

TP6 : méthodes d'intelligence artificielle

1 Algorithme des k-moyennes

Cet algorithme a été décrit dans le cours

Programmer cet algorithme et le tester sur l'exemple ci-dessus.

Les données ainsi qu'une trame suggérant certaines fonctions auxiliaires sont fournies dans le fichier `Kmoy-Acompleter.py`.

2 Application de l'algorithme kNN à la reconnaissance de chiffres

Dans cette partie on applique l'algorithme kNN dans une situation plus proche d'une situation réelle.

On va pour cela se servir d'un jeu de données fourni par la bibliothèque `scikit-learn`. Cette bibliothèque est dédiée aux algorithmes d'apprentissage. Nous n'allons pas utiliser ses fonctionnalités, mais juste un jeu de données qu'elle fournit.

On importe tout d'abord ce jeu de données (voir le fichier `kNN_chiffres_Acompleter.py` pour les détails). On en tire un tableau numpy à 3 dimensions représentant un ensemble d'images en noir et blanc figurant des chiffres.

On va séparer cet ensemble d'images en une partie qui servira à l'apprentissage et une autre qui servira aux tests.

Il faut adapter à cette situation l'algorithme kNN vu précédemment. En particulier :

- la distance doit être modifiée (on prend la distance euclidienne dans un espace où chaque pixel est une coordonnée) ;
- il est préférable ici de ne pas utiliser un tri mais une fonction créée spécialement pour sélectionner les k images du jeu d'apprentissage qui sont les plus proches de l'image que l'on souhaite classifier.