

Homework 2

$$\text{Info}(D) = I\left(\begin{smallmatrix} \text{yes} \\ 9, 5 \end{smallmatrix}\right) = -\frac{9}{14} \log_2\left(\frac{9}{14}\right) - \frac{5}{14} \log_2\left(\frac{5}{14}\right) = 0.940$$

$$\begin{aligned} \text{Info}_{\text{age}}(D) &= \frac{5}{14} I\left(\begin{smallmatrix} \leq 30 \\ 2, 3 \end{smallmatrix}\right) + \frac{4}{14} I\left(\begin{smallmatrix} 31-40 \\ 4, 0 \end{smallmatrix}\right) + \frac{5}{14} I\left(\begin{smallmatrix} > 40 \\ 3, 2 \end{smallmatrix}\right) \\ &= \frac{5}{14} \left[-\frac{2}{5} \log_2\left(\frac{2}{5}\right) - \frac{3}{5} \log_2\left(\frac{3}{5}\right) \right] + \frac{4}{14} \left[-\frac{4}{4} \log_2\left(\frac{4}{4}\right) - \frac{0}{4} \log_2\left(\frac{0}{4}\right) \right] + \\ &\quad \frac{5}{14} \left[-\frac{3}{5} \log_2\left(\frac{3}{5}\right) - \frac{2}{5} \log_2\left(\frac{2}{5}\right) \right] = 0.694 \end{aligned}$$

$$\begin{aligned} \text{Info}_{\text{income}}(D) &= \frac{4}{14} I\left(\begin{smallmatrix} \text{high} \\ 2, 2 \end{smallmatrix}\right) + \frac{6}{14} I\left(\begin{smallmatrix} \text{medium} \\ 4, 2 \end{smallmatrix}\right) + \frac{4}{14} I\left(\begin{smallmatrix} \text{low} \\ 3, 1 \end{smallmatrix}\right) \\ &= \frac{4}{14} \left[-\frac{2}{4} \log_2\left(\frac{2}{4}\right) - \frac{2}{4} \log_2\left(\frac{2}{4}\right) \right] + \frac{6}{14} \left[-\frac{4}{6} \log_2\left(\frac{4}{6}\right) - \frac{2}{6} \log_2\left(\frac{2}{6}\right) \right] + \\ &\quad \frac{4}{14} \left[-\frac{3}{4} \log_2\left(\frac{3}{4}\right) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) \right] = 0.911 \end{aligned}$$

$$\begin{aligned} \text{Info}_{\text{student}}(D) &= \frac{7}{14} I\left(\begin{smallmatrix} \text{yes} \\ 6, 1 \end{smallmatrix}\right) + \frac{7}{14} I\left(\begin{smallmatrix} \text{no} \\ 3, 4 \end{smallmatrix}\right) \\ &= \frac{7}{14} \left[-\frac{6}{7} \log_2\left(\frac{6}{7}\right) - \frac{1}{7} \log_2\left(\frac{1}{7}\right) \right] + \frac{7}{14} \left[-\frac{3}{7} \log_2\left(\frac{3}{7}\right) - \frac{4}{7} \log_2\left(\frac{4}{7}\right) \right] = 0.788 \end{aligned}$$

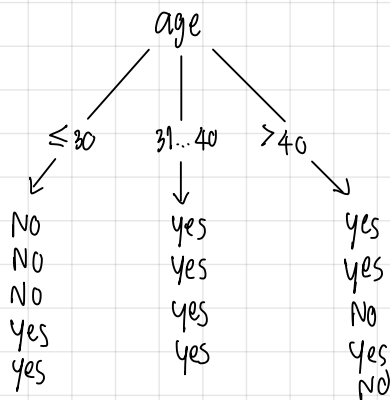
$$\begin{aligned} \text{Info}_{\text{credit-rating}}(D) &= \frac{6}{14} I\left(\begin{smallmatrix} \text{excellent} \\ 3, 1 \end{smallmatrix}\right) + \frac{8}{14} I\left(\begin{smallmatrix} & \\ 6, 2 \end{smallmatrix}\right) \\ &= \frac{6}{14} \left[-\frac{3}{6} \log_2\left(\frac{3}{6}\right) - \frac{3}{6} \log_2\left(\frac{3}{6}\right) \right] + \frac{8}{14} \left[-\frac{6}{8} \log_2\left(\frac{6}{8}\right) - \frac{2}{8} \log_2\left(\frac{2}{8}\right) \right] = 0.892 \end{aligned}$$

$$\text{Gain}_{\text{age}} = 0.940 - 0.694 = 0.246$$

$$\text{Gain}_{\text{income}} = 0.940 - 0.911 = 0.029$$

$$\text{Gain}_{\text{student}} = 0.940 - 0.788 = 0.152$$

$$\text{Gain}_{\text{credit-rating}} = 0.940 - 0.892 = 0.048$$



Age ≤ 30

$$\text{Info}_{\text{age} \leq 30}(D) = I\left(\begin{smallmatrix} 2, 3 \end{smallmatrix}\right) = -\frac{2}{5} \log_2\left(\frac{2}{5}\right) - \frac{3}{5} \log_2\left(\frac{3}{5}\right) = 0.971$$

$$\begin{aligned} \text{Info}_{\text{income}}(D) &= \frac{2}{5} I\left(\begin{smallmatrix} \text{high} \\ 0, 2 \end{smallmatrix}\right) + \frac{2}{5} I\left(\begin{smallmatrix} \text{medium} \\ 1, 1 \end{smallmatrix}\right) + \frac{1}{5} I\left(\begin{smallmatrix} \text{low} \\ 1, 0 \end{smallmatrix}\right) \\ &= \frac{2}{5} \left[-\frac{0}{2} \log_2\left(\frac{0}{2}\right) - \frac{2}{2} \log_2\left(\frac{2}{2}\right) \right] + \frac{2}{5} \left[-\frac{1}{2} \log_2\left(\frac{1}{2}\right) - \frac{1}{2} \log_2\left(\frac{1}{2}\right) \right] + \\ &\quad \frac{1}{5} \left[-\frac{1}{1} \log_2\left(\frac{1}{1}\right) - \frac{0}{1} \log_2\left(\frac{0}{1}\right) \right] = 0.4 \end{aligned}$$

$$\begin{aligned} \text{Info}_{\text{student}}(D) &= \frac{2}{5} I\left(\begin{smallmatrix} \text{yes} \\ 3, 0 \end{smallmatrix}\right) + \frac{3}{5} I\left(\begin{smallmatrix} \text{no} \\ 0, 3 \end{smallmatrix}\right) \\ &= \frac{2}{5} \left[-\frac{3}{2} \log_2\left(\frac{3}{2}\right) - \frac{0}{2} \log_2\left(\frac{0}{2}\right) \right] + \frac{3}{5} \left[-\frac{0}{3} \log_2\left(\frac{0}{3}\right) - \frac{3}{3} \log_2\left(\frac{3}{3}\right) \right] = 0 \end{aligned}$$

$$\text{Info}_{\text{credit-rating}}(D) = \frac{2}{5} I\left(\begin{smallmatrix} \text{excellent} \\ 1, 1 \end{smallmatrix}\right) + \frac{3}{5} I\left(\begin{smallmatrix} & \\ 1, 2 \end{smallmatrix}\right)$$

