HW 14

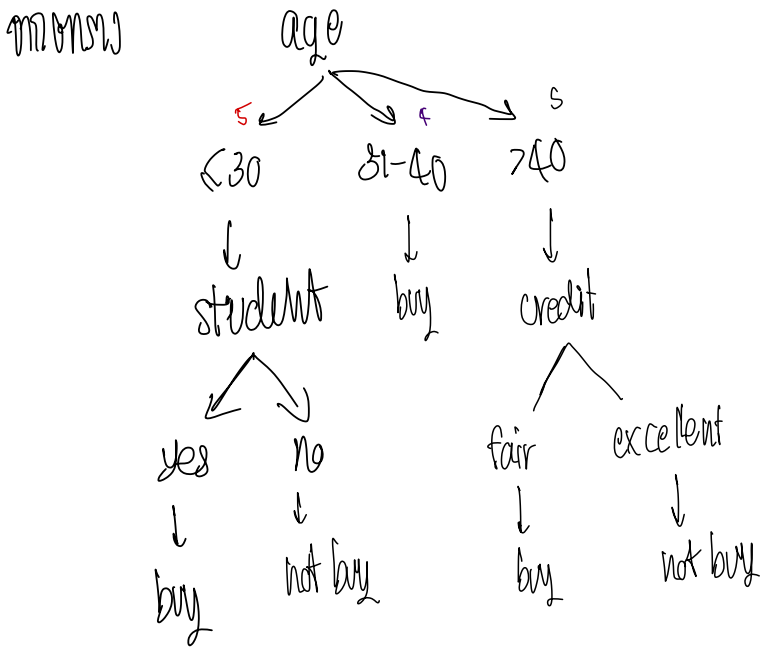
$$Gain(age) = Info(D) - Info_{age}(D) = 0.246$$

Similarly, we can get

$$Gain(income) = 0.029$$

$$Gain(student) = 0.151$$

$$Gain(credit\_rating) = 0.048$$



Age ≤30; $Info(D) = I(2,3) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5}$

$= 0.5288 + 0.4422$

$= 0.971$ *

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

$= \frac{2}{5}\left[-\frac{0}{2}\log_2\frac{0}{2} - \frac{2}{2}\log_2\frac{2}{2}\right] + \frac{2}{5}\left[-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}\right]$

$+ \frac{1}{5}\left[-\frac{1}{1}\log_2\frac{1}{1} - \frac{0}{1}\log_2\frac{0}{1}\right]$

$= 0 + \frac{2}{5}(0.5+0.5) + 0 = 0.40$ *

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

$= \frac{2}{5}\left[-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}\right] + \frac{3}{5}\left[-\frac{0}{3}\log_2\frac{0}{3} - \frac{3}{3}\log_2\frac{3}{3}\right]$

$= 0 - 0 = 0$ *

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

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

$= 0.5510 + 0.4 = 0.9510$ *

$Gain(income) = Info(D) - Info_{income}(D) = 0.971 - 0.40 = 0.571$ *

$Gain(student) = Info(D) - Info_{student}(D) = 0.971 - 0 = 0.971$ *

$Gain(credit) = Info(D) - Info_{credit}(D) = 0.971 - 0.9510 = 0.02$ *

Age 31-40; $\text{Info}(D) = I(4,0) = -\frac{4}{4}\log_2\frac{4}{4} - \frac{0}{4}\log_2\frac{0}{4} = 0$ ✱

$\text{Info income}(D) = \frac{2}{4}I(2,0) + \frac{1}{4}I(1,0) + \frac{1}{4}I(1,0)$

$= \frac{2}{4}\left[-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}\right] + \frac{1}{4}\left[-\frac{1}{1}\log_2\frac{1}{1} - \frac{0}{1}\log_2\frac{1}{1}\right] + \frac{1}{4}\left[-\frac{1}{1}\log_2\frac{1}{1} - \frac{0}{1}\log_2\frac{1}{1}\right] = 0$ ✱

$\therefore$ 31-40 boy ada yes semua

$\text{Info student}(D) = \frac{2}{4}I(2,0) + \frac{2}{4}I(2,0)$

$= \frac{2}{4}\left[-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}\right] + \frac{2}{4}\left[-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}\right] = 0$ ✱

$\text{Info credit}(D) = \frac{2}{4}I(2,0) + \frac{2}{4}(2,0)$

$= \frac{2}{4}\left[-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}\right] + \frac{2}{4}\left[-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}\right] = 0$ ✱

Age >40; $\text{Info}(D) = I(3,2) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.4422 + 0.5288 = 0.9710$ ✱

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

$= \frac{3}{5}\left[-\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3}\right] + \frac{2}{5}\left[-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}\right] = 0.5510 + 0.4 = 0.9510$ ✱

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

$= \frac{3}{5}\left[-\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3}\right] + \frac{2}{5}\left[-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}\right] = 0.5510 + 0.4 = 0.9510$ ✱

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

$= \frac{3}{5}\left[-\frac{3}{3}\log_2\frac{3}{3} - \frac{0}{3}\log_2\frac{0}{3}\right] + \frac{2}{5}\left[-\frac{0}{2}\log_2\frac{0}{2} - \frac{2}{2}\log_2\frac{2}{2}\right] = 0$ ✱

$\text{Gain (income)} = \text{Info}(D) - \text{Info income}(D)$

$= 0.9710 - 0.9510$

$= 0.020$ ✱

$\text{Gain (student)} = \text{Info}(D) - \text{Info student}(D)$

$= 0.9710 - 0.9510$

$= 0.020$ ✱

$\text{Gain (credit)} = \text{Info}(D) - \text{Info credit}(D)$

$= 0.9710 - 0$

$= 0.9710$ ✱