

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31...40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

$$Info(D) = I(9,5) = -\frac{9}{14} \log_2 \left(\frac{9}{14} \right) - \frac{5}{14} \log_2 \left(\frac{5}{14} \right) = 0.409 - (-0.531) = 0.940$$

$$Info_{age}(D) = \frac{5}{14} I(2,3) + \frac{4}{14} I(4,0) + \frac{5}{14} I(3,2)$$

$$= \frac{5}{14} \left(-\frac{2}{3} \log_2 \left(\frac{2}{5} \right) - \frac{3}{5} \log_2 \left(\frac{3}{5} \right) \right) + \frac{4}{14} \left(-\frac{4}{4} \log_2 \left(\frac{4}{4} \right) - \frac{0}{4} \log_2 \left(\frac{0}{4} \right) \right) + \frac{5}{14} \left(-\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) \right)$$

$$= 0.347 + 0 + 0.347$$

$$= 0.694$$

$$Info_{income}(D) = \frac{4}{14} I(3,1) + \frac{6}{14} I(4,2) + \frac{4}{14} I(2,2)$$

$$= \frac{4}{14} \left(-\frac{3}{4} \log_2 \left(\frac{3}{4} \right) - \frac{1}{4} \log_2 \left(\frac{1}{4} \right) \right) + \frac{6}{14} \left(-\frac{4}{6} \log_2 \left(\frac{4}{6} \right) - \frac{2}{6} \log_2 \left(\frac{2}{6} \right) \right) + \frac{4}{14} \left(-\frac{2}{4} \log_2 \left(\frac{2}{4} \right) - \frac{2}{4} \log_2 \left(\frac{2}{4} \right) \right)$$

$$= 0.232 + 0.393 + 0.286$$

$$= 0.911$$

$$Info_{student}(D) = \frac{7}{14} I(6,1) + \frac{7}{14} I(3,4)$$

$$= \frac{7}{14} \left(-\frac{6}{7} \log_2 \left(\frac{6}{7} \right) - \frac{1}{7} \log_2 \left(\frac{1}{7} \right) \right) + \frac{7}{14} \left(-\frac{3}{7} \log_2 \left(\frac{3}{7} \right) - \frac{4}{7} \log_2 \left(\frac{4}{7} \right) \right)$$

$$= 0.296 + 0.493$$

$$= 0.789$$

$$Info_{credit_rating}(D) = \frac{6}{14} I(3,3) + \frac{8}{14} I(6,2)$$

$$= \frac{6}{14} \left(-\frac{3}{6} \log_2 \left(\frac{3}{6} \right) - \frac{3}{6} \log_2 \left(\frac{3}{6} \right) \right) + \frac{8}{14} \left(-\frac{6}{8} \log_2 \left(\frac{6}{8} \right) - \frac{2}{8} \log_2 \left(\frac{2}{8} \right) \right)$$

$$= 0.429 + 0.463$$

$$= 0.892$$

వారిచ్చు సమితి

663020283-2

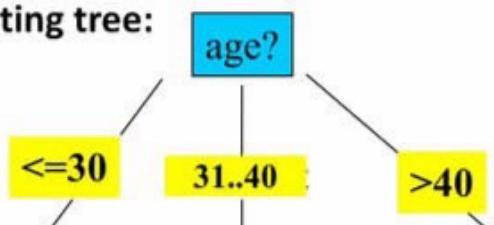
$$Gain(age) = Info(D) - Info_{age}(D) = 0.940 - 0.694 = 0.246$$

$$Gain(income) = Info(D) - Info_{income}(D) = 0.940 - 0.911 = 0.029$$

$$Gain(student) = Info(D) - Info_{student}(D) = 0.940 - 0.789 = 0.151$$

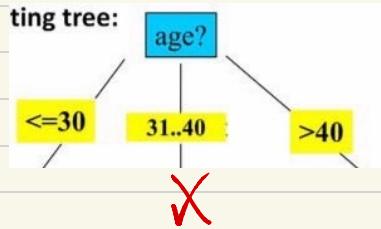
$$Gain(credit_rating) = Info(D) - Info_{credit_rating}(D) = 0.940 - 0.892 = 0.048$$

ting tree:



Age > student > credit_rating > income

age	income	student	credit rating	buys computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31..40	high	no	fair	yes
31..40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
<=30	medium	yes	excellent	yes
31..40	medium	no	excellent	yes
31..40	high	yes	fair	yes



Age			
Buy	≤ 30	$31-40$	> 40
yes	2	4	3
no	3	0	2

X

$$\text{Info}(D) = I(5,5) = -\frac{5}{10} \log_2\left(\frac{5}{10}\right) - \frac{5}{10} \log_2\left(\frac{5}{10}\right) = 1$$

$$\text{Info}_{\leq 30}(D) = \frac{2}{5} I(2,0) + \frac{3}{5} I(0,3)$$

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

$$= 0+0 = 0$$

$$\text{Info}_{> 40}(D) = \frac{3}{5} I(2,1) + \frac{2}{5} I(1,1)$$

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

$$= 0.551 + 0.4$$

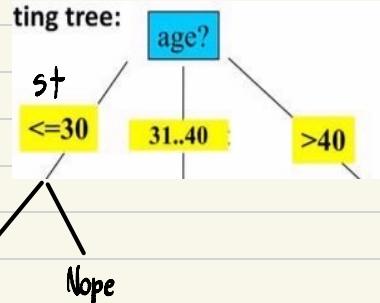
$$= 0.951$$

$$\text{Gain}(\leq 30) = \text{Info}(D) - \text{Info}_{\leq 30}(D) = 1 - 0 = 1$$

$$\text{Gain}(> 40) = \text{Info}(D) - \text{Info}_{> 40}(D) = 1 - 0.951 = 0.049$$

then ≤ 30

Buy	Student	
	yes	no
yes	3	2
no	0	3



మార్కెట్ లో వున్నాయి

663020283-2

>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
>40	medium	yes	fair	yes
>40	medium	no	excellent	no

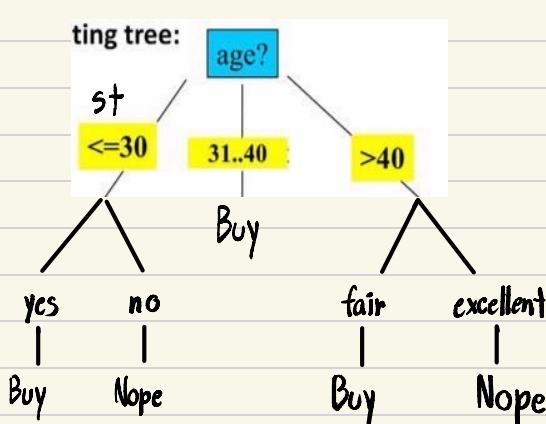
$$Info(D) = I(3,2) = -\frac{3}{5} \log_2 \left(\frac{3}{5}\right) - \frac{2}{5} \log_2 \left(\frac{2}{5}\right) = 0.442 - (-0.529) = 0.971$$

$$\begin{aligned} Info_{income}(D) &= \frac{3}{5} I(2,1) + \frac{2}{5} I(1,1) \\ &= \frac{3}{5} \left(-\frac{2}{3} \log_2 \left(\frac{2}{3}\right) - \frac{1}{3} \log_2 \left(\frac{1}{3}\right) \right) + \frac{2}{5} \left(-\frac{1}{2} \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} \log_2 \left(\frac{1}{2}\right) \right) \\ &= 0.551 + 0.4 \\ &= 0.951 \end{aligned}$$

$$\begin{aligned} Info_{credit_rating}(D) &= \frac{2}{5} I(0,2) + \frac{3}{5} I(3,0) \\ &= \frac{2}{5} \left(-\frac{0}{2} \log_2 \left(\frac{0}{2}\right) - \frac{2}{2} \log_2 \left(\frac{2}{2}\right) \right) + \frac{3}{5} \left(-\frac{2}{3} \log_2 \left(\frac{2}{3}\right) - \frac{0}{3} \log_2 \left(\frac{0}{3}\right) \right) \\ &= 0 + 0 = 0 \end{aligned}$$

$$\begin{aligned} Gain(income) &= Info(D) - Info_{income}(D) = 0.971 - 0.951 = 0.020 \\ Gain(credit_rating) &= Info(D) - Info_{credit_rating}(D) = 0.971 - 0 = 0.971 \end{aligned}$$

non credit_rating



Age > 40	
G-B	
ex	fair
yes	0
no	2

వారికు లేదని

663020283-2