Information partagée par deux variables

Dans l'épisode précédent…

- Notion de valeur / prix de l'information :
  - L'information est-elle facile à obtenir ?
  - Y a-t-il beaucoup de possibilités et d'incertitude ?
  - Mesure de **quantité d'information** (relative à un évènement) : $I = \log_2(N/n)$

Dans l’épisode précédent...

- Notion de valeur / prix de l’information :
  - L’information est-elle facile à obtenir ?
  - Y a-t-il beaucoup de possibilités et d’incertitude ?
  - Mesure de **quantité d’information** (relative à un évènement) : $\mathrm{I} = \log_2(N/n)$

- Quantité moyenne d’information dans une distribution :
  - **Entropie**
  - $H(I) = -\sum_{i \in I} p_i \log_2(p_i)$

Information partagée par deux variables

Nous pouvons considérer deux types d'information :

- La couleur de la boîte $c \in C$ = {bleue, grise, jaune, orange, verte}
- La taille de la boîte $t \in T$ = {petite, moyenne, grande}

Comment définir une probabilité pour une paire de deux variables aléatoires $\{c, t\}$?

Comment définir la quantité d'information et l'entropie pour une paire de deux variables aléatoires $\{c, t\}$?

Rappel de probabilité

- Probabilité conjointe :

Si on considère deux variables aléatoires, il s’agit de la probabilité que leur tirages respectifs génèrent une paire de valeur.

Rappel de probabilité

- Probabilité conjointe :

Si on considère deux variables aléatoires, il s'agit de la probabilité que leur tirages respectifs génèrent une paire de valeur.

Exemple :

1. Variable « couleur de la boîte », avec valeurs {bleue, grise, jaune, rouge}
2. Variable « taille de la boîte », avec valeurs {petite, grande}

Probabilité conjointe $p_{bleue,petite} = \frac{n_{bleue\&petite}}{N}$

Rappel de probabilité

- Probabilité marginale :

Si on considère deux variables aléatoires, il s'agit de la probabilité sur le tirage d'une des variables, quelque soit le tirage de l'autre.

Rappel de probabilité

- Probabilité marginale :

Si on considère deux variables aléatoires, il s'agit de la probabilité sur le tirage d'une des variables, quelque soit le tirage de l'autre.

Exemple :

1. Variable « couleur de la boîte », avec valeurs {bleue, grise, jaune, rouge}
2. Variable « taille de la boîte », avec valeurs {petite, grande}

Probabilité marginale $p_{bleue} = \frac{n_{bleue\&petite}}{N} + \frac{n_{bleue\&grande}}{N}$

Rappel de probabilité

- Table de probabilités :

Présentation « graphique » des probabilités conjointes et marginales

	Petite	**Grande**	**Proba marginale**
Bleue			
Grise			
Jaune			
Rouge			
Proba marginale			

Rappel de probabilité

- Table de probabilités :

Présentation « graphique » des probabilités conjointes et marginales

	Petite	Grande	Proba marginale
Bleue	1/8	1/8	2/8
Grise	2/8	2/8	4/8
Jaune	1/8	0	1/8
Rouge	0	1/8	1/8
Proba marginale	4/8	4/8	

Rappel de probabilité

- Probabilité conditionnelle :

Si on considère deux variables aléatoires, il s’agit de la probabilité sur le tirage d’une des variables quand on sait que l’autre variable a tiré une valeur donnée.

Rappel de probabilité

- Probabilité conditionnelle :

Si on considère deux variables aléatoires, il s’agit de la probabilité sur le tirage d’une des variables quand on sait que l’autre variable a tiré une valeur donnée.

Formule de Bayes :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

ou encore : $P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$

Rappel de probabilité

- Probabilité conditionnelle :

Si on considère deux variables aléatoires, il s'agit de la probabilité sur le tirage d'une des variables quand on sait que l'autre variable a tiré une valeur donnée.

Exemple :

1. Variable « couleur de la boîte », avec valeurs {bleue, grise, jaune, rouge}
2. Variable « taille de la boîte », avec valeurs {petite, grande}

Probabilité conditionnelle $p_{grand|gris} = \frac{p_{grand\&gris}}{p_{gris}}$

Entropie conjointe

Nous pouvons considérer deux types d’information :

- La couleur de la boîte $c \in C$ = {bleue, grise, jaune, orange, verte}
- La taille de la boîte $t \in T$ = {petite, moyenne, grande}

Une paire $\{c, t\}$ a une probabilité conjointe $p_{c,t} = \frac{n_{c,t}}{N}$.

L’**entropie** se généralise donc comme :

$$H(C,T) = -\sum_{c \in C, t \in T} p_{c,t} \log_2(p_{c,t})$$

Rappel : entropie = valeur moyenne de l’information (conjointe)

Entropie conjointe

Quelques propriétés :

- Il est plus dur de connaitre la couleur ET la taille, que la couleur seule (ou la taille seule) :

$$H(C,T) \geq H(C)$$
$$H(C,T) \geq H(T)$$

Rappel : entropie = valeur moyenne de l'information

Entropie conjointe

Quelques propriétés :

- Il est plus dur de connaitre la couleur ET la taille, que la couleur seule (ou la taille seule) :

$$H(C,T) \geq H(C)$$

$$H(C,T) \geq H(T)$$

- Il ne peut *pas* être plus de dur déterminer la couleur ET la taille **ensemble**, **conjointement**, que de déterminer la couleur **seule**, puis la taille **seule**, **séparément** :

$$H(C,T) \leq H(C) + H(T)$$

Quantité d'information commune à deux événements

Rappel : mesure de quantité d’information relative à un évènement (bit) :

$$\mathbb{I} = \log_2\left(\frac{N}{n}\right)$$

$\mathbb{I}$ est une mesure (sous-)additive :

- Deux (types d’)informations renforcent la définition d’un système

Rappel : $H(C,T) \geq H(C)$

- Mais attention : pour « la boîte est bleue » + « la boîte est petite »

$\mathbb{I}_{\text{global}} \neq \mathbb{I}_{\text{bleue}} + \mathbb{I}_{\text{petite}}$ à cause du logarithme

Rappel : $H(C,T) \leq H(C) + H(T)$

Entropie conditionnelle

$$H(T|C) = H(C,T) - H(C)$$

- $H(C,T)$: quantité moyenne d'information, mesure de l'incertitude sur $\{c, t\}$
- $H(T|C)$: incertitude restante sur T si on a déjà payé le prix de l'information sur C

Indépendance

*On dit que **deux variables*** (ou types d'information) *sont **indépendantes** si la réalisation* (ou valeur) *de l'une n'apporte aucune information sur la réalisation de l'autre.* [Wikipedia]

Indépendance

*On dit que **deux variables*** (ou types d'information) *sont **indépendantes** si la réalisation* (ou valeur) *de l'une n'apporte aucune information sur la réalisation de l'autre.* [Wikipedia]

Exemple de variables / types d'information indépendantes :

la couleur et les propriétés nutritives des oursons

Exemple de variables / types d'information non indépendantes :

la couleur et le goût (?) des oursons

Indépendance

*On dit que **deux variables*** (ou types d'information) *sont **indépendantes** si la réalisation* (ou valeur) *de l'une n'apporte aucune information sur la réalisation de l'autre.* [Wikipedia]

Entropie conjointe pour des variables indépendantes :

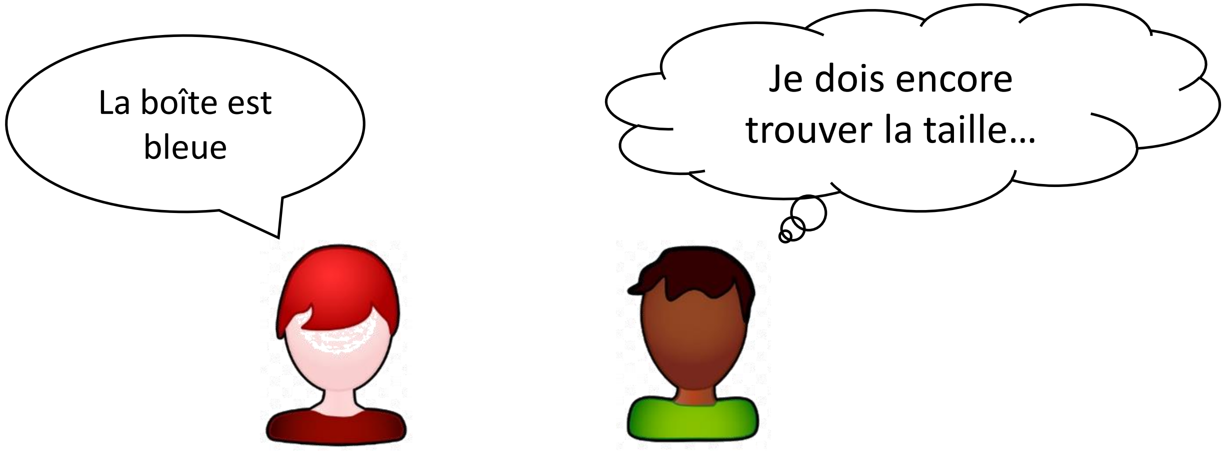
$$H(C, T) = H(C) + H(T)$$

Entropie conditionnelle pour des variables indépendantes :

