

Section	1 ING IDSD
Matière	TP Théorie de l'information et de l'incertain
Enseignante	Trabelsi Nessrine

TP 1: Information propre et Entropie

Objectifs du TP :

- Se familiariser avec les notions “quantité d’information” et “entropie”
- Calcul de l’information propre d’un événement
- Vérification de certaines propriétés de l’information propre
- Calcul de l’entropie d’une source
- Vérification de certaines propriétés de l’entropie
- Calcul de l’information et entropie conjointe de deux sources
- Utilisation de python

Exercice 1 :

Dans un disque dur se trouvent deux dossiers : **Bureau** et **Personnel**.

Supposons que **Bureau** contient 50 fichiers et **Personnel** en contient 150.

1. Quelle est la quantité d’information I_1 associée à l’événement **e1** “le fichier appartient au dossier **Bureau**” ?

On suppose aussi que **Bureau** contient 20 fichiers de type .jpg et 30 de type .txt, et que **Personnel** contient 110 .jpg et 40 .txt.

2. Quelle est la quantité d’information I_2 associée à l’événement **e2** “le fichier est de type .jpg” ?
3. Quelle est la quantité d’information I_3 associée à l’événement **e3** “le fichier est de type .jpg et se trouve dans le dossier **Bureau**” ?
4. Est-ce que $I_3 = I_1 + I_2$?
5. Peut-on en conclure que les deux événements **e1** et **e2** sont dépendants ou indépendants ?
6. En supposant que l’on peut changer le type d’un seul fichier (de .txt en .jpg, ou bien l’inverse), quel fichier choisirez-vous de changer pour que la quantité de l’information I_2 (calculée en 2.) augmente?
7. Calculer l’entropie de la source $A = \{.jpg, .txt\}$ et de la source $B = \{\text{Bureau}, \text{Personnel}\}$.
8. Proposer une nouvelle répartition des fichiers dans les dossiers **Bureau** et **Personnel** de façon à diminuer l’entropie de la source B.

9. Que vaudrait l'entropie de la source B, si les dossiers **Bureau** et **Personnel** contenaient 100 fichiers chacun?
10. Calculer l'entropie conjointe des sources A et B.
11. Comparer en expliquant: $H(A,B)$ et $H(A)+H(B)$

Exercice 2 :

Le but de cet exercice est d'utiliser python afin de calculer et dessiner des courbes de l'information propre et de l'entropie.

Prérequis:

- Télécharger Python et l'installer
 - Télécharger Pycharm (IDE pour python) et l'installer
 - Installer les packages scipy et matplotlib (File / Settings)
1. Ouvrir à l'aide de Pycharm, le fichier "entropy_information_test.py" se trouvant sur le bureau.
 2. Compléter le code afin de calculer l'information propre d'un événement ayant une probabilité P.
 3. Exécuter le code et afficher la courbe de l'information propre en fonction de P.
 4. Compléter le code afin de calculer l'entropie d'une source comportant N événements équiprobables (N entre 1 et 100).
 - Exemple pour remplir un tableau à 3 éléments avec la même valeur 0.2:
`np.full(3, 0.2)`
 - Exemple pour l'utilisation de la fonction entropy de scipy pour une distribution de 2 probabilités de 0.5 et utilisation de log2:
`entropy([1/2, 1/2], None, 2)`
 5. Exécuter le code et afficher la courbe de l'entropie en fonction du nombre des événements équiprobables.
 6. Compléter le code afin de calculer l'entropie d'une source comportant 2 événements avec les probabilités P et 1-P.
 - Exemple pour remplir un tableau à 3 éléments avec les valeurs 0.2, 0.3 et 0.5
`np.array([0.2, 0.3, 0.5])`
 7. Exécuter le code et afficher la courbe de l'entropie en fonction de P.
 8. Supposons qu'on ait une source à 5 événements. Sans calculer l'entropie, ordonnez en expliquant l'entropie de ces 3 distributions:
 Dist1 = [0.2, 0.2, 0.2, 0.2, 0.2]
 Dist2 = [0.05, 0.05, 0.05, 0.05, 0.8]
 Dist3 = [0.1, 0.2, 0.3, 0.15, 0.25]
 9. Vérifier votre réponse en faisant le calcul sur Python. Pour cela, sur la console Python, importer la fonction d'entropie "from scipy.stats import entropy", puis l'utiliser avec les paramètres adéquats.

