**Задания**

**для решения задачи кластеризации методом k-средних**

* **Даны** шесть объектов, каждый определяется двумя признаками X и Y.
* **Задание:** Разбить объекты на три кластера методом k-средних. Первоначально первые три объекта образуют начальные кластеры, метрика – квадрат евклидова расстояния.

**Вариант 3.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Объект | A | B | C | D | E | F |
| Признак X | 0 | -1 | 7 | 5 | 6 | 3 |
| Признак Y | 2 | 0 | 3 | -3 | 2 | -2 |

**Итерация №1**

Эталонные точки -  центроиды кластеров:

e1 = B = (-1; 0)

e2 = E = (6; 2)

e3 = F = (3; -2)

Сравниваем расстояние от точки D до эталонных точек.

D(D,e1) = **√** (5 – (-1))2 +(-3 - 0)2 = **√** 45

D(D,e2) = **√** (5 - 6)2 + (-3 – 2)2 = **√** 26

D(D,e3) = **√** (5 - 3)2 + (-3 –(-2))2 = **√** **5**  
Минимальным является расстояние d(D;e3)

Сравниваем расстояние от точки E до эталонных точек.

D(А,e1) = **√** (0 – (-1))2 + (2 – 0)2 = **√5**

D(А,e2) = **√** (0 - 6)2 + (2 – 2)2 = 6

D(А,e3) = **√** (0 - 3)2 + (2 – (-2))2 = 5

Минимальным является расстояние d(А,e1)

Сравниваем расстояние от точки F до эталонных точек.

D(С,e1) = **√** (7 – (-1))2 +(3 - 0)2 = 10

D(С,e2) = **√** (7 - 6)2 +(3 – 2)2 = **√ 2**

D(С,e3) = **√** (7 - 3)2 +(3 - (-2))2 = **√** 41

Минимальным является расстояние d(С;e2)

Выписываем в таблице принадлежность объектов кластерам

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Объект | A | B | C | D | E | F |
| Номер кластера | 1 | 1 | 2 | 3 | 2 | 3 |

|  |  |  |
| --- | --- | --- |
| e1 | e2 | e3 |
| A,B | C,E | D,F |

Пересчитываем значения для эталонных точек.

e1 = (-0,5; 1)

e2 = (6,5; 2,5)

e3 = (4; -2,5)

**Итерация №2**

Эталонные точки -  центроиды кластеров:

e1 = (-0,5; 1)

e2 = (6,5; 2,5)

e3 = (4; -2,5)

Сравниваем расстояние от точки A(0; 2) до эталонных точек.

D(A,e1) = **√** (0 – (-0,5))2 +(2 - 1)2 = **√ 1,25**

D(A,e2) = **√** (0 - 6,5)2 +(2 - 2,5)2 = **√** 42,5

D(A,e3) = **√** (0 – 4)2 +(2 – (-2,5))2 = **√** 36,25

Минимальным является расстояние d(A,e1)

Сравниваем расстояние от точки B(-1; 0) до эталонных точек.

D(B,e1) = **√** (-1 – (-0,5))2 +(0 - 1)2 = **√ 1,25**

D(B,e2) = **√** (-1- 6,5)2 +(0 - 2,5)2 = **√** 62,5

D(B,e3) = **√** (-1 – 4)2 +(0 – (-2,5))2 = √ 31,25

Минимальным является расстояние d(B,e1)

Сравниваем расстояние от точки C(7; 3) до эталонных точек.

D(C,e1) = **√** (7 – (-0,5))2 +(3 - 1)2 = **√** 60,25

D(C,e2) = **√** (7- 6,5)2 +(3 - 2,5)2 = **√ 0,5**

D(C,e3) = **√** (7 – 4)2 +(3 – (-2,5))2 = **√** 39,25

Минимальным является расстояние d(C,e2)

Сравниваем расстояние от точки