

Informatique S6 - CC2

Le but de ces exercices est de construire une méthode d'apprentissage de type "clustering". Etant donné un nuage de points, elle consiste à regrouper les points par paquets. Ces points appartiennent à un espace abstrait : cela peut être des points de \mathbb{R}^n ou bien des objets complexes comme des voitures, des logements, des pays, etc.

Exercice 1 (Nuage de points (12.25 points)).

Pour commencer on va construire une classe "cloud" qui modélise un nuage de points de type `T`.

Question (Classe (1.5 points)).

Construire une classe `cloud` paramétrée par une classe `T` et contenant trois attributs :

- un entier qui correspond au nombre de points,
- un pointeur de type `T` qui contiendra les points,
- un pointeur de fonction double `(*metric)(T a, T b)` qui contiendra une fonction de distance entre deux points.

Question (Constructeurs I (2.25 points)). Coder le constructeur par défaut, le destructeur et un constructeur qui prend le nombre de points en entrée et initialise la mémoire du tableau.

Question (Constructeurs II (3.5 points)). Coder le constructeur par copie et surcharger l'opérateur `=`.

Question (Accesseurs/mutateurs (5.0 points)). Coder les accesseurs/mutateurs :

- un accesseur pour la taille du nuage de points,
- un accesseur au point `i`,
- un mutateur pour le point `i`,
- un mutateur pour le pointeur de fonction `metric`
- un accesseur qui prend deux points et qui renvoie un double qui est le résultat de `metric`.

Exercice 2 (K-moyenne (9.5 points)).

L'algorithme des K -moyennes consiste à définir K groupes à partir de K points donnés C_1, C_2, \dots, C_K , appelés centres des groupes. Pour cela, on associe à chaque point du nuage X_i le centre C_j le plus proche. On a ainsi formé K groupes.

Ensuite on met à jour les centres en calculant les barycentres des groupes, puis on reforme les groupes en associant à chaque point le nouveau centre le plus proche. On continue ainsi jusqu'à convergence. On ne codera pas cet algorithme intégralement dans l'exercice.

Pour commencer on va construire une classe `K_mean` qui modélise une K moyenne de type `T`.

Question (Classe (2.0 points)).

Construire un template de classe `K_mean` de paramètre la classe `T` et un entier `K` (le nombre de groupes) qui contient trois attributs :

- un nuage de points `cl`,
- un pointeur sur les entiers `gp_by_point` qui contiendra le numéro du groupe associé à chaque point,
- un tableau statique de taille `K` et de type `T` appelé `grp` qui contiendra les centres de groupes.

Question (Constructeur I, destructeur (1.0 points)).

Coder le constructeur par défaut et le destructeur.

Question (Constructeur II (1.0 points)).

Coder un constructeur qui prend un nuage de points, le copie et initialise la mémoire du tableau `gp_by_point` de taille le nombre de point dans le nuage.

Question (Constructeur III (1.5 points)).

Coder le constructeur par copie (mais pas le =).

Question (Méthode (4.0 points)).

Coder une fonction `initgroup` qui :

- prend les `K` premier points du nuage et les copie dans le tableaux des centres `grp`
- parcourt tous les points du nuage et fait :
 - déclare un double `dis=1000` et un entier `index=0`,
 - parcourt les centres de groupe,
 - si la distance (fonction `metric` du nuage) entre le point courant et le groupe courant est plus petite que `dis`, on met cette distance dans `dis` et le numéro de centre dans `index`.
- on met la valeur d' "index" dans "gp_by_point" pour le point courant.