

① Age

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

$$\text{Info}_{\text{age}} = 0.694$$

$$\text{Gain}_{\text{age}} = 0.940 - 0.694 = 0.246$$

② Income

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

root node (income)

	L.	M.	H.
Yes	3	4	2
No.	1	2	2

$$\text{Info}_{\text{income}}(D) = \frac{4}{14} I(3,1) + \frac{6}{14} I(4,2) + \frac{4}{14} I(2,2) \approx 0.911$$

$$\text{Gain}_{\text{income}} = 0.940 - 0.911 = 0.029$$

③ Student

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

root node (student)

	std.	Not std.
Yes	6	3
No.	1	4

$$\text{Info}_{\text{std.}}(D) = \frac{7}{14} I(6,1) + \frac{7}{14} I(3,4) \approx 0.789$$

$$\text{Gain}_{\text{income}} = 0.940 - 0.789 = 0.151$$

③ Student

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

root node (credit)

	fair	excellent
Yes	6	3
No.	2	3

$$\text{Info}_{\text{Credit}}(D) = \frac{8}{14} I(6,2) + \frac{6}{14} I(3,3) \approx 0.892$$

$$\text{Gain}_{\text{income}} = 0.940 - 0.892 = 0.048$$

② student

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
<=30	medium	yes	excellent	yes

$$\text{Info}(D) = I(2,3) = -\frac{2}{5} \log_2\left(\frac{2}{5}\right) - \frac{3}{5} \log_2\left(\frac{3}{5}\right) \approx 0.970$$

$$\text{Info}_{\text{income}}(D) = \frac{1}{5} I(1,0) + \frac{2}{5} I(1,1) + \frac{2}{5} I(0,2) \approx 0.8$$

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

$$\text{Info}_{\text{credit}}(D) = \frac{3}{5} I(1,2) + \frac{2}{5} I(1) \approx 0.951$$

$$\text{Gain}(\text{income}) = 0.970 - 0.8 = 0.170$$

$$\text{Gain}(\text{std}) = 0.970 - 0 = 0.970$$

$$\text{Gain}(\text{credit}) = 0.970 - 0.951 = 0.019$$

$\therefore \text{study} > \text{income} > \text{credit_rating}$

① Credit or Income

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

1.1) Income

	Low	Medium	High
Yes	1	2	0
No	1	1	0

1.2) Credit

	fair	excellent
Yes	3	0
No	0	2

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

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

$Info_{income}(D) = \frac{3}{5} I(1,1) + \frac{2}{5} I(2,1) + \frac{0}{5} I(0,0) = 0.951$

$Gain_{income} = \overset{Info(D)}{0.971} - \overset{Info_{income}(D)}{0.951} = 0.020$

$Gain_{credit} = \overset{Info(D)}{0.971} - \overset{Info_{credit}(D)}{0} = 0.971 \checkmark$

$\therefore Credit > income$ เลือก Credit

