

Decision-Tree

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

2.1) $Info_{age}(D) = \frac{5}{14} I(2,3) + \frac{4}{14} I(4,0) + \frac{5}{14} I(3,2) = 0.694$

2.2) $Info_{income}(D) = \frac{4}{14} I(2,2) + \frac{6}{14} I(4,2) + \frac{4}{14} I(3,1) = 0.911$

2.3) $Info_{student}(D) = \frac{7}{14} I(6,1) + \frac{7}{14} I(3,4) = 0.789$

2.4) $Info_{credit_rating}(D) = \frac{8}{14} I(6,2) + \frac{6}{14} I(3,3) = 0.892$

3) คำนวณ Information Gain โดยนำ Grain ที่คำนวณได้ root node

3.1) $Grain(age) = 0.940 - 0.694 = 0.246$ มีค่ามากที่สุด ~~✗~~

3.2) $Grain(income) = 0.940 - 0.911 = 0.029$

3.3) $Grain(student) = 0.940 - 0.789 = 0.151$

3.4) $Grain(credit_rating) = 0.940 - 0.892 = 0.048$

4. เลือกค่า feature ที่มีค่าใน root

4.1) $<= 30$

$Info(D) = I(2,3) = 0.971$

$Info_{income}(D) = \frac{2}{5} I(0,2) + \frac{2}{5} I(1,1) + \frac{1}{5} I(1,0) = 0.4$

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

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

คำนวณ Information Gain

$$\text{Grain (Income)} = 0.971 - 0.4 = 0.571$$

$$\text{Grain (Student)} = 0.971 - 0 = 0.971 \quad \text{มีค่ามากที่สุด} \quad \#$$

$$\text{Grain (Credit_rating)} = 0.971 - 0.951 = 0.02$$

4.2) 31...40

$$\text{มี yes} = 4 \quad \text{no} = 0$$

สามารถถาม buy ได้เลย

4.3) >40

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

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

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

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

คำนวณ Information Gain

$$\text{Grain (Income)} = 0.971 - 0.951 = 0.02$$

$$\text{Grain (Income)} = 0.971 - 0.951 = 0.02$$

$$\text{Grain (Credit_rating)} = 0.971 - 0 = 0.971 \quad \text{มีค่ามากที่สุด} \quad \#$$