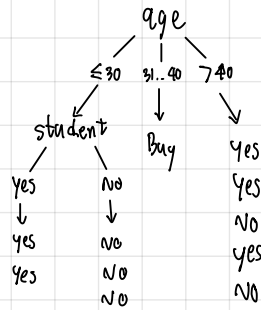
age	income	student	credit rating	buys computer
≤30	high	no	fair	no
≤30	high	no	excellent	no
31...40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
≤30	medium	no	fair	no
≤30	low	yes	fair	yes
>40	medium	yes	fair	yes
≤30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

$$= \frac{2}{5} \left[-\frac{1}{2} \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \log_2 \left(\frac{1}{2} \right) \right] + \frac{3}{5} \left[-\frac{1}{3} \log_2 \left(\frac{1}{3} \right) - \frac{2}{3} \log_2 \left(\frac{2}{3} \right) \right] = 0.951$$

$$\text{Gain}_{\text{income}} = 0.971 - 0.4 = 0.571$$

$$\text{Gain}_{\text{student}} = 0.971 - 0 = 0.971$$

$$\text{Gain}_{\text{credit-rating}} = 0.971 - 0.951 = 0.020$$



Age > 40

$$\text{Info}_{\text{age} > 40}(D) = I(3,2) = -\frac{3}{5} \log_2 \left(\frac{3}{5} \right) - \frac{2}{5} \log_2 \left(\frac{2}{5} \right) = 0.971$$

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

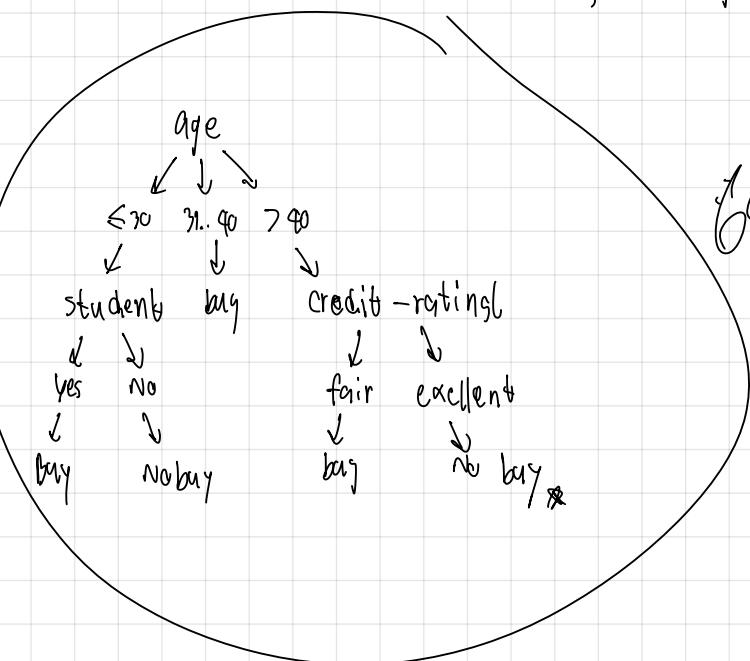
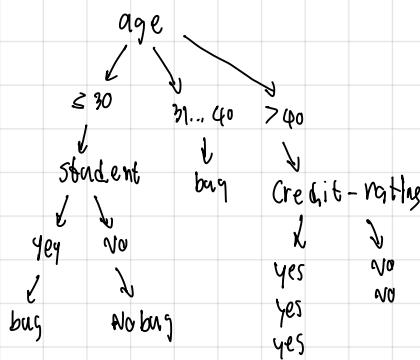
$$\begin{aligned} \text{Info}_{\text{student}}(D) &= \frac{3}{5} I(2,1) + \frac{2}{5} I(1,1) \\ &= 0.951 \end{aligned}$$

$$\begin{aligned} \text{Info}_{\text{credit-rating}}(D) &= \frac{2}{3} I(0,2) + \frac{3}{3} I(3,0) \\ &= 0 \end{aligned}$$

$$\text{Gain}_{\text{income}} = 0.971 - 0.951 = 0.020$$

$$\text{Gain}_{\text{student}} = 0.971 - 0.951 = 0.020$$

$$\text{Gain}_{\text{credit-rating}} = 0.971 - 0 = 0.971$$



643020522

اسم اسم