

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

$$y = 8 \Rightarrow n = 4$$

$$\text{Info}(D) = I(3,4) = -\frac{3}{12} \log_2 \left(\frac{3}{12}\right) - \frac{4}{12} \log_2 \left(\frac{4}{12}\right) = \log_2 3 - \frac{2}{3} \approx 0.9183$$

age	p_i	n_i	$I(p_i, n_i)$
<=30	2	2	1
31...40	3	0	0
>40	3	2	0.9710

student	p_i	n_i	$I(p_i, n_i)$
Yes	5	1	0.65
No	3	3	1

$$\text{Info}_{age}(D) = I(2,2) = -\frac{3}{4} \log_2 \left(\frac{3}{4}\right) - \frac{2}{4} \log_2 \left(\frac{2}{4}\right) = 1$$

$$\text{Info}_{age}(D) = I(3,0) = -\frac{3}{3} \log_2 \left(\frac{3}{3}\right) - \frac{0}{3} \log_2 \left(\frac{0}{3}\right) = 0$$

$$\text{Info}_{age}(D) = I(3,2) = -\frac{3}{5} \log_2 \left(\frac{3}{5}\right) - \frac{2}{5} \log_2 \left(\frac{2}{5}\right) = 0.9710$$

$$\text{Info}_{age}(D) = \frac{4}{12} I(2,2) + \frac{3}{12} I(3,0) + \frac{5}{12} I(3,2)$$

$$\text{Info}_{age}(D) = \frac{4}{12} (1) + \frac{3}{12} (0) + \frac{5}{12} (0.9710) = 0.9399$$

$$\text{Gain}_{age} = \text{Info}(D) - \text{Info}_{age}(D) = 0.9183 - 0.9399 = -0.0204$$

income	p_i	n_i	$I(p_i, n_i)$
high	2	2	1
medium	4	1	0.7219
low	2	1	0.9713

$$\text{Info}(D) = I(2,2) = -\frac{3}{4} \log_2 \left(\frac{3}{4}\right) - \frac{2}{4} \log_2 \left(\frac{2}{4}\right) = 1$$

$$\text{Info}(D) = I(4,1) = -\frac{4}{5} \log_2 \left(\frac{4}{5}\right) - \frac{1}{5} \log_2 \left(\frac{1}{5}\right) = 0.7219$$

$$\text{Info}(D) = I(2,1) = -\frac{2}{3} \log_2 \left(\frac{2}{3}\right) - \frac{1}{3} \log_2 \left(\frac{1}{3}\right) = 0.9713$$

$$\text{Info}_{income}(D) = \frac{4}{12} I(2,2) + \frac{3}{12} I(4,1) + \frac{5}{12} I(2,1)$$

$$\text{Info}_{income}(D) = \frac{4}{12} (1) + \frac{3}{12} (0.7219) + \frac{5}{12} (0.9713) = 0.8637$$

$$\text{Gain}_{income} = \text{Info}(D) - \text{Info}_{income}(D) = 0.9183 - 0.8637 = 0.0546$$

$$\text{Gain}_{age} = 0.1804$$

$$\text{Gain}_{income} = 0.0546$$

$$\text{Gain}_{student} = 0.0933$$

$$\text{Gain}_{credit_rating} = 0.1686$$

$$\text{Info}(D) = I(5,1) = -\frac{5}{6} \log_2 \left(\frac{5}{6}\right) - \frac{1}{6} \log_2 \left(\frac{1}{6}\right) = 0.65$$

$$\text{Info}(D) = I(3,3) = -\frac{3}{6} \log_2 \left(\frac{3}{6}\right) - \frac{3}{6} \log_2 \left(\frac{3}{6}\right) = 1$$

$$\text{Info}_{student}(D) = \frac{1}{12} I(5,1) + \frac{1}{12} I(3,3)$$

$$\text{Info}_{student}(D) = \frac{1}{12} (0.65) + \frac{1}{12} (1) = 0.885$$

$$\text{Gain}_{student} = \text{Info}(D) - \text{Info}_{student}(D) = 0.9183 - 0.885 = 0.0353$$

credit_rating	p_i	n_i	$I(p_i, n_i)$
fair	6	1	0.5917
excellent	2	3	0.9710

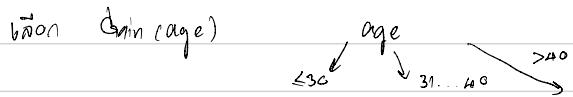
$$\text{Info}(D) = I(6,1) = -\frac{3}{7} \log_2 \left(\frac{3}{7}\right) - \frac{2}{7} \log_2 \left(\frac{2}{7}\right) = 0.5917$$

$$\text{Info}(D) = I(2,3) = -\frac{2}{5} \log_2 \left(\frac{2}{5}\right) - \frac{3}{5} \log_2 \left(\frac{3}{5}\right) = 0.9710$$

$$\text{Info}_{credit_rating}(D) = \frac{7}{12} I(6,1) + \frac{5}{12} I(2,3)$$

$$\text{Info}_{credit_rating}(D) = \frac{7}{12} (0.5917) + \frac{5}{12} (0.9710) = 0.7407$$

$$\text{Gain}_{credit_rating} = \text{Info}(D) - \text{Info}_{credit_rating}(D) = 0.9183 - 0.7407 = 0.1786$$



x_1	y	x_1	y	x_1	y
≤ 30	no	31...40	yes	> 40	yes
≤ 30	no	31...40	yes	> 40	yes
≤ 30	yes	31...40	yes	> 40	yes
≤ 30	yes			> 40	no

