Théorie de l'information MA331

Pierre-Alain TOUPANCE

Esisar
Grenoble INP

2 novembre 2020

Mesure de l'information

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, .
La quantité d'information d'information d'un évènement A que l'on note $I(A)$ doit vérifier :

- $I(A)$ est fonction de $\mathbb{P}(A)$: $I(A) = f(\mathbb{P}(A))$.
- f est décroissante.
- Si $\mathbb{P}(A) = 1$ alors $I(A) = 0$.
- Si A et B sont deux deux évènements indépendants alors $I(A \cap B) = I(A) + I(B)$

Mesure de l'information

Dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ la quantité d'information d'un évènement A est :

$$I(A) = -\log_2(\mathbb{P}(A))$$

L'unité est le Shannon.

La théorie de l'information est un outil pour définir le codage de messages,
généralement la base du logarithme est choisie en fonction du type de codage (base 2 lorsque le codage est binaire).

Information conditionnelle et mutuelle

A et B deux évènements, L'information conditionnelle de A sachant B est :

$$I(A|B) = -\mathrm{Log}_2(\mathbb{P}(A|B) = -\mathrm{Log}_2\left(\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}\right)$$

L'information mutuelle est :

$$\begin{aligned} I(A,B) &= I(A) - I(A|B) \\ &= I(B) - I(B|A) \\ &= I(A) + I(B) - I(A \cap B) \end{aligned}$$

L'information conditionnelle est l'incertitude de A sachant B.
L'information mutuelle est l'information apportée par B sur un évènement A.

Entropie d'une VA

Soit X une variable aléatoire telle que $X(\omega) = \{x_1, x_2, ..., x_n\}$, on appelle entropie de X, l'information moyenne de X, notée $H(X)$

$$H(X) = \mathbb{E}(I(X)) = -\sum_{i=1}^{n} \mathbb{P}(X = x_i)\mathrm{Log}_2(\mathbb{P}(X = x_i))$$

$H(X)$ quantifie l'incertitude, le désordre.

Propriétés de l'entropie

X une VA de loi de probabilité $(x_i, p_i)_{1 \leqslant i \leqslant n}$, on a :

- $H(X) \geqslant 0$
- $H(X) \leqslant \text{Log}_2(n)$
- $H(X) = \text{Log}_2(n)$ dans le cas d'équiprobabilité.

Soient X et Y deux variables aléatoires, $X(\Omega) = \{x_1, x_2, ..., x_p\}$
et $Y(\Omega) = \{y_1, y_2, ..., y_q\}$
On note $\forall (i, j)$ $p(x_i) = \mathbb{P}(X = x_i)$
$p(y_j) = \mathbb{P}(X = y_j)$
et $p_{i,j} = \mathbb{P}(X = x_i \cap Y = y_j)$

Entropie mutuelle

L'entropie mutuelle de X et Y est :

$$H(X, Y) = -\sum_i \sum_j p_{ij} \text{Log}\,(p_{ij})$$

Entropie conditionnelle

L'entropie conditionnelle de X sachant $Y = y_j$ est :

$$H(X|y_j) = -\sum_{i=1}^{p} p(x_i/y_j)\text{Log}\left(p(x_i|y_j)\right)$$

L'entropie conditionnelle de X sachant Y est :

$$H(X|Y) = \sum_{j=1}^{q} p(y_j)H(X|y_j) = \sum_{i=1}^{q}\sum_{j=1}^{p} p_{ij}\log(p(x_i|y_j))$$

On a :

$$\begin{aligned} H(X,Y) &= H(X) + H(Y|X) \\ &= H(Y) + H(X|Y) \end{aligned}$$

Information mutuelle

Soient X et Y deux variables aléatoires, telles que $X(\Omega) = \{x_1, x_2, ..., x_p\}$ et $Y(\Omega) = \{y_1, y_2, ..., y_q\}$
L'information de 2 évènements $(X = x_i)$ et $(Y = y_j)$ est :

$$\begin{aligned} I(x_i, y_j) &= I(x_i) + I(y_j) - I(x_i \cap y_j) \\ &= \mathrm{Log}\left(\frac{p_{ij}}{p(x_i)p(y_j)}\right) \end{aligned}$$

L'information mutuelle de 2 variables aléatoires X et Y est :

$$\begin{aligned} I(X, Y) &= \sum_i \sum_j p_{ij} I(x_i, y_j) \\ &= \sum_i \sum_j p_{ij} \mathrm{Log}\left(\frac{p_{ij}}{p(x_i)p(y_j)}\right) \end{aligned}$$

Propriété

$$
\begin{aligned}
H(X,Y) &= H(X) + H(Y|X) \\
&= H(Y) + H(X|Y)
\end{aligned}
$$

$$
\begin{aligned}
I(X,Y) &= H(Y) - H(Y|X) \\
&= H(X) - H(X|Y)
\end{aligned}
$$

En entrée d'un canal on a des données $X_1, X_2, ...X_p$ sur un alphabet $\mathcal{A} = \{x_1, x_2, ...x_n\}$ et en sortie Y_1, $Y_2, ...Y_p$ sur un alphabet $\mathcal{B} = \{y_1, y_2, ..., y_m\}$.
On dit que le canal est sans mémoire si à tout instant les sorties Y_i ne dépendent pas des entrées et sorties passées.

Canal

Un canal sans mémoire est défini par un triplet :

$$\mathcal{A}, P(X|Y), \mathcal{B})$$

$\mathcal{A}$ l'alphabet d'entrée
$\mathcal{B}$ l'alphabet de sortie.
$P(Y|X)$ la matrice de probabilité de transition
$P(Y|X) = (\mathbb{P}(y_j|x_i))$

Canal Binaire symétrique

Le canal binaire symétrique a une matrice de transition

$$P = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}$$

Capacité d'un canal

La capacité d'un canal est

$$C = sup_{\mathbb{P}_X}(I(X,Y))$$

La capacité est l'information maximale que l'on peut transmettre dans le canal.

Bibliographie