Sol^n 1:- Here there is two class C_1 and C_2 namely ~~Info (D) =~~ "yes" and "no"

C_1 = Yes = 9 ; C_2 = No = 5.

So,

$$Info(D) = -\sum p_i \log_2(p_i)$$

$$Info(D) = -\frac{9}{14}\log_2\left(\frac{9}{14}\right) - \frac{5}{14}\log_2\left(\frac{5}{14}\right)$$

$$\boxed{Info(D) = \cancel{0.409}\ 0.940}$$

1) Information required for age attribute

$$Info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{D} \times Info(D)$$

Age attribute has three distribution youth, Middle Aged and senior respectively.

So,

$$Info_A(D) = \frac{5}{14} \times \left(\frac{-2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)\right) +$$

$$\frac{4}{14} \times \left(\frac{-4}{4}\log_2\left(\frac{4}{4}\right)\right) + \frac{5}{14} \times \left(\frac{-3}{5}\log_2\left(\frac{3}{5}\right) - \frac{2}{5}\log\right.$$

$Info_A(D) = 0.347 + 0 + 0.347$

$$\boxed{Info_A(D) = 0.694}$$ Age attribute

2) Information required for Income attribute

$$Info_I(D) = \sum_{j=1}^{v} \frac{|D_j|}{D} \times Info(D)$$

$Info_I(D)$: Income has three distribution High, Medium and Low respectively so,

$$Info_I(D) = \frac{4}{14} \times \left(\frac{-2}{4}\log_2\left(\frac{2}{4}\right) - \frac{2}{4}\log_2\left(\frac{2}{4}\right)\right) +$$

$$\frac{6}{14}\left(\frac{-4}{6}\log_2\left(\frac{4}{6}\right) \cdot \frac{-2}{6}\log_2\left(\frac{2}{6}\right)\right) +$$

$$\frac{4}{14}\left(\frac{-3}{4}\log_2\left(\frac{3}{4}\right) \frac{-1}{4}\log_2\left(\frac{1}{4}\right)\right)$$

$$\boxed{Info_I(D) = 0.912}$$

3) Student information required

$$Info_S(D) = \sum_{j=1}^{2} \frac{|D_j|}{D} \times Info(D)$$

Student has two distribution "Yes" and "No" respectively,

So,

$$Info_S(D) = \sum_{j=1}^{2} \frac{|D_j|}{D} \times Info(D)$$

$$Info_S(D) = \frac{7}{14} \times \left(\frac{-3}{7} \log_2\left(\frac{3}{7}\right) - \frac{4}{7} \log_2\left(\frac{4}{7}\right) \right) +$$

$$\frac{7}{14} \times \left(\frac{-6}{7} \log_2\left(\frac{6}{7}\right) - \frac{1}{7} \log_2\left(\frac{1}{7}\right) \right)$$

$$\boxed{Info_S(D) = 0.789}$$

4) Credit - rating information required.

$$Info_C(D) = \sum_{j=1}^{x} \frac{|D_j|}{D} \times Info(D)$$

Credit - rating have two distribution "fair" and "excellent" respectively.
So,

$$Info_C(D) = \frac{8}{14} \times \left(\frac{-6}{8} \log_2\left(\frac{6}{8}\right) - \frac{2}{8} \log_2\left(\frac{2}{8}\right) \right)$$

$$+ \frac{6}{14} \times \left(\frac{-3}{6} \log_2\left(\frac{3}{6}\right) - \frac{3}{6} \log_2\left(\frac{3}{6}\right) \right)$$

$$Info_C(D) = 0.464 + 0.429.$$

$$\boxed{Info_C(D) = 0.893}$$

Now the Gain of all attributes

$$Gain(age) = Info(D) - Info_A(D)$$
$$Gain(age) = 0.940 - 0.694$$

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

2) Gain (income) = Gain $Info(D) - Info_I(D)$

$= 0.940 - 0.912$

Gain (income) = 0.028

3) Gain (student) = $Info(D) - Info_s(D)$

$= 0.940 - 0.789$

Gain (student) = 0.151

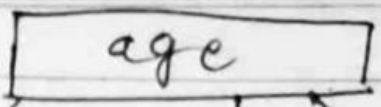
4) Gain (credit-rating) = $Info(D) - Info_c(D)$

Gain (credit_rating) = $0.940 - 0.893$

Gain (credit_rating) = 0.047

On comparing all the gain's of all 4 attributes "age" is the perfect root node.

Decision-Tree



youth

income	student	credit rating	class
high	no	fair	no
high	no	excellent	no
medium	no	fair	no
low	yes	fair	yes
medium	yes	excellent	yes

middle aged

income	student	credit rating	class
high	no	fair	yes
low	yes	excellent	yes
medium	no	excellent	no
high	yes	fair	yes

senior.

income	student	credit rating	class
medium	no	fair	yes
low	yes	fair	yes
low	yes	excellent	no
medium	yes	fair	yes
medium	no	excellent	no.