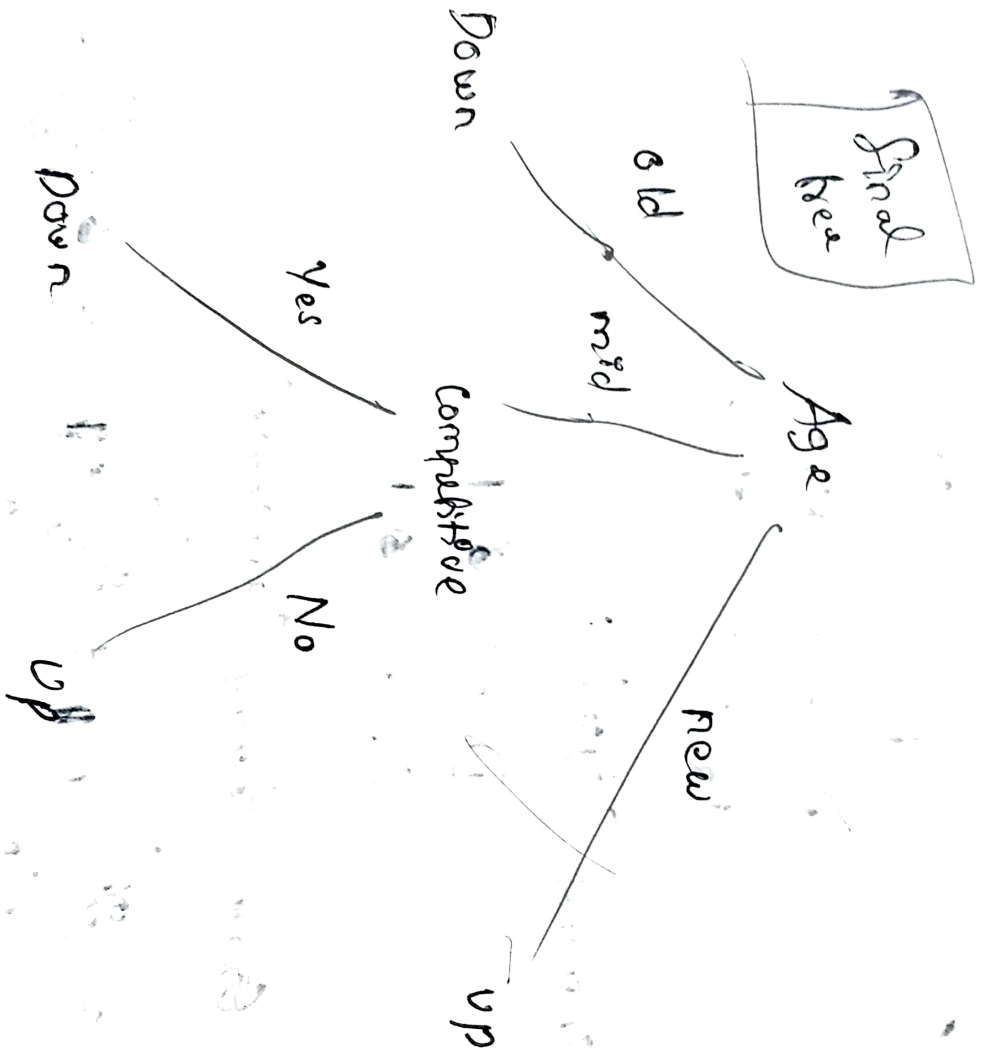
Cometive = 1

Type = 0

high

→

Cometive



we don't need type because

mid mid mid	Yes	Software	Down	3
	Yes	Hardware	Down	3
	No	Hardware	up	3
	No	Software	up	3

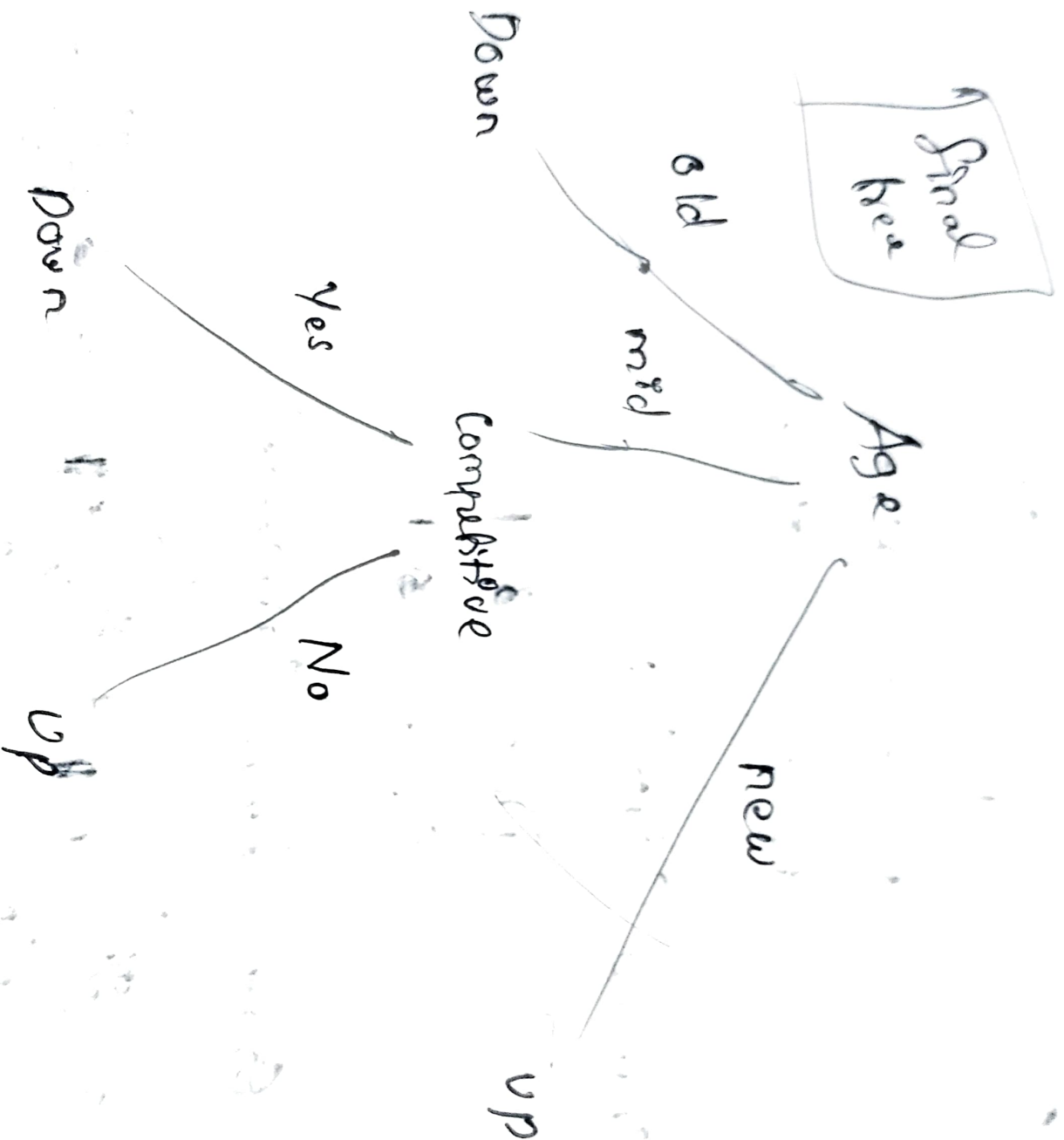
if Yes then Down
 if No then up

type

high



competition



we don't need type
because