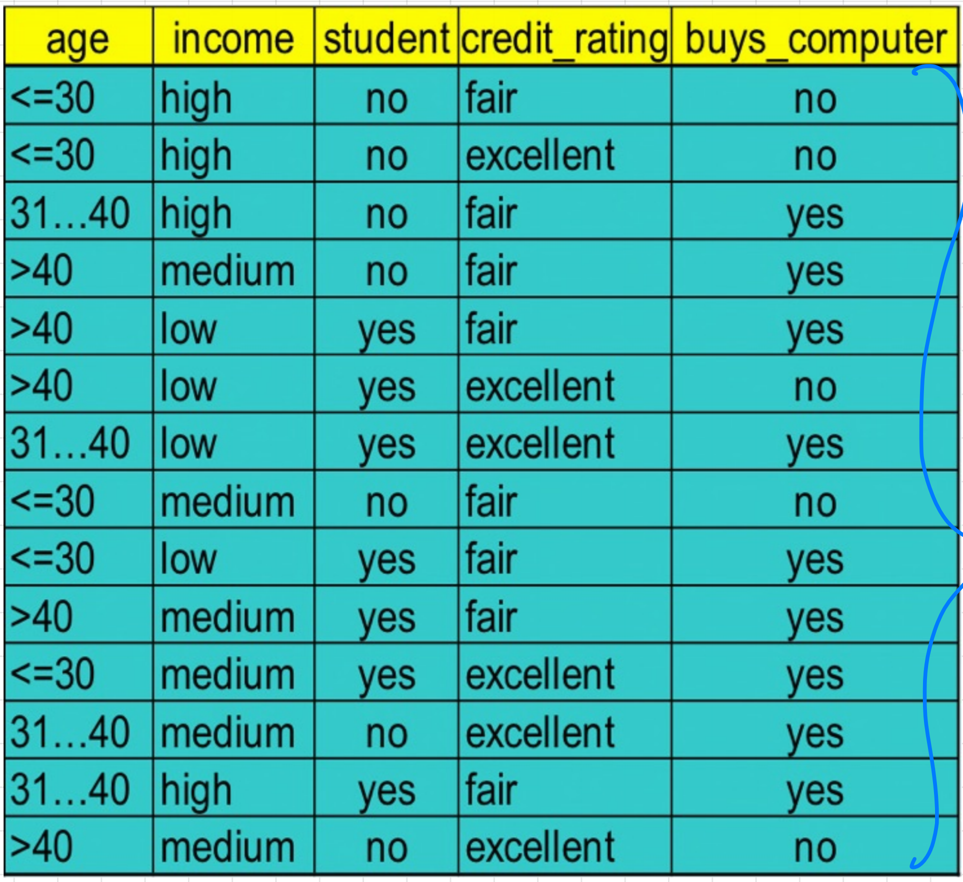
Decision-Tree



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

14

1. $\mathrm{Info}(D) = -\sum_{i=1}^{m} p_i \log_2(p_i)$

Yes : 9

No : 5

$\mathrm{Info}(D) = I(\overset{Y}{9},\overset{N}{5}) = -\frac{9}{14}\log_2\left(\frac{9}{14}\right) - \frac{5}{14}\log_2\left(\frac{5}{14}\right) = 0.940$

2. $\mathrm{Info}_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times \mathrm{Info}(D_j)$ Feature

1 $\mathrm{Info}_{age}(D) = \frac{5}{14} I(2,3) + \frac{4}{14} I(4,0) + \frac{5}{14}(3,2)$

$= \frac{5}{14}\left[-\frac{2}{5}\overset{Y}{\log_2}\left(\frac{2}{5}\right) - \frac{3}{5}\overset{N}{\log_2}\left(\frac{3}{5}\right)\right] + \frac{4}{14}\left[-\frac{4}{4}\overset{Y}{\log_2}\left(\frac{4}{4}\right)\right] + \frac{5}{14}\left[-\frac{3}{5}\overset{Y}{\log_2}\left(\frac{3}{5}\right) - \frac{2}{5}\overset{N}{\log_2}\left(\frac{2}{5}\right)\right] = 0.694$

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

$= \frac{4}{14}\left[-\frac{2}{5}\overset{Y}{\log_2}\left(\frac{2}{4}\right) - \frac{2}{4}\overset{N}{\log_2}\left(\frac{2}{4}\right)\right] + \frac{6}{14}\left[-\frac{4}{4}\overset{Y}{\log_2}\left(\frac{4}{6}\right) - \frac{2}{6}\overset{N}{\log_2}\left(\frac{2}{6}\right)\right] + \frac{4}{14}\left[-\frac{3}{4}\overset{Y}{\log_2}\left(\frac{3}{4}\right) - \frac{1}{5}\overset{N}{\log_2}\left(\frac{1}{5}\right)\right] = 0.911$

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

$= \frac{7}{14}\left[-\frac{3}{7}\overset{Y}{\log_2}\left(\frac{3}{7}\right) - \frac{4}{7}\overset{N}{\log_2}\left(\frac{4}{7}\right)\right] + \frac{7}{14}\left[-\frac{6}{7}\overset{Y}{\log_2}\left(\frac{6}{7}\right) - \frac{1}{7}\overset{N}{\log_2}\left(\frac{1}{7}\right)\right] = 0.789$

4. $\mathrm{Info}_{credit\_rating}(D) = \frac{8}{14} I(6,2) + \frac{6}{14} I(3,3)$

$= \frac{8}{14}\left[-\frac{6}{8}\overset{Y}{\log_2}\left(\frac{6}{8}\right) - \frac{2}{8}\overset{N}{\log_2}\left(\frac{2}{8}\right)\right] + \frac{6}{14}\left[-\frac{3}{6}\overset{Y}{\log_2}\left(\frac{3}{6}\right) - \frac{3}{6}\overset{N}{\log_2}\left(\frac{3}{6}\right)\right] = 0.892$

3. $Gain(A) = Info(D) - Info_A(D)$ ค่า Grain สูงสุดจะเป็น root node

3.1 Grain(age) = 0.940 - 0.694 = 0.246 //

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. แยกกลุ่ม feature ตามค่าใน root node

4.1 <= 30

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

$Info_{income}(D) = \frac{2}{3} I(0,2) + \frac{2}{5} I(1,1) + \frac{1}{5} I(1,0) = 0.4$

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

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

คำนวณ Informtion Grain

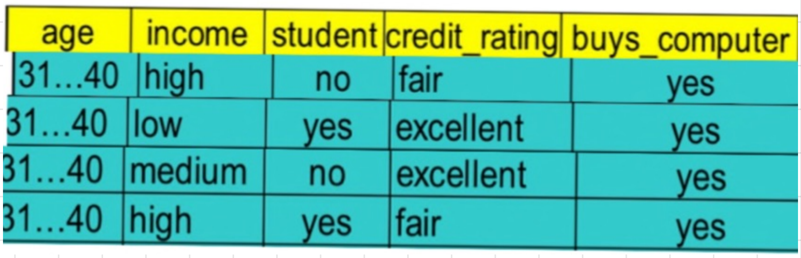
Grain(Income) = 0.971 - 0.4 = 0.571

Grain(Student) = 0.971 - 0 = 0.971 //

Grain(Credit_rating) = 0.971 - 0.951 = 0.02

เลือก Grain(Student) เป็น Node

4.2 30…40



age	income	student	credit_rating	buys_computer
31…40	high	no	fair	yes
31…40	low	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes

yes : 4 no : 0

เมื่ออายุ 31…40 ตอบ Yes

ใน buys_computer

4.3 > 40

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

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

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

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

คำนวณ Informtion Grain

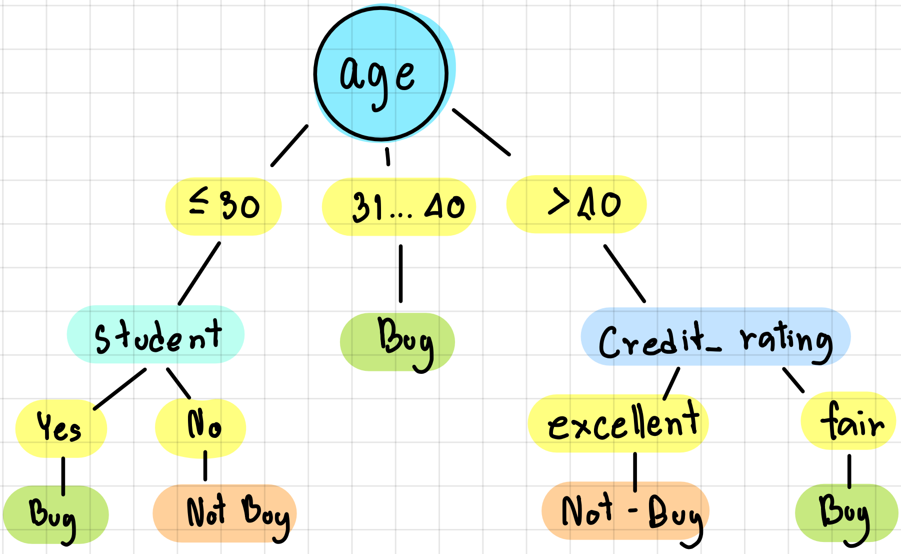
Grain(Income) = 0.971 - 0.951 = 0.02

Grain(Student) = 0.971 - 0.951 = 0.02

Grain(Credit_rating) = 0.971 - 0 = 0.971 //

เลือก Grain(Credit_rating) เป็น Node

5.


age
≤30
31...40
>40
Student
Buy
Credit_ rating
Yes
No
excellent
fair
Buy
Not Buy
Not - Buy
Buy