$$H(T|C) = H(T)$$

Indépendance et information mutuelle

- Si la couleur et la taille sont **indépendantes** :



$$H(C,T) = H(C) + H(T)$$

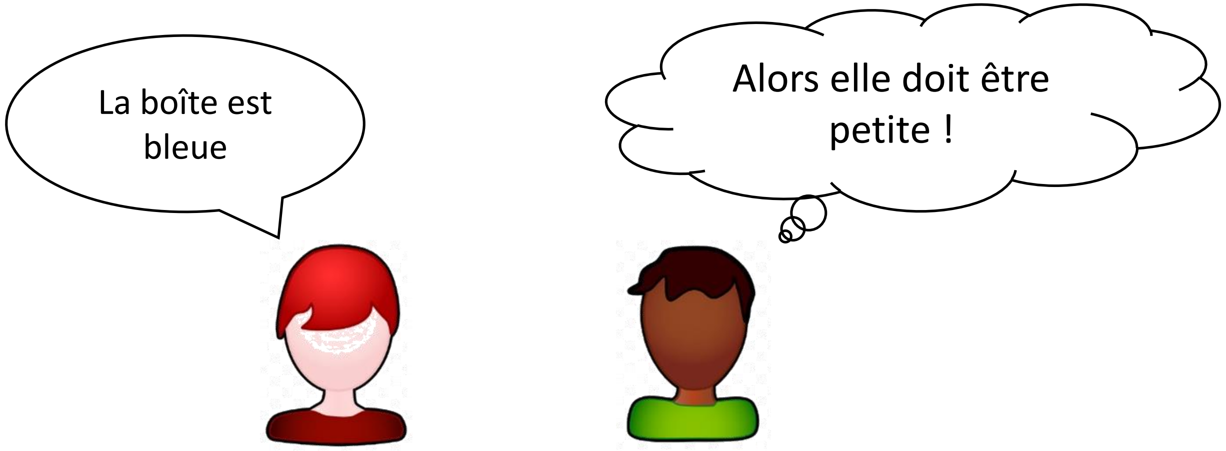
$$H(T|C) = H(T)$$

Pour connaitre {couleur, taille}, il faut encore payer le **prix de l’information « taille »**

Indépendance et information mutuelle

- Si les boîtes bleues sont **toujours** petites :

Variables fortement dépendantes

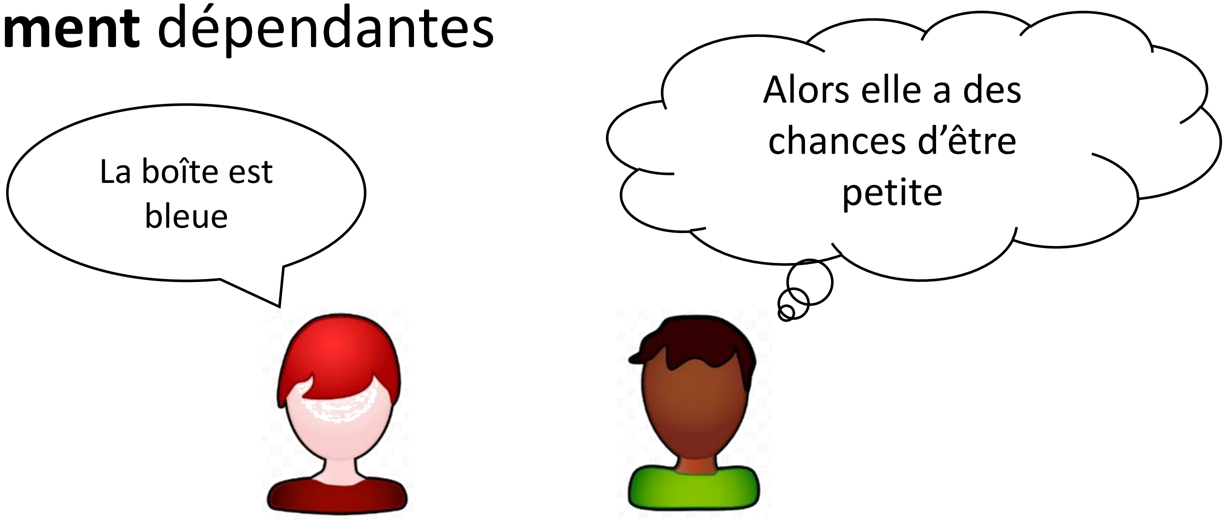


Pour connaitre {couleur, taille}, il n'y a plus rien à payer car la couleur nous a aussi renseigné sur la taille.

