

Homework

Gain (age)

Age		
≤ 30	$31 \dots 40$	> 40
N		Y
N	Y	Y
N	Y	
Y		Y
Y		N

$$\text{Info}(D) = I(9,5)$$

$$= -\frac{9}{14} \log_2 \frac{9}{14} - \frac{5}{14} \log_2 \frac{5}{14}$$

$$= 0.940 *$$

Gain (age)

age	P_i	n_i	$I(pini)$
≤ 30	2	3	0.971
$31 \dots 40$	4	0	0
> 40	3	2	0.971

$$\text{Info}_{\text{age}}(D) = \frac{5}{14} I(2,3) + \frac{4}{14} I(4,0) + \frac{5}{14} I(3,2) = 0.694 *$$

$$\text{Gain (age)} = 0.940 - 0.694 = 0.246 *$$

Gain (income)

income	P_i	n_i	$I(pini)$
≤ 30	2	2	1
$31 \dots 40$	4	2	0.918
> 40	3	1	0.811

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

$$\text{Gain (income)} = 0.940 - 0.911 = 0.029 *$$

Gain (student)

Student	P _i	n _i	I(p _i n _i)
yes	6	1	0.592
no	3	4	0.985

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

$$\text{Gain (student)} = 0.940 - 0.789 = 0.151 *$$

Gain (credit_rating)

credit_rating	P _i	n _i	I(p _i n _i)
yes	6	2	0.811
no	3	3	1

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

$$\text{Gain (credit_rating)} = 0.940 - 0.892 = 0.048 *$$

$$\text{Gain (age)} = 0.246$$

$$\text{Gain (income)} = 0.024$$

$$\text{Gain (student)} = 0.151$$

$$\text{Gain (credit_rating)} = 0.048$$

ກົດຕົວມີຄວາມສັບສົນຂອງ Gain (age) ຕັ້ງ root mode

U1 age ≤ 30

$$\text{Info}(D) = I(2,3) = 0.971 *$$

Info_{income}(D) von age (≤ 30)

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

$$\begin{aligned}\text{Info}_{\text{income}}(D) \text{ von age } (\leq 30) &= \frac{2}{5} I(0,2) + \frac{2}{5} I(1,1) + \frac{1}{5} I(1,0) \\ &= 0.4 *\end{aligned}$$

$$\text{Gain(income) von age } (\leq 30) = 0.971 - 0.4 = 0.571 *$$

Info_{Student}(D) von age (≤ 30)

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

ja → ja (buy_computer) → 75%

nein → nein (buy_computer) → 25%

age (> 40)

$$\text{Info } (D) = I(3,2) = 0.971 *$$

Info income (D) vs age (> 40)

income	p_i	n_i	$I(p_i n_i)$
low	1	1	1
medium	2	1	0.918

$$\text{Info}_{\text{income}} (D) \text{ vs } \text{age} (> 40) = \frac{2}{5} I(1,1) + \frac{3}{5} I(2,1) = 0.951 *$$

$$\text{Gain}(\text{income}) \text{ vs } \text{age} (> 40) = 0.971 - 0.951 = 0.020 *$$

Info student (D) vs age (> 40)

student	p_i	n_i	$I(p_i n_i)$
yes	2	1	0.916
no	1	1	1

$$\text{Info}_{\text{student}} (D) \text{ vs } \text{age} (> 40) = \frac{3}{5} I(2,1) + \frac{2}{5} I(1,1) = 0.951 *$$

$$\text{Gain}(\text{student}) \text{ vs } \text{age} (> 40) = 0.971 - 0.951 = 0.020 *$$

Info credit-rating (D) vs age (> 40)

$$\text{Info}_{\text{credit-rating}} (D) \text{ vs } \text{age} (> 40) = \frac{3}{5} I(3,0) + \frac{2}{5} I(0,2)$$

9:10 fair → yes ≈ 80

excellent → no ≈ 30

