

Reconnaissance de chiffres en Python

Objectif:

Le but des exercices sur cette feuille est de réussir à créer un algorithme en Python permettant de reconnaître un chiffre à travers une image.

On implémentera ainsi différentes méthodes, plus ou moins simples, pour ce qui est des images on pourra utiliser cette [archive](#). (Pour l'extraire: `tar -xzf chiffres.tar.gz`)

Partie 1: Les k plus proches voisins

Dans cette partie on va considérer qu'une image de largeur l et de hauteur h peut être représentée comme un **point** appartenant à $[0, 1]^{l \cdot h}$. Ainsi en calculant la **distance** entre deux points, on peut avoir une mesure de la ressemblance d'une image à une autre et utiliser la méthode des [k plus proches voisins](#).

1. Tout d'abord on aura besoin d'une fonction pour lire une image depuis un fichier avec son nom, on pourra par exemple utiliser la classe `Image` de la librairie `PIL`.

(Pour l'installer: `pip3 install Pillow`)

2. Ensuite il faut une fonction pour passer d'une **image** à un point, on pourra le représenter comme une matrice de pixel de taille $l \cdot h$, par exemple avec la librairie `numpy`.

On trouvera aussi un moyen pour transformer les couleurs en une valeur dans $[0, 1]$.

3. Puis on va créer une fonction pour calculer la distance entre deux **points** de $[0, 1]^{l \cdot h}$. On pourra appliquer la formule de la distance euclidienne de la même manière que dans l'espace, ou une autre si vous trouvez plus intéressant.

(Remarque: pour les performances on pourra enlever la racine carré dans la formule de la distance euclidienne qui n'est pas nécessaire ici)

4. Enfin on crée une fonction pour reconnaître un chiffre sur une **image**, pour ça on peut trouver les k images les plus ressemblantes parmi les références et en déduire le chiffre.

(Remarque: pour lister tous les fichiers d'un répertoire on pourra utiliser la fonction `os.listdir(repertoire)` après avoir fait `import os`)

Partie 2: Un réseau neuronal

TODO