aufgabenblatt: K-Means-Clustering mit Manhattan-Distanz

Einführung

In dieser Übung wirst du den k-Means-Clustering-Algorithmus anwenden, um eine Menge von Datenpunkten in Cluster zu gruppieren. Dabei verwendest du die **Manhattan-Distanz** zur Berechnung der Distanzen zwischen Punkten und Zentroiden.

Aufgabe

1. k-Means-Algorithmus durchführen:

- (a) Berechne die Manhattan-Distanzen zwischen jedem Datenpunkt und den Zentroiden.
- (b) Weise jeden Datenpunkt dem nächstgelegenen Zentroiden zu und markiere die Datenpunkte entsprechend ihrer Clusterzugehörigkeit, indem du sie mit einem Symbol umschließt:
 - O für Cluster 1
 - △ für Cluster 2
 - □ für Cluster 3
- (c) Aktualisiere die Position der Zentroiden, indem du den Mittelwert der X- und Y-Koordinaten der zugewiesenen Punkte berechnest.
- (d) Wiederhole die Schritte a) bis c), bis sich die Clusterzuweisungen nicht mehr ändern.

Datenpunkte und initiale Zentroiden

Punkt	\mathbf{X}	\mathbf{Y}
a	2	10
b	2	5
c	8	4
d	5	8
e	7	5
f	6	4
g	1	2
h	4	9
i	6	2
j	3	3
k	5	6
l	9	7

- \bigcirc **Z**entroid Z_1 : a(2,10)
- \triangle **Zentroid** Z_2 : d(5,8)
- \square **Z**entroid Z_3 : g(1,2)

	Y
	11 \uparrow Z_1
	10 a
	9 + 7
Iteration 1	8 +
O Z_1 : (2, 10)	7 + d
· · · · ·	6
$\triangle Z_2$: (5, 8)	$b \mid b \mid e$
$\square Z_3$: $(1, 2)$	4
2 3. (1, 2)	$\frac{3}{Z_3}$
	$\frac{2}{q}$
	1
	0 1 2 3 4 5 6 7 8 9
	Y
	11
	10
	9 a
Iteration 2	8
	7
$\bigcirc Z_1$:	6
$\triangle Z_2$:	5
	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
$\square Z_3$:	$\frac{1}{2}$
	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$
	0
	0 1 2 3 4 5 6 7 8 9
	Y •
	11 7
Iteration 3	9 h
	7 d
$\bigcirc Z_1$:	6
$\triangle Z_2$:	5 k
	4 b e
$\square Z_3$:	$_{3}$ f c
	$_{2}$
	$_{1}$ g i
	0
	0 1 2 3 4 5 6 7 8 9
	Y

11 10

9

8

6

3

 \overline{a}

b

h

e

f

Punkt	$\mathbf{D}(Z_1)$	$\mathbf{D}(Z_2)$	$\mathbf{D}(Z_3)$	Cluster
a	0	5	9	$\bigcirc Z_1$
b	5	6	4	$\square Z_3$
С	10	7	10	ΔZ_3
d				
e				
f				
g				
h				
i				
j				
k				
l				
D 14	D(Z)	D(Z)	D(Z)	C 1 4
Punkt	$D(Z_1)$	$D(Z_2)$	$\mathbf{D}(Z_3)$	Cluster

Punkt	$\mathbf{D}(Z_1)$	$\mathbf{D}(Z_2)$	$\mathbf{D}(Z_3)$	Cluster
a				
b				
c				
d				
e				
f				
g				
h				
i				
j				
k				
l				

Punkt	$\mathbf{D}(Z_1)$	$\mathbf{D}(Z_2)$	$\mathbf{D}(Z_3)$	Cluster
a				
b				
С				
d				
e				
f				
g				
h				
i				
j				
k				
1				

Punkt	$\mathbf{D}(Z_1)$	$\mathbf{D}(Z_2)$	$\mathbf{D}(Z_3)$	Cluster
a				
b				
С				
d				
e				
f				
g				
h				
i				
j				
k				
1				

Iteration 4

 $\square Z_3$: