## Training data set: Who buys computer?

| 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 |

$$Info(D) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.940 \text{ \#}$$

$y = 9$

$n = 5$

$$Age \quad \leq 30 \quad y\ 2 \quad n\ 3$$
$$31-40 \quad y\ 4 \quad n\ 0$$
$$>40 \quad y\ 3 \quad n\ 2$$

$$Info_{age} = \frac{5}{14}I(2,3) + \frac{4}{14}I(4,0) + \frac{5}{14}I(3,2) = 0.694$$

$$= Info(D) - Info_{age} = 0.940 - 0.694 = \boxed{0.246}$$

---

$$Income \quad low \quad y\ 3 \quad n\ 1$$
$$medium \quad y\ 4 \quad n\ 2$$
$$high \quad y\ 2 \quad n\ 2$$

$$Info_{Income} = \frac{4}{14}I(3,1) + \frac{6}{14}I(4,2) + \frac{4}{14}I(2,2) = 0.911$$

$$Gain\ income = Info(D) - Info_{income} = 0.940 - 0.911 = 0.029$$

---

$$Student \quad yes \quad y\ 6 \quad n\ 1$$
$$no \quad y\ 3 \quad n\ 4$$

$$Info_{student} = \frac{7}{14}I(6,1) + \frac{7}{14}I(3,4) = 0.7885$$
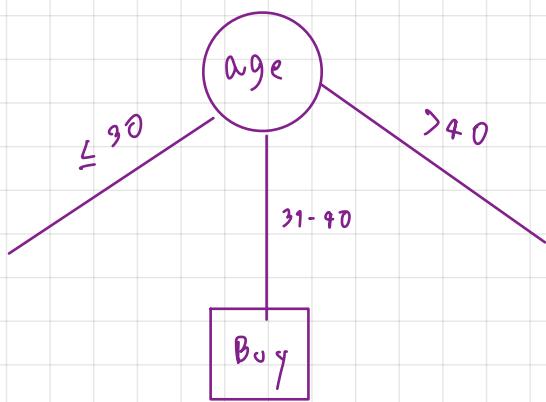
$$Gain_{student} = Info(D) - Info_{student} = 0.940 - 0.789 = 0.151$$

---

$$Credit \quad fair \quad y\ 6 \quad n\ 2$$
$$excellent \quad y\ 3 \quad n\ 3$$

$$Info_{credit} = \frac{8}{14}I(6,2) + \frac{6}{14}I(3,3) = 0.892$$

$$Gain = Info(D) - Info_{credit} = 0.940 - 0.892 = 0.048$$

---

$\therefore$ Age เป็นตัวที่ดี ที่ suด

Top tree:
- age
  - ≤30
  - 31-40 → Buy
  - >40

---
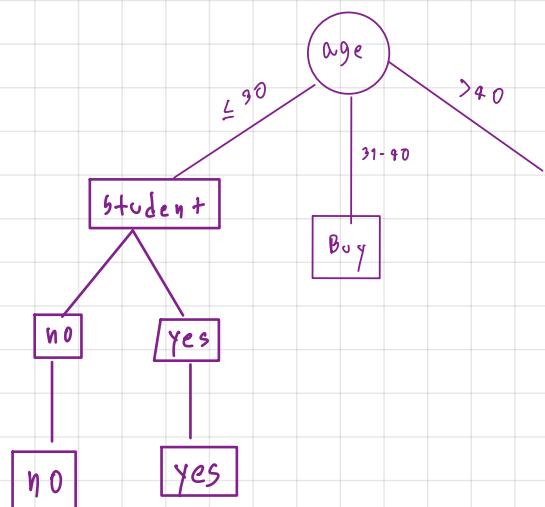
หัวข้อ  Student > credit > income

Info (D): $I(5,5) = 1$

Info$_{≤30}$ = $\frac{2}{5}(2,0) + \frac{3}{5}(0,3) = 0$

Info$_{>40}$ = $\frac{3}{5}(2,1) + \frac{2}{5}(1,1) = 0.951$

Gain = ≤30 = $1.0 = 1$

>40 = $1 - 0.951 = 0.049$



age
- ≤30 → student
  - no → no
  - yes → yes
- 31-40 → Buy
- >40

---

หัวข้อ  Credit > income

Info (D) = $I(3,2) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.970$

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

Info$_{income}$ = $\frac{3}{5}I(2,1) + \frac{2}{5}(1,1) = 0.951$

Gain = Credit = $0.970 - 0 = 0.970$

Income = $0.970 - 0.951 = 0.019$



age
- ≤30 → student
  - no → no
  - yes → yes
- 31-40 → Buy
- >40 → Credit
  - fair → yes
  - excellent → no

หาย  วิสุ ณ  แสงหรีง     6630202 96-3