

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:

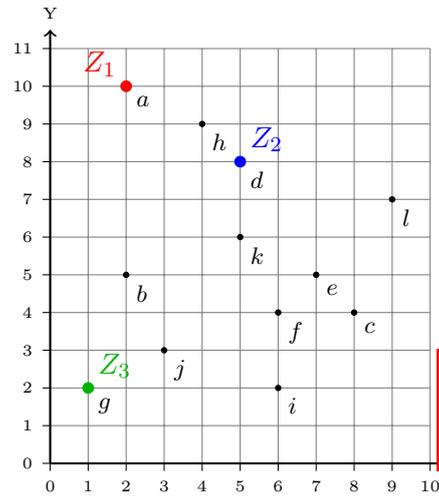
- Berechne die Manhattan-Distanzen zwischen jedem Datenpunkt und den Zentroiden.
- Weise jeden Datenpunkt dem nächstgelegenen Zentroiden zu und markiere die Datenpunkte entsprechend ihrer Clusterzugehörigkeit, indem du sie mit einem Symbol umschließt:
 -  für Cluster 1
 -  für Cluster 2
 -  für Cluster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

Datenpunkte und initiale Zentroiden

Punkt	X	Y
<i>a</i>	2	10
<i>b</i>	2	5
<i>c</i>	8	4
<i>d</i>	5	8
<i>e</i>	7	5
<i>f</i>	6	4
<i>g</i>	1	2
<i>h</i>	4	9
<i>i</i>	6	2
<i>j</i>	3	3
<i>k</i>	5	6
<i>l</i>	9	7

-  Cluster 1, Zentroid $Z_1: a(2, 10)$
-  Cluster 2, Zentroid $Z_2: d(5, 8)$
-  Cluster 3, Zentroid $Z_3: g(1, 2)$

Punkt	D(Z ₁)	D(Z ₂)	D(Z ₃)	Cluster
a	0	5	9	○ Z ₁
b	5	6	4	□ Z ₃
c	12	7	9	△ Z ₂
d	5	0	10	△ Z ₂
e	10	5	9	△ Z ₂
f	10	5	7	△ Z ₂
g	9	10	0	□ Z ₃
h	3	2	10	△ Z ₂
i	12	7	5	□ Z ₃
j	8	7	3	□ Z ₃
k	7	2	8	△ Z ₂
l	10	5	13	△ Z ₂



Zentroiden Iteration 1

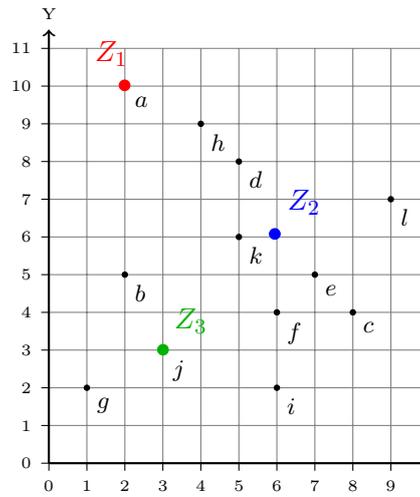
○ Z₁: 2,10

△ Z₂: 6,6

□ Z₃: 3,3

Cluster 1: ['a']
Cluster 2: ['c', 'd', 'e', 'f', 'h', 'k', 'l']
Cluster 3: ['b', 'g', 'i', 'j']

Punkt	D(Z ₁)	D(Z ₂)	D(Z ₃)	Cluster
a				
b				
c				
d				
e				
f				
g				
h				
i				
j				
k				
l				



Zentroiden Iteration 2

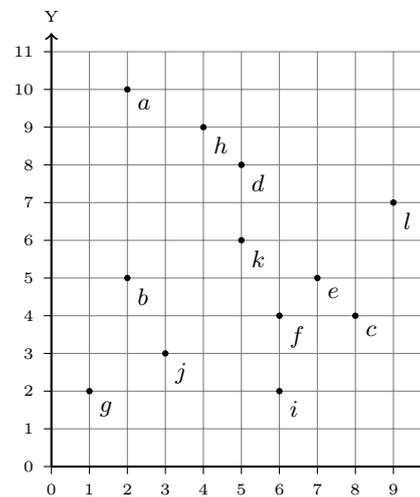
○ Z₁: 3,10

△ Z₂: 7,5

□ Z₃: 2,3

Cluster 1: ['a', 'h']
Cluster 2: ['c', 'd', 'e', 'f', 'i', 'k', 'l']
Cluster 3: ['b', 'g', 'j']

Punkt	D(Z ₁)	D(Z ₂)	D(Z ₃)	Cluster
a				
b				
c				
d				
e				
f				
g				
h				
i				
j				
k				
l				



Zentroiden Iteration 3

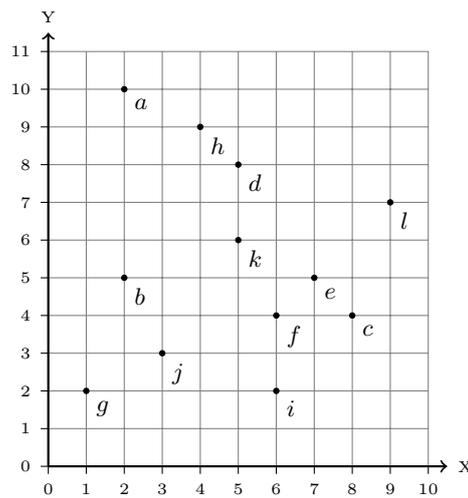
○ Z₁: 4,9

△ Z₂: 7,5

□ Z₃: 2,3

Cluster 1: ['a', 'd', 'h']
Cluster 2: ['c', 'e', 'f', 'i', 'k', 'l']
Cluster 3: ['b', 'g', 'j']

Punkt	D(Z ₁)	D(Z ₂)	D(Z ₃)	Cluster
a				
b				
c				
d				
e				
f				
g				
h				
i				
j				
k				
l				



Zentroiden Iteration 4

○ Z₁:

△ Z₂:

□ Z₃: