# Série de TD N°4 Clustering 1 (K-means et K-medoids)

## Exercice1 :

On désire classifier l'ensemble des points suivants en trois classes: A1(2, 10), A2(2, 5), A3(8, 4), B1(5, 8), B2(7, 5), B3(6, 4), C1(1, 2), C2(4, 9)

On suppose initialement que les points A1, B1 et C1 sont choisis comme centres. Utiliser l'algorithme K-means pour déterminer:

1. Les trois centres calculés après la première itération.
2. Les trois classes finales résultant de l'application de l'algorithme.

## Exercice2 :

Soient 10 points (a,b,c,d,e,f,g,h,i,j) du plan dont les coordonnées sont les colonnes du tableau suivant:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| a | b | c | d | e | f | g | h | i | j |
| 1 | 2 | 3 | 5 | 4 | 6 | 8 | 10 | 12 | 10 |
| 2 | 5 | 2 | 4 | 7 | 7 | 3 | 4 | 1 | 0 |

Donner les partitions stabilisées ps1 et ps2 correspondants:

pour ps1 aux centres de départ {a,b,i} pour ps2 aux centres de départ {a,d,i}

On fera fonctionner l'algorithme directement sur le graphique joint. Que peut-on conclure?

## Exercice3:

On considère les 6 points M1 = (−2, 3), M2 = (−2, 1), M3 = (−2,−1), M4 = (2,−1), M5 = (2, 1)

et M6 = (1, 0).

En supposant que les deux premiers points M1 et M2 sont les centres initiaux, décrire l’algorithme k-medoids pour la répartition de ces points.

# Série de TD N°4- Corrigé Clustering 1 (K-means et K-medoids)

## Exercice1 :

On désire classifier l'ensemble des points suivants en trois classes: A1(2, 10), A2(2, 5), A3(8, 4), B1(5, 8), B2(7, 5), B3(6, 4), C1(1, 2), C2(4, 9)

On suppose initialement que les points A1, B1 et C1 sont choisis comme centres. Utiliser l'algorithme K-means pour déterminer:

1. Les trois centres calculés après la première itération.
2. Les trois classes finales résultant de l'application de l'algorithme.

## Réponse

1. Les trois centres calculés après la première itération.

Initialisation : G1 = {A1}, G2 = {B1}, G3 = {C1} avec c1 = A1, c2 = B1 et c3=C1

### Itération 1 :

Pour A2 : d(A2,c1) = (|2-2|²+|5-10|²)=5, d(A2,c2) = (|2-5|²+|5-8|²)= 4.24, d(A2,c3) = (|2-1|²+|5-2|²)= 3.16 alors G3 = G3 + A2

Pour A3 : d(A3,c1) = (|8-2|²+|4-10|²)=8.48, d(A3,c2) = (|8-5|²+|4-8|²)= 5, d(A3,c3) = (|8-1|²+|4-2|²)= 7.28 alors G2 = G2 + A3

Pour B2 : d(B2,c1) = (|7-2|²+|5-10|²)=5, d(B2,c2) = (|7-5|²+|5-8|²)= 3.60, d(B2,c3) = (|7-1|²+|5-2|²)= 6.70 alors G2 = G2 + B2

Pour B3 : d(B3,c1) = (|6-2|²+|4-10|²)=7.21, d(B3,c2) = (|6-5|²+|4-8|²)= 4.12, d(B3,c3) = (|6-1|²+|4-2|²)= 5.38 alors G2 = G2 + B3

Pour C2 : d(C2,c1) = (|4-2|²+|9-10|²)=2.23, d(C2,c2) = (|4-5|²+|9-8|²)= 1.41, d(C2,c3) = (|4-1|²+|9-2|²)= 7.61 alors G2 = G2 + C2

Les trois clusters après la première itération sont :

G1 = {A1}, G2 = {B1,A3,B2,B3,C2}, G3 = {C1,A2}

Les trois centres après la première itération sont :

c1 = A1, c2 = ((8+5+7+6+4)/5, (4+8+5+4+9)/5)=(6,6), c3 = ((1+2)/2, (5+2)/2)=(1.5,3.5)

1. Les trois classes finales résultant de l'application de l'algorithme.

### Itération 2 :

G1 = {}, G2 = {}, G3 = {} avec c1 = A1, c2 = (6,6) et c3=(1.5,3.5)

Pour A2 : d(A2,c1) = (|2-2|²+|5-10|²)=5, d(A2,c2) = (|2-6|²+|5-6|²)= 4.12, d(A2,c3) = (|2-1.5|²+|5-3.5|²)= 1.58

alors G3 = G3 + A2

Pour A3 : d(A3,c1) = (|8-2|²+|4-10|²)=8.48, d(A3,c2) = (|8-6|²+|4-6|²)=2.82, d(A3,c3) = (|8-1.5|²+|4-3.5|²)=6.51

alors G2 = G2 + A3

Pour B1 : d(B1,c1) = (|5-2|²+|8-10|²)=3.60, d(B1,c2) = (|5-6|²+|8-6|²)=2.23, d(B1,c3) = (|5-1.5|²+|8-3.5|²)=5.70

alors G2 = G2 + B1

Pour B2 : d(B2,c1) = (|7-2|²+|5-10|²)=5, d(B2,c2) = (|7-6|²+|5-6|²)=1.41, d(B2,c3) = (|7-1.5|²+|5-3.5|²)=5.70

alors G2 = G2 + B2

Pour B3 : d(B3,c1) = (|6-2|²+|4-10|²)=7.21, d(B3,c2) = (|6-6|²+|4-6|²)=2, d(B3,c3) = (|6-1.5|²+|4-3.5|²)=4.52

alors G2 = G2 + B3

Pour C1 : d(C1,c1) = (|1-2|²+|2-10|²)=8.06, d(C1,c2) = (|1-6|²+|2-6|²)=6.40, d(C1,c3) = (|1-1.5|²+|2-3.5|²)=1.58

alors G3 = G3 + C1

Pour C2 : d(C2,c1) = (|4-2|²+|9-10|²)=2.23, d(C2,c2) = (|4-6|²+|9-6|²)=3.6, d(C2,c3) = (|4-1.5|²+|9-3.5|²)=6.04

alors G1 = G1 + C2

Les trois clusters après la première itération sont :

G1 = {A1,C2}, G2 = {A3,B1,B2,B3}, G3 = {A2,C1}

Les trois centres après la première itération sont :

c1 = ((2+4)/2, (10+9)/2)=(3,9.5), c2 = ((8+5+7+6)/4, (4+8+5+4)/4)=(6.5,5.25), c3 = ((1+2)/2, (5+2)/2)=(1.5,3.5)