Indépendance et information mutuelle

- Si les boîtes bleues sont **souvent** petites :

Variables **partiellement** dépendantes



Pour connaitre {couleur, taille}, il faut encore payer le **prix de l’information « taille »**, mais sa valeur est **réduite** car la couleur nous a **partiellement** renseignés.

Indépendance et information mutuelle

Définitions **Information Mutuelle** :

- Mesure la quantité d'information en commun entre deux variables

- $$IM(C,T) = \underbrace{H(C) + H(T)}_{\text{Prix pour connaitre \{couleur\} et \{taille\} indépendamment l'un de l'autre}} - \underbrace{H(C,T)}_{\text{Prix pour connaitre \{couleur, taille\} en tenant compte de leur (éventuelle) dépendance}}$$

- $$IM(C,T) = \underbrace{H(T)}_{\text{Prix pour connaitre \{taille\}, incertitude sur sa valeur}} - \underbrace{H(T|C)}_{\text{Incertitude restante quand on connait \{couleur\}}} = H(C) - H(C|T)$$

Indépendance et information mutuelle

$$IM(C,T) = H(C) + H(T) - H(C,T)$$

$$IM(C,T) = H(T) - H(T|C) = H(C) - H(C|T)$$

- Variables indépendantes : $IM(C,T) = 0$

Rappel: $H(C,T) = H(C) + H(T)$

- Plus les variables sont dépendantes, et plus $IM(C,T)$ est grande.

- $IM(C,T) \geq 0$

Information mutuelle entre deux signaux / messages

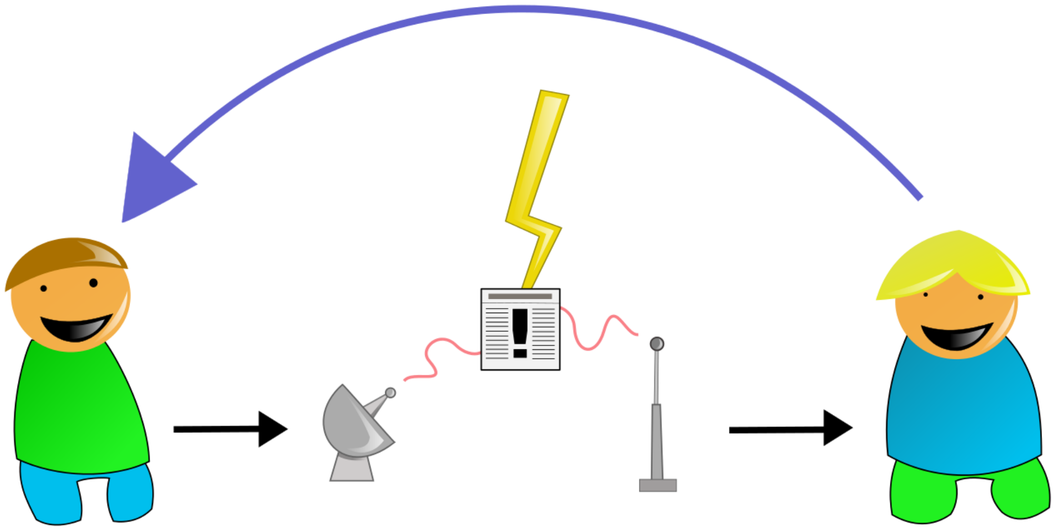
Même chose qu'entre deux variables :

$$IM(X,Y) = H(X) + H(Y) - H(X,Y)$$

$$IM(X,Y) = H(Y) - H(Y|X) = H(X) - H(X|Y)$$

X : signal d'origine

Y : signal après transmission



Information mutuelle

- L'information mutuelle peut elle être négative ?

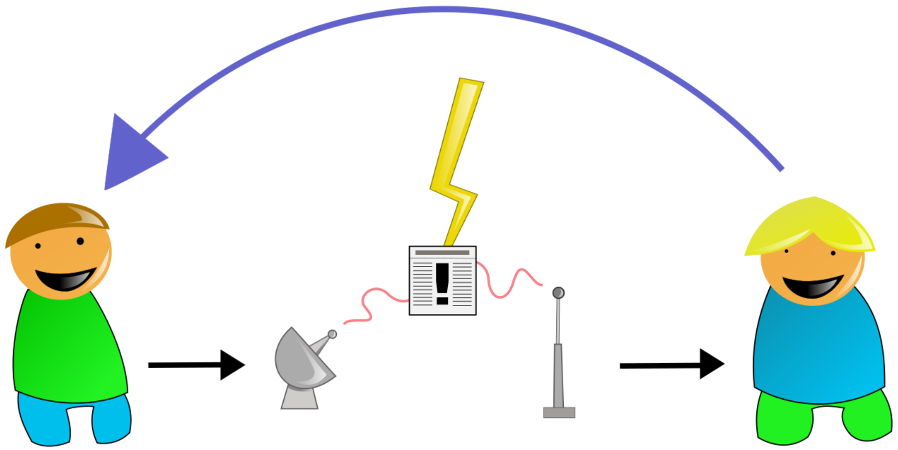
Cas à deux variables :

Pour des évènements donnés : oui ; mais en moyenne, sur les deux variables : non.

Exemple : Un canal de communication reçoit X, le transmet, et on récupère Y en sortie.

Pour un message donné x, il se peut que la sortie y soit ambiguë et ne permette pas de retrouver x.

Mais en moyenne, l'ensemble des messages reçus Y permet de comprendre l'entrée X (pour un canal de transmission raisonnable).



Information mutuelle

- L’information mutuelle peut elle être négative ?

Cas à *n*>2 variables :

