

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
<del>31...40</del>	<del>low</del>	<del>yes</del>	<del>excellent</del>	<del>yes</del>
<del>&lt;=30</del>	<del>medium</del>	<del>no</del>	<del>fair</del>	<del>no</del>
<=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

age				student			
<=30	Y	N	4	No	3	3	6
31...40	3	0	3	Yes	5	1	6
>40	3	2	5				
income				credit			
h	2	2	4	f	6	1	7
m	4	1	5	e	2	3	5
L	2	1	3				

$$\text{Class Info (D)} = \sum_{i=1}^n p_i \log_2 p_i$$

$$= I(8, 4)$$

$$= -\frac{8}{12} \log_2 \left(\frac{8}{12}\right) - \frac{4}{12} \log_2 \left(\frac{4}{12}\right) = 0.918$$

$$\text{Feature Info age (D)} = \frac{4}{12} I(2, 2) + \frac{3}{12} I(3, 0) + \frac{5}{12} I(3, 2) = 0.809$$

$$\text{Info}_{\text{income}} (D) = \frac{4}{12} I(2, 2) + \frac{5}{12} I(4, 1) + \frac{3}{12} I(2, 1) = 0.832$$

$$\text{Info}_{\text{std}} (D) = \frac{6}{12} I(3, 3) + \frac{6}{12} I(5, 1) = 0.825$$

$$\text{Info}_{\text{credit}} (D) = \frac{7}{12} I(6, 1) + \frac{5}{12} I(2, 3) = 0.747$$

$$\text{Gain in Root node } \text{Gain}(\text{age}) = 0.918 - 0.809 = 0.109$$

$$\text{Gain}(\text{income}) = 0.918 - 0.832 = 0.086$$

$$\text{Gain}(\text{std}) = 0.918 - 0.825 = 0.093$$

$$\text{Gain}(\text{credit}) = 0.918 - 0.747 = \underline{0.171} \quad \text{gain credit with root node}$$

credit

again credit = 2

income<sub>h</sub> = 2 Entropy = 0.918

income<sub>m</sub> = 4 Entropy = 0

income<sub>L</sub> = 3 Entropy = 0

fair

excellent

no

no

yes

no

yes

yes

yes

yes

yes

no

yes

income

Buy

h

m

L

Buy

No

No

∴ in credit = fair we buy Yes

in credit = excellent we buy Income = high we buy yes