

| หมายเลข | สัดส่วน | หน้าตา | คำถาม | ประพจน์ | ผล |
|---------|---------|---------|-------|---------|------|
| 1 | ปานกลาง | ปานกลาง | 10 | ปานกลาง | ตก |
| 2 | ดี | ปานกลาง | 8 | ดี | เข้า |
| 3 | ปานกลาง | ต่ำ | 8 | ต่ำ | ตก |
| 4 | ดี | ดี | 9 | ปานกลาง | เข้า |
| 5 | ปานกลาง | ดี | 10 | ปานกลาง | เข้า |
| 6 | ดี | ปานกลาง | 7 | ดี | เข้า |
| 7 | ต่ำ | ปานกลาง | 10 | ต่ำ | ตก |
| 8 | ต่ำ | ต่ำ | 7 | ต่ำ | ตก |
| 9 | ปานกลาง | ดี | 8 | ปานกลาง | เข้า |
| 10 | ดี | ดี | 8 | ดี | เข้า |
| 11 | ปานกลาง | ปานกลาง | 9 | ต่ำ | เข้า |
| 12 | ต่ำ | ปานกลาง | 7 | ต่ำ | ตก |
| 13 | ปานกลาง | ดี | 8 | ต่ำ | เข้า |
| 14 | ปานกลาง | ต่ำ | 8 | ปานกลาง | ตก |
| 15 | ต่ำ | ต่ำ | 7 | ต่ำ | ตก |
| 16 | ดี | ปานกลาง | 6 | ดี | เข้า |
| 17 | ต่ำ | ปานกลาง | 7 | ดี | ตก |
| 18 | ปานกลาง | ต่ำ | 8 | ปานกลาง | เข้า |
| 19 | ต่ำ | ต่ำ | 9 | ดี | ตก |
| 20 | ต่ำ | ปานกลาง | 10 | ต่ำ | เข้า |

ชั้นที่ 1

$$H(x) = - \left[\frac{11}{20} * \log_2 \left(\frac{11}{20} \right) + \frac{9}{20} * \log_2 \left(\frac{9}{20} \right) \right] = 0.993$$

สัดส่วน

$$\text{Gain (ผล, สัดส่วน)} = P(\text{ผล}|\text{ดี}) * H(\text{ผล}|\text{ดี}) + P(\text{ผล}|\text{ปานกลาง}) * H(\text{ผล}|\text{ปานกลาง}) + P(\text{ผล}|\text{ต่ำ}) * H(\text{ผล}|\text{ต่ำ})$$

$$H(\text{ผล}|\text{ดี}) = - \left[\frac{5}{5} * \log_2 \left(\frac{5}{5} \right) + \frac{0}{5} * \log_2 \left(\frac{0}{5} \right) \right] = 0$$

$$H(\text{ผล}|\text{ปานกลาง}) = - \left[\frac{5}{8} * \log_2 \left(\frac{5}{8} \right) + \frac{3}{8} * \log_2 \left(\frac{3}{8} \right) \right] = 0.954$$

$$H(\text{ผล}|\text{ต่ำ}) = - \left[\frac{1}{7} * \log_2 \left(\frac{1}{7} \right) + \frac{6}{7} * \log_2 \left(\frac{6}{7} \right) \right] = 0.592$$

$$\text{Gain (ผล, สัดส่วน)} = 0.993 - \left[\frac{5}{20} * 0 + \frac{8}{20} * 0.954 + \frac{7}{20} * 0.592 \right] = 0.405$$

$$H(\text{สัดส่วน}) = - \left[\frac{5}{20} * \log_2 \left(\frac{5}{20} \right) + \frac{8}{20} * \log_2 \left(\frac{8}{20} \right) + \frac{7}{20} * \log_2 \left(\frac{7}{20} \right) \right] = 1.559$$

$$\therefore \text{GainRatio(ผล, สัดส่วน)} = \frac{0.405}{1.559} = 0.26$$

หน้าตา

$$\text{Gain (ผล, หน้าตา)} = P(\text{ผล|ดี}) * H(\text{ผล|ดี}) + P(\text{ผล|ปานกลาง}) * H(\text{ผล|ปานกลาง}) + P(\text{ผล|ต่ำ}) * H(\text{ผล|ต่ำ})$$

$$H(\text{ผล|ดี}) = -\left[\frac{5}{5} * \log_2\left(\frac{5}{5}\right) + \frac{0}{5} * \log_2\left(\frac{0}{5}\right)\right] = 0$$

$$H(\text{ผล|ปานกลาง}) = -\left[\frac{5}{9} * \log_2\left(\frac{5}{9}\right) + \frac{4}{9} * \log_2\left(\frac{4}{9}\right)\right] = 0.991$$

$$H(\text{ผล|ต่ำ}) = -\left[\frac{1}{6} * \log_2\left(\frac{1}{6}\right) + \frac{5}{6} * \log_2\left(\frac{5}{6}\right)\right] = 0.65$$

$$\text{Gain (ผล, หน้าตา)} = 0.993 - \left[\frac{5}{20} * 0 + \frac{9}{20} * 0.991 + \frac{6}{20} * 0.65\right] = 0.352$$

$$H(\text{หน้าตา}) = -\left[\frac{5}{20} * \log_2\left(\frac{5}{20}\right) + \frac{9}{20} * \log_2\left(\frac{9}{20}\right) + \frac{6}{20} * \log_2\left(\frac{6}{20}\right)\right] = 1.539$$

$$\therefore \text{GainRatio (ผล, หน้าตา)} = \frac{0.352}{1.539} = 0.229$$

คำถาม₆

$$\text{Gain (ผล, คำถาม}_6) = P(\text{ผล} \leq 6) * H(\text{ผล|ดี} \leq 6) + P(\text{ผล} > 6) * H(\text{ผล|ดี} > 6)$$

$$H(\text{ผล} \leq 6) = -\left[\frac{1}{1} * \log_2\left(\frac{1}{1}\right) + \frac{0}{1} * \log_2\left(\frac{0}{1}\right)\right] = 0$$

$$H(\text{ผล} > 6) = -\left[\frac{10}{19} * \log_2\left(\frac{10}{19}\right) + \frac{9}{19} * \log_2\left(\frac{9}{19}\right)\right] = 0.998$$

$$\text{Gain (ผล, คำถาม}_6) = 0.993 - \left[\frac{1}{20} * 0 + \frac{19}{20} * 0.998\right] = 0.045$$

คำถาม₇

$$\text{Gain (ผล, คำถาม}_7) = P(\text{ผล} \leq 7) * H(\text{ผล|ดี} \leq 7) + P(\text{ผล} > 7) * H(\text{ผล|ดี} > 7)$$

$$H(\text{ผล} \leq 7) = -\left[\frac{2}{6} * \log_2\left(\frac{2}{6}\right) + \frac{4}{6} * \log_2\left(\frac{4}{6}\right)\right] = 0.918$$

$$H(\text{ผล} > 7) = -\left[\frac{9}{14} * \log_2\left(\frac{9}{14}\right) + \frac{5}{14} * \log_2\left(\frac{5}{14}\right)\right] = 0.94$$

$$\text{Gain (ผล, คำถาม}_7) = 0.993 - \left[\frac{6}{20} * 0.918 + \frac{14}{20} * 0.94\right] = 0.06$$

คำถาม₈

$$\text{Gain (ผล, คำถาม}_8) = P(\text{ผล} \leq 8) * H(\text{ผล|ดี} \leq 8) + P(\text{ผล} > 8) * H(\text{ผล|ดี} > 8)$$

$$H(\text{ผล} \leq 8) = -\left[\frac{7}{13} * \log_2\left(\frac{7}{13}\right) + \frac{6}{13} * \log_2\left(\frac{6}{13}\right)\right] = 0.996$$

$$H(\text{ผล} > 8) = -\left[\frac{4}{7} * \log_2\left(\frac{4}{7}\right) + \frac{3}{7} * \log_2\left(\frac{3}{7}\right)\right] = 0.985$$

$$\text{Gain (ผล, คำถาม}_8) = 0.993 - \left[\frac{13}{20} * 0.996 + \frac{7}{20} * 0.985\right] = 0.001$$

คำถาม₉

$$\text{Gain}(\text{ผล}, \text{คำถาม}_9) = P(\text{ผล} \leq 9) * H(\text{ผล} | \text{ดี} \leq 9) + P(\text{ผล} > 9) * H(\text{ผล} | \text{ดี} > 9)$$

$$H(\text{ผล} | \leq 9) = - \left[\frac{9}{16} * \log_2 \left(\frac{9}{16} \right) + \frac{7}{16} * \log_2 \left(\frac{7}{16} \right) \right] = 0.989$$

$$H(\text{ผล} | > 9) = - \left[\frac{2}{4} * \log_2 \left(\frac{2}{4} \right) + \frac{2}{4} * \log_2 \left(\frac{2}{4} \right) \right] = 1$$

$$\text{Gain}(\text{ผล}, \text{คำถาม}_9) = 0.993 - \left[\frac{16}{20} * 0.989 + \frac{4}{20} * 1 \right] = 0.002$$

$$H(\text{คำถาม}_7) = - \left[\frac{6}{20} * \log_2 \left(\frac{6}{20} \right) + \frac{14}{20} * \log_2 \left(\frac{14}{20} \right) \right] = 0.881$$

$$\therefore \text{GainRatio}(\text{ผล}, \text{คำถาม}_7) = \frac{0.06}{0.881} = 0.068$$

ประพฤติ

$$\text{Gain}(\text{ผล}, \text{ประพฤติ}) = P(\text{ผล} | \text{ดี}) * H(\text{ผล} | \text{ดี}) + P(\text{ผล} | \text{ปานกลาง}) * H(\text{ผล} | \text{ปานกลาง}) + P(\text{ผล} | \text{ต่ำ}) * H(\text{ผล} | \text{ต่ำ})$$

$$H(\text{ผล} | \text{ดี}) = - \left[\frac{4}{6} * \log_2 \left(\frac{4}{6} \right) + \frac{2}{6} * \log_2 \left(\frac{2}{6} \right) \right] = 0.918$$

$$H(\text{ผล} | \text{ปานกลาง}) = - \left[\frac{4}{6} * \log_2 \left(\frac{4}{6} \right) + \frac{2}{6} * \log_2 \left(\frac{2}{6} \right) \right] = 0.918$$

$$H(\text{ผล} | \text{ต่ำ}) = - \left[\frac{3}{8} * \log_2 \left(\frac{3}{8} \right) + \frac{5}{8} * \log_2 \left(\frac{5}{8} \right) \right] = 0.954$$

$$\text{Gain}(\text{ผล}, \text{ประพฤติ}) = 0.993 - \left[\frac{6}{20} * 0.918 + \frac{6}{20} * 0.918 + \frac{8}{20} * 0.954 \right] = 0.061$$

$$H(\text{ประพฤติ}) = - \left[\frac{6}{20} * \log_2 \left(\frac{6}{20} \right) + \frac{6}{20} * \log_2 \left(\frac{6}{20} \right) + \frac{8}{20} * \log_2 \left(\frac{8}{20} \right) \right] = 1.571$$

$$\therefore \text{GainRatio}(\text{ผล}, \text{ประพฤติ}) = \frac{0.061}{1.571} = 0.039$$

เลือก feature สัดส่วนในชั้นที่ 1



ชั้นที่ 2

สัดส่วน = ดี

$$H(x) = 0$$

สัดส่วน = ปานกลาง

$$H(x) = 0.954$$

หน้าตา

$$\text{Gain (ผล, หน้าตา)} = P(\text{ผล|ดี}) * H(\text{ผล|ดี}) + P(\text{ผล|ปานกลาง}) * H(\text{ผล|ปานกลาง}) + P(\text{ผล|ต่ำ}) * H(\text{ผล|ต่ำ})$$

$$H(\text{ผล|ดี}) = -\left[\frac{3}{3} * \log_2\left(\frac{3}{3}\right) + \frac{0}{3} * \log_2\left(\frac{0}{3}\right)\right] = 0$$

$$H(\text{ผล|ปานกลาง}) = -\left[\frac{1}{2} * \log_2\left(\frac{1}{2}\right) + \frac{1}{2} * \log_2\left(\frac{1}{2}\right)\right] = 1$$

$$H(\text{ผล|ต่ำ}) = -\left[\frac{1}{3} * \log_2\left(\frac{1}{3}\right) + \frac{2}{3} * \log_2\left(\frac{2}{3}\right)\right] = 0.918$$

$$\text{Gain (ผล, หน้าตา)} = 0.954 - \left[\frac{3}{8} * 0 + \frac{2}{8} * 1 + \frac{3}{8} * 0.918\right] = 0.36$$

$$H(\text{หน้าตา}) = -\left[\frac{3}{8} * \log_2\left(\frac{3}{8}\right) + \frac{2}{8} * \log_2\left(\frac{2}{8}\right) + \frac{3}{8} * \log_2\left(\frac{3}{8}\right)\right] = 1.561$$

$$\therefore \text{GainRatio(ผล, สัดส่วน)} = \frac{0.36}{1.561} = 0.231$$

คำถาม₈

$$\text{Gain (ผล, คำถาม}_8) = P(\text{ผล} \leq 8) * H(\text{ผล|ดี} \leq 8) + P(\text{ผล} > 8) * H(\text{ผล} > 8)$$

$$H(\text{ผล} \leq 8) = -\left[\frac{3}{5} * \log_2\left(\frac{3}{5}\right) + \frac{2}{5} * \log_2\left(\frac{2}{5}\right)\right] = 0.971$$

$$H(\text{ผล} > 8) = -\left[\frac{2}{3} * \log_2\left(\frac{2}{3}\right) + \frac{1}{3} * \log_2\left(\frac{1}{3}\right)\right] = 0.918$$

$$\text{Gain (ผล, คำถาม}_8) = 0.954 - \left[\frac{5}{8} * 0.971 + \frac{3}{8} * 0.918\right] = 0.003$$

คำถาม₉

$$\text{Gain (ผล, คำถาม}_9) = P(\text{ผล} \leq 9) * H(\text{ผล|ดี} \leq 9) + P(\text{ผล} > 9) * H(\text{ผล} > 9)$$

$$H(\text{ผล} \leq 9) = -\left[\frac{4}{6} * \log_2\left(\frac{4}{6}\right) + \frac{2}{6} * \log_2\left(\frac{2}{6}\right)\right] = 0.918$$

$$H(\text{ผล} > 9) = -\left[\frac{1}{2} * \log_2\left(\frac{1}{2}\right) + \frac{1}{2} * \log_2\left(\frac{1}{2}\right)\right] = 1$$

$$\text{Gain (ผล, คำถาม}_9) = 0.954 - \left[\frac{6}{8} * 0.918 + \frac{2}{8} * 1\right] = 0.016$$

$$H(\text{คำถาม}_9) = -\left[\frac{6}{8} * \log_2\left(\frac{6}{8}\right) + \frac{2}{8} * \log_2\left(\frac{2}{8}\right)\right] = 0.811$$

$$\therefore \text{GainRatio}(\text{ผล, ค่าถาม}_9) = \frac{0.016}{0.811} = 0.02$$

ประพฤติ

$$\text{Gain}(\text{ผล, ประพฤติ}) = P(\text{ผล|ปานกลาง}) * H(\text{ผล|ปานกลาง}) + P(\text{ผล|ต่ำ}) * H(\text{ผล|ต่ำ})$$

$$H(\text{ผล|ปานกลาง}) = -\left[\frac{3}{5} * \log_2\left(\frac{3}{5}\right) + \frac{2}{5} * \log_2\left(\frac{2}{5}\right)\right] = 0.971$$