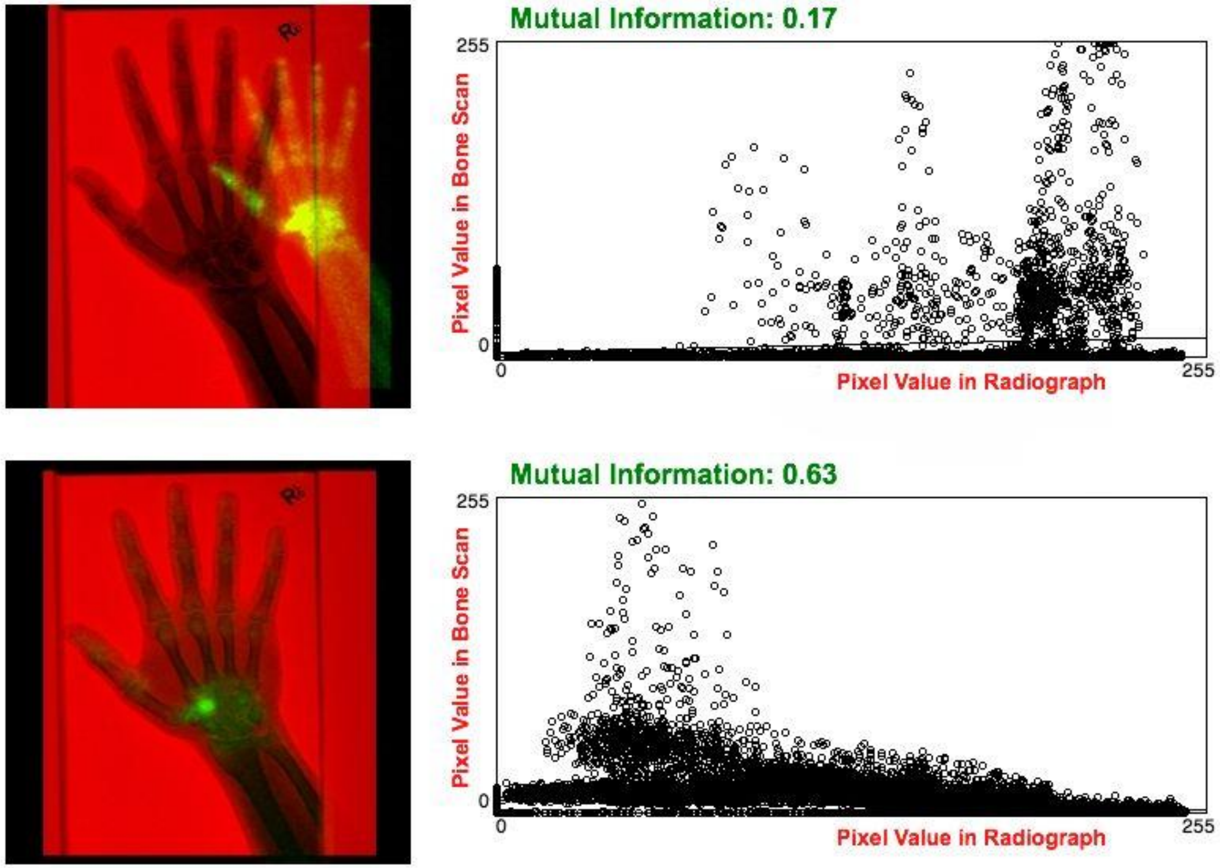
Oui

Il y a $2^n - 1$ degrés de liberté sur les interactions possibles entre variables. Ces interactions sont donc complexes et peuvent aboutir à une information mutuelle totale négative (ou nulle, ou positive).

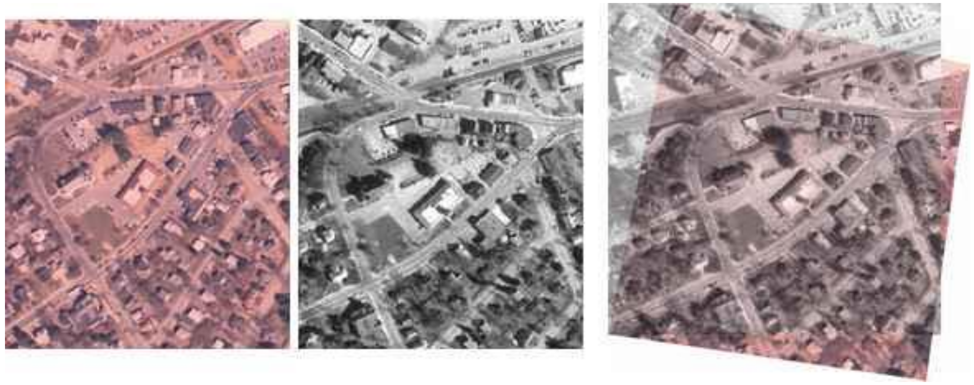
Ce sera abordé dans un prochain cours.

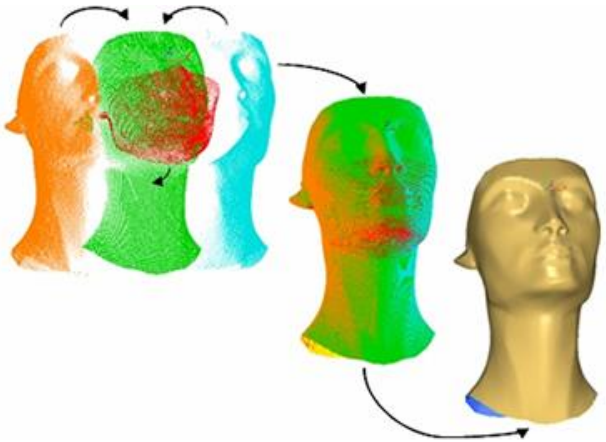
Information mutuelle en science des données

Exemples d’applications : recalage / alignement



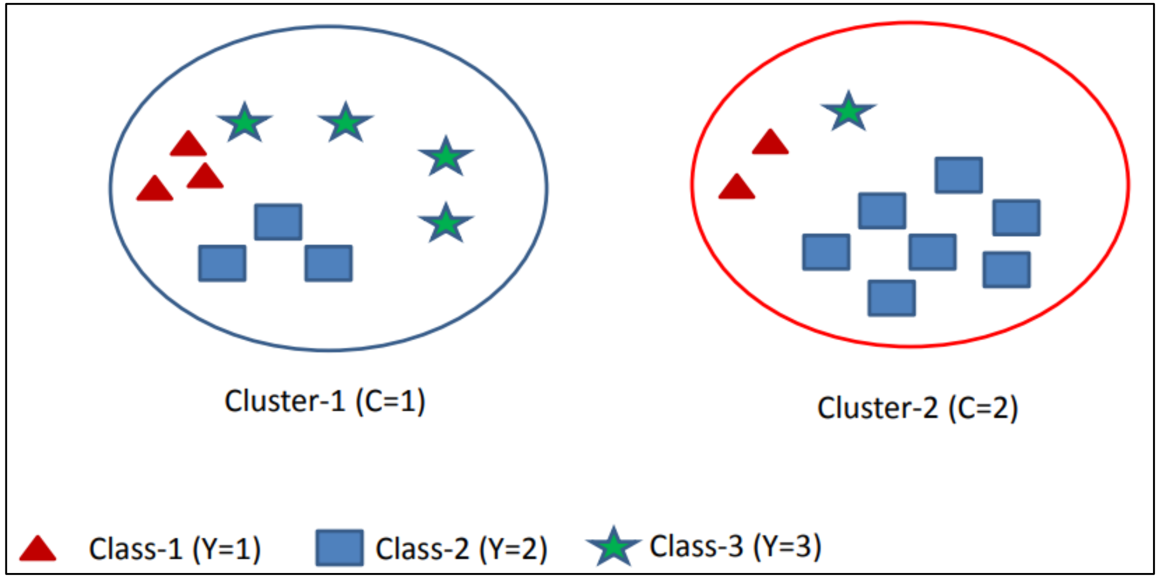
[Wikipedia]

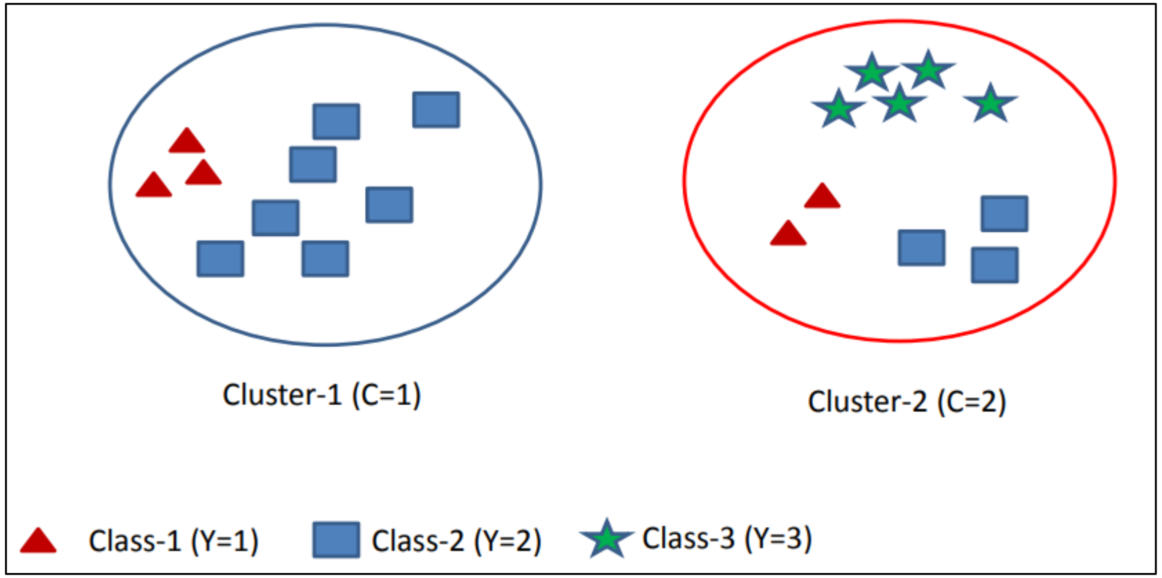




Information mutuelle en science des données

Exemples d'applications : clustering

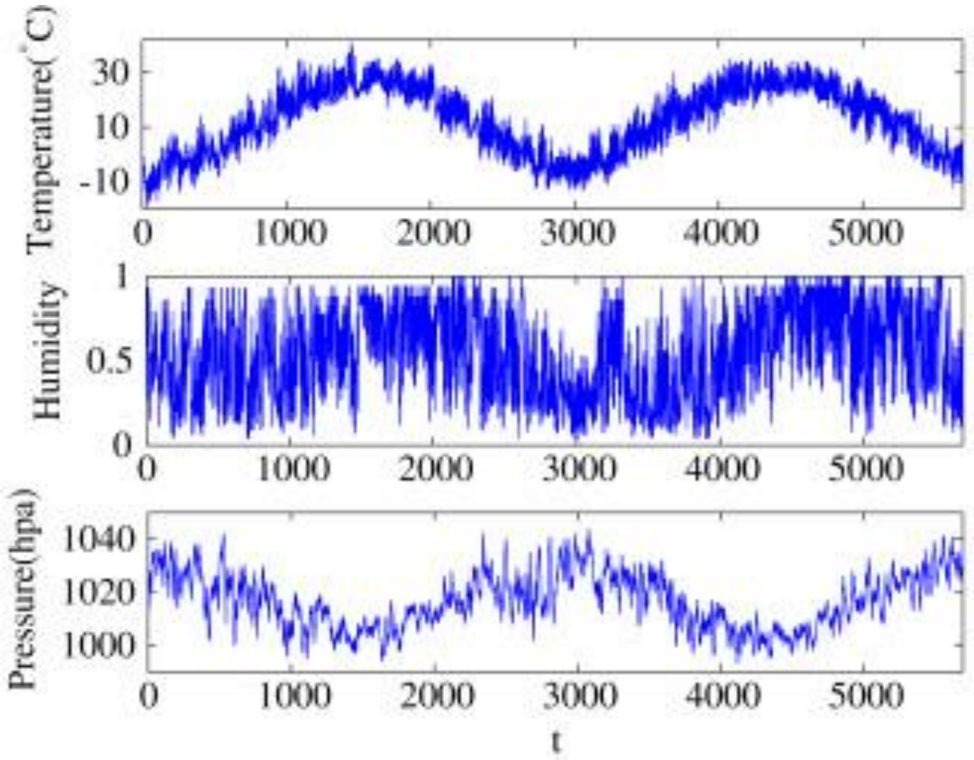


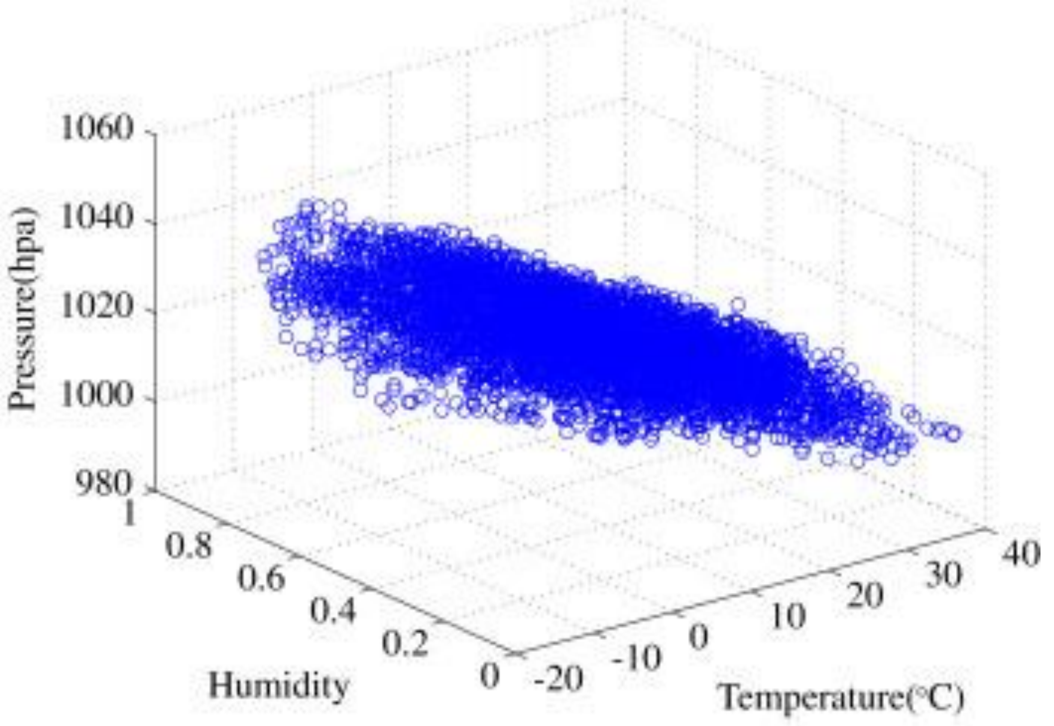


Par exemple : Association de phrases et de contextes dans les moteurs de recherche

Information mutuelle en science des données

Exemples d’applications : synchronisation de séries temporelles / mesure du délai





Information mutuelle en science des données

Exemples d’applications : vérifier la dépendance de phénomènes

Par exemple : Les propriétés des galaxies sont-elles influencées par leur environnement intergalactique ?

[Pandey and Sarkar: *How much a galaxy knows about its large-scale environment?: An information theoretic perspective*, MNRAS 2018]

A creuser en TD et TP…

Utilisation de l'information mutuelle en pratique