

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
<=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

Overall entropy

$$\frac{8}{12} \log_2 \frac{8}{12} + \frac{4}{12} \log_2 \frac{4}{12}$$

$$\text{entropy} = 0.918$$

$$\text{entropy}_{\leq 30} = - \left(\frac{3}{4} \log_2 \frac{3}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right)$$

$$\text{entropy}_{31-40} = - \left(\frac{3}{4} \log_2 \frac{3}{4} + \frac{1}{4} \log_2 \frac{1}{4} \right)$$

$$\text{entropy}_{>40} = - \left(\frac{2}{4} \log_2 \frac{2}{4} + \frac{2}{4} \log_2 \frac{2}{4} \right)$$

sum = 0.049

node student (y, n)

$$\left. \begin{aligned} \text{entropy } y &= -\left(\frac{5}{6} \log_2 \frac{5}{6} + \frac{1}{6} \log_2 \frac{1}{6}\right) \\ \text{entropy } n &= -\left(\frac{3}{6} \log_2 \frac{3}{6} + \frac{3}{6} \log_2 \frac{3}{6}\right) \end{aligned} \right\} \text{sum} = 0.093$$

$$\text{root node} = 0.093 \quad \text{which is } \cancel{\text{not}} \geq n$$

$$\text{credit_rating} \leq 0.5$$

$$\text{entropy} = 0.918$$

$$\text{samples} = n$$