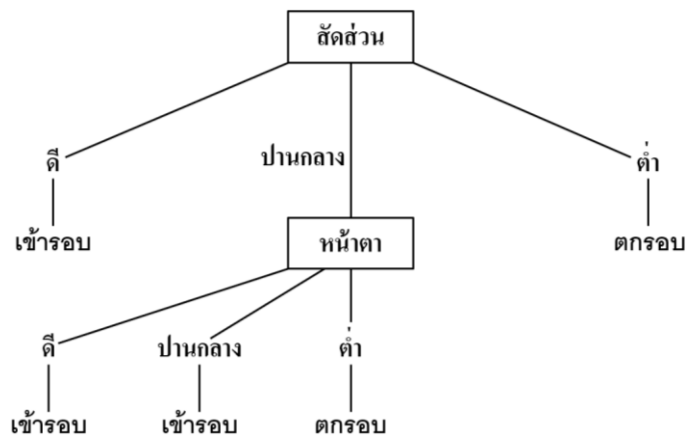
$$H(\text{ผล|ต่ำ}) = -\left[\frac{1}{3} * \log_2\left(\frac{1}{3}\right) + \frac{2}{3} * \log_2\left(\frac{2}{3}\right)\right] = 0.918$$

$$\text{Gain}(\text{ผล, ประพฤติ}) = 0.954 - \left[\frac{5}{8} * 0.971 + \frac{3}{8} * 0.918\right] = 0.03$$

$$H(\text{ประพฤติ}) = -\left[\frac{5}{8} * \log_2\left(\frac{5}{8}\right) + \frac{3}{8} * \log_2\left(\frac{3}{8}\right)\right] = 0.954$$

$$\therefore \text{GainRatio}(\text{ผล, ประพฤติ}) = \frac{0.003}{0.954} = 0.003$$

เลือก feature หน้าตาในชั้นที่ 2 node สัดส่วน = ปานกลาง



สัดส่วน = ต่ำ

$$H(x) = 0.592$$

หน้าตา

$$\text{Gain}(\text{ผล, หน้าตา}) = P(\text{ผล|ปานกลาง}) * H(\text{ผล|ปานกลาง}) + P(\text{ผล|ต่ำ}) * H(\text{ผล|ต่ำ})$$

$$H(\text{ผล|ปานกลาง}) = -\left[\frac{1}{4} * \log_2\left(\frac{1}{4}\right) + \frac{3}{4} * \log_2\left(\frac{3}{4}\right)\right] = 0.811$$

$$H(\text{ผล|ต่ำ}) = -\left[\frac{0}{3} * \log_2\left(\frac{0}{3}\right) + \frac{3}{3} * \log_2\left(\frac{3}{3}\right)\right] = 0$$

$$\text{Gain}(\text{ผล, หน้าตา}) = 0.592 - \left[\frac{4}{7} * 0.811 + \frac{3}{7} * 0\right] = 0.128$$

$$H(\text{หน้าตา}) = -\left[\frac{4}{7} * \log_2\left(\frac{4}{7}\right) + \frac{3}{7} * \log_2\left(\frac{3}{7}\right)\right] = 0.985$$

$$\therefore \text{GainRatio}(\text{ผล, หน้าตา}) = \frac{0.128}{0.958} = 0.134$$

คำถาม₇

$$\text{Gain}(\text{ผล, คำถาม}_7) = P(\text{ผล} \leq 7) * H(\text{ผล} | \text{ดี} \leq 7) + P(\text{ผล} > 7) * H(\text{ผล} | \text{ดี} > 7)$$

$$H(\text{ผล} \leq 7) = -\left[\frac{0}{4} * \log_2\left(\frac{0}{4}\right) + \frac{4}{4} * \log_2\left(\frac{4}{4}\right)\right] = 0$$

$$H(\text{ผล} > 7) = -\left[\frac{1}{3} * \log_2\left(\frac{1}{3}\right) + \frac{2}{3} * \log_2\left(\frac{2}{3}\right)\right] = 0.918$$

$$\text{Gain}(\text{ผล, คำถาม}_7) = 0.592 - \left[\frac{4}{7} * 0 + \frac{3}{7} * 0.918\right] = 0.198$$

คำถาม₉

$$\text{Gain}(\text{ผล, คำถาม}_9) = P(\text{ผล} \leq 9) * H(\text{ผล} | \text{ดี} \leq 9) + P(\text{ผล} > 9) * H(\text{ผล} | \text{ดี} > 9)$$

$$H(\text{ผล} \leq 9) = -\left[\frac{0}{5} * \log_2\left(\frac{0}{5}\right) + \frac{5}{5} * \log_2\left(\frac{5}{5}\right)\right] = 0$$

$$H(\text{ผล} > 9) = -\left[\frac{1}{2} * \log_2\left(\frac{1}{2}\right) + \frac{1}{2} * \log_2\left(\frac{1}{2}\right)\right] = 1$$

$$\text{Gain}(\text{ผล, คำถาม}_9) = 0.592 - \left[\frac{5}{7} * 0 + \frac{2}{7} * 1\right] = 0.306$$

$$H(\text{คำถาม}_7) = -\left[\frac{5}{7} * \log_2\left(\frac{5}{7}\right) + \frac{2}{7} * \log_2\left(\frac{2}{7}\right)\right] = 0.863$$

$$\therefore \text{GainRatio}(\text{ผล, คำถาม}_7) = \frac{0.306}{0.863} = 0.355$$

ประพฤติ

$$\text{Gain}(\text{ผล, ประพฤติ}) = P(\text{ผล} | \text{ดี}) * H(\text{ผล} | \text{ดี}) + P(\text{ผล} | \text{ต่ำ}) * H(\text{ผล} | \text{ต่ำ})$$

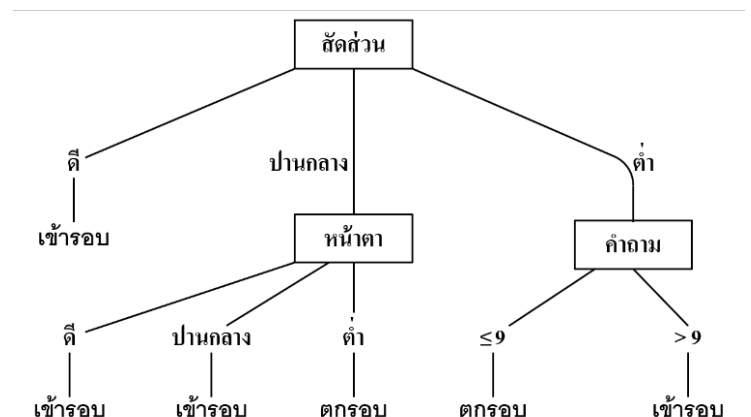
$$H(\text{ผล} | \text{ดี}) = -\left[\frac{0}{2} * \log_2\left(\frac{0}{2}\right) + \frac{2}{2} * \log_2\left(\frac{2}{2}\right)\right] = 0$$

$$H(\text{ผล} | \text{ต่ำ}) = -\left[\frac{1}{5} * \log_2\left(\frac{1}{5}\right) + \frac{4}{5} * \log_2\left(\frac{4}{5}\right)\right] = 0.722$$

$$\text{Gain}(\text{ผล, ประพฤติ}) = 0.592 - \left[\frac{2}{7} * 0 + \frac{5}{7} * 0.722\right] = 0.076$$

$$H(\text{ประพฤติ}) = -\left[\frac{2}{7} * \log_2\left(\frac{2}{7}\right) + \frac{5}{7} * \log_2\left(\frac{5}{7}\right)\right] = 0.863$$

$$\therefore \text{GainRatio}(\text{ผล, ประพฤติ}) = \frac{0.076}{0.863} = 0.088$$



เลือก feature คำถามที่ threshold = 9 ในชั้นที่ 2 node สัดส่วน = ปานกลาง