

AI2 10th week

情報量 (情報エントロピー) ← 情報理論

試行の
 対象は確率事象 $X = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p(x_1) & p(x_2) & \dots & p(x_n) \end{pmatrix}$

(注: x_i は結果, $p(x_i)$ は確率)

$$\forall i \in \{1, \dots, n\} \quad 0 \leq p(x_i) \leq 1, \quad \sum_{i=1}^n p(x_i) = 1.0$$

情報量の定義

X の情報量 $I(X)$

$$I(X) \stackrel{\text{def}}{=} \sum_{i=1}^n -p(x_i) \log_2 p(x_i)$$

(例) コイン投げ $X = \begin{pmatrix} T & H \\ 1/2 & 1/2 \end{pmatrix}$

$$\begin{aligned}
 I(X) &= I(1/2, 1/2) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} \\
 &= +\frac{1}{2} \log_2 2 + \frac{1}{2} \log_2 2 = 1 \text{ (bit)}
 \end{aligned}$$

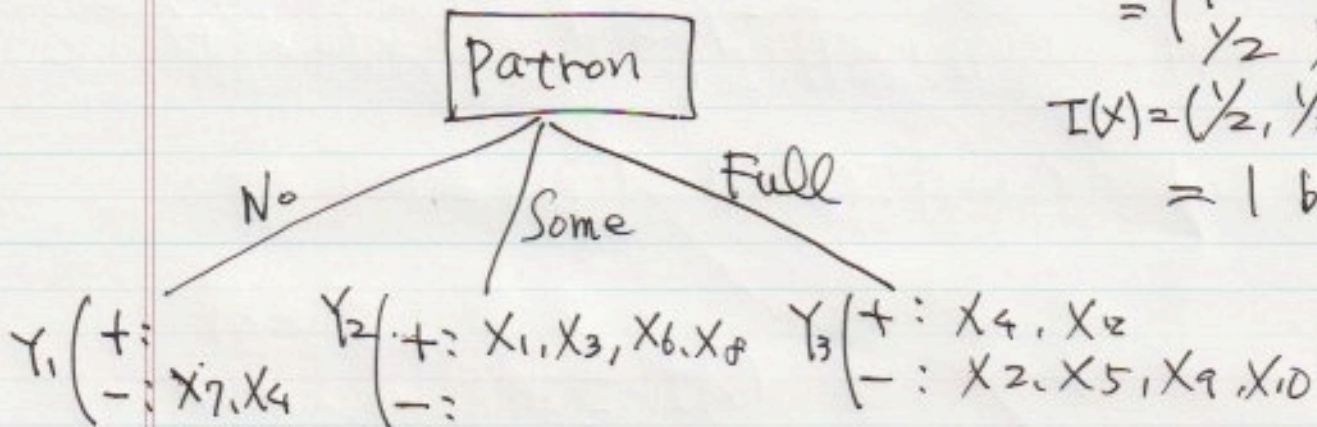
(例) サイコロ投げ $X = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$

$$\begin{aligned}
 I(X) &= I(1/6, 1/6, 1/6, 1/6, 1/6, 1/6) = 6 \left(-\frac{1}{6} \log_2 \frac{1}{6} \right) \\
 &= \log_2 6 = \log_2 2 + \log_2 3 \\
 &= 1 + \log_2 3 \text{ (bit)}
 \end{aligned}$$

$$\left(\begin{array}{l} 0 \log_2 0 = 0 \\ \therefore \lim_{x \rightarrow +0} x \log x = 0 \end{array} \right)$$

X (正例 + : $X_1, X_3, X_4, X_6, X_8, X_{12}$
 (Yes)
 負例 - : $X_2, X_5, X_7, X_9, X_{10}, X_{11}$
 (No)

$$X = \begin{pmatrix} \text{Yes} & \text{No} \\ 6/12 & 6/12 \end{pmatrix} \\ = \begin{pmatrix} \text{Yes} & \text{No} \\ 1/2 & 1/2 \end{pmatrix} \\ I(X) = (1/2, 1/2) \\ = 1 \text{ bit}$$



$$Y_1 = \begin{pmatrix} \text{Yes} & \text{No} \\ 0 & 2/2 \\ = 1 \end{pmatrix} \quad Y_2 = \begin{pmatrix} \text{Yes} & \text{No} \\ 4/4 & 0/4 \\ = 1 \end{pmatrix} \quad Y_3 = \begin{pmatrix} \text{Yes} & \text{No} \\ 2/6 & 4/6 \\ = 1/3 & = 2/3 \end{pmatrix}$$

$$I(Y_1) = -0 \log 0 + 1 \log 1 = 0 \text{ bit} \quad I(Y_2) = 0 \text{ bit}$$

$$I(Y_3) = I(1/3, 2/3) \\ = -\frac{1}{3} \log \frac{1}{3} - \frac{2}{3} \log \frac{2}{3} \\ = \frac{1}{3} \log 3 - \frac{2}{3} (\log 2 - \log 3) \\ = \log 3 - \frac{2}{3} \approx 0.91 \text{ bit}$$

重み付けの必要性

(属性リストの枝の下にある
 テータ数の全データ数に対する
 比率で各情報量を重み付けする)

属性リストの残余情報量

$$\text{Remainder}(\text{Patron}) = \frac{2}{12} \times 0 + \frac{4}{12} \times 0 + \frac{6}{12} (\log 3 - \frac{2}{3}) \\ = \frac{1}{2} (\log 3 - \frac{2}{3}) = \frac{1}{2} \log 3 - \frac{1}{3}$$

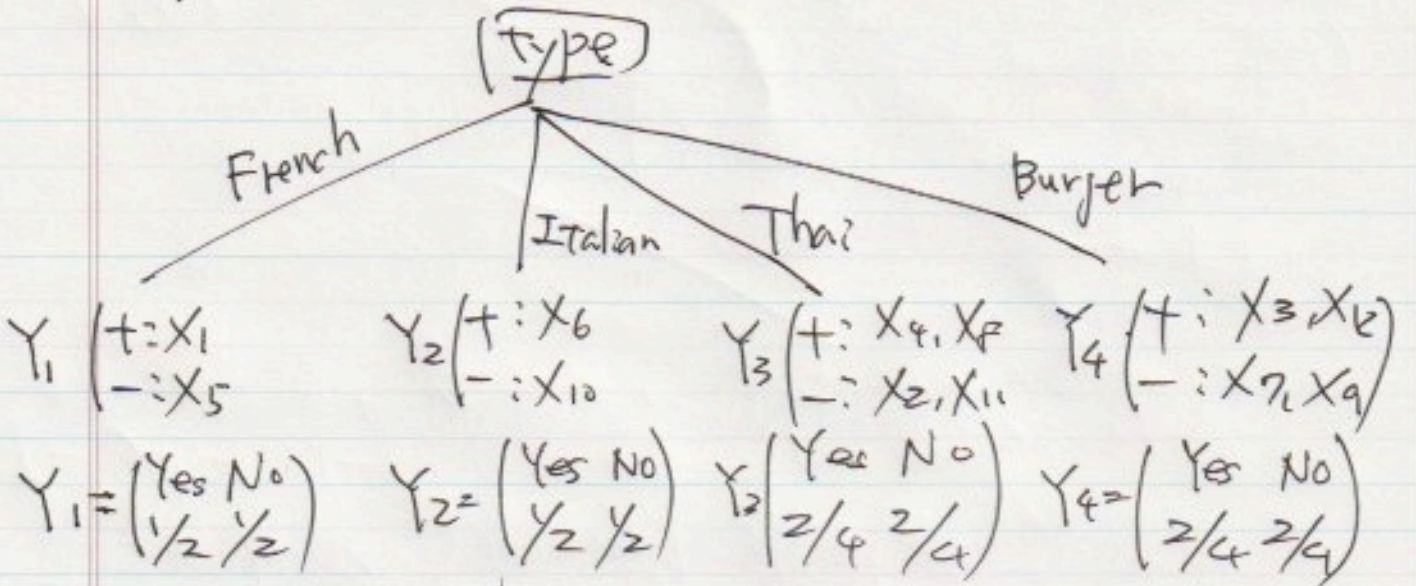
解釈としては、Patron属性リスト $\approx 0.45 \text{ bit}$

(よ) $1 - (\frac{1}{2} \log 3 - \frac{1}{3}) = \frac{4}{3} - \frac{1}{2} \log 3 \approx 0.55 \text{ bit}$
 の情報量が得られた。

$$\text{よ} \approx \text{Gain}(\text{Patron}) = 0.55 \text{ bit}$$

利得

$$X \begin{cases} +: X_1, X_3, X_4, X_6, X_8, X_{12} \\ -: X_2, X_5, X_7, X_9, X_{10}, X_{11} \end{cases} \quad I(X) = 1 \text{ bit}$$



$$I(Y_1) = 1 \text{ bit} \quad I(Y_2) = 1 \text{ bit} \quad I(Y_3) = 1 \text{ bit} \quad I(Y_4) = 1 \text{ bit}$$

$$\begin{aligned} \text{Remainder}(\text{Type}) &= \frac{2}{12} \times 1 + \frac{2}{12} \times 1 + \frac{4}{12} \times 1 + \frac{4}{12} \times 1 \\ &= 1 \text{ bit} \end{aligned}$$

$$\text{Gain}(\text{Type}) = 1 - 1 = 0 \text{ bit}$$