50%

income	p_i	n_i	$I(p_i, n_i)$
high	0	2	0
medium	1	0	0
low	1	0	0

$$I(0,2) = -\frac{1}{2} \log_2 \left(\frac{1}{2}\right) - \frac{1}{2} \log_2 \left(\frac{1}{2}\right) = 0$$

$$I(1,0) = -\frac{1}{1} \log_2 \left(\frac{1}{1}\right) - \frac{0}{1} \log_2 \left(\frac{0}{1}\right) = 0$$

$$I(1,0) = -\frac{1}{1} \log_2 \left(\frac{1}{1}\right) - \frac{0}{1} \log_2 \left(\frac{0}{1}\right) = 0$$

$$\text{Info}_{\text{income}}(D_1 \leq 30) = \frac{3}{4} I(0,2) + \frac{1}{4} I(1,0) + \frac{1}{4} (1,0)$$

$$= \frac{3}{4}(0) + \frac{1}{4}(0) + \frac{1}{4}(0)$$

$$\text{Info}_{\text{income}}(D_1 \leq 30) = 0$$

$$\text{Gain}(\text{income}) = \text{Info}(D_1 \leq 30) - \text{Info}_{\text{income}}(D_1 \leq 30)$$

$$= 1 - 0$$

$$\text{Gain}(\text{income}) = 1$$

$$\text{Gain}(\text{student}) = 1$$

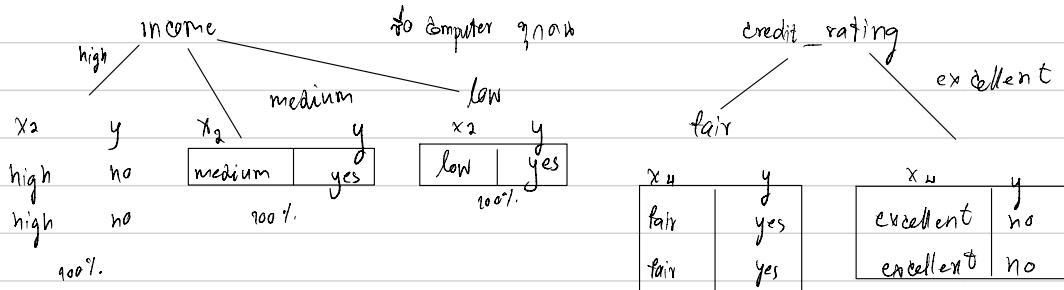
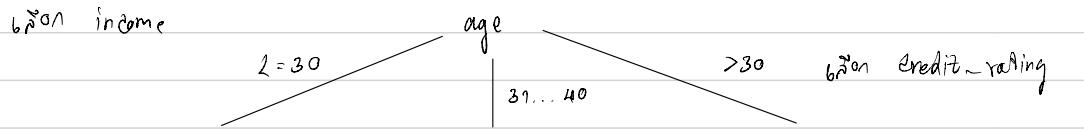
$$\text{Gain}(\text{credit_rating}) = 0$$

$$\text{Gain}(\text{income}) = \text{Info}(D_1 > 40) - \text{Info}_{\text{income}}(D_1 > 40)$$

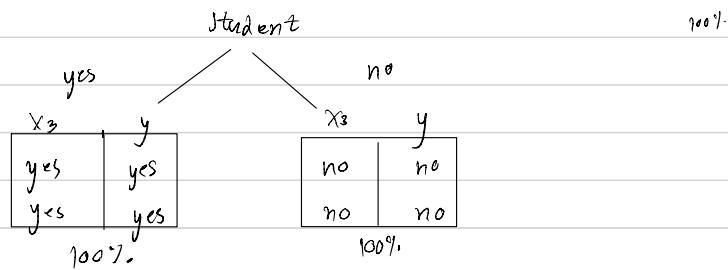
$$= 0.991 - 0.9183$$

$$\text{Gain}(\text{income}) = 0.0527$$

$$\text{Gain}(\text{credit_rating}) = \text{Info}(D_1 > 40) - \text{Info}_{\text{credit_rating}}(D_1 > 40) = 0.991 - 0 = 0.991$$



branch student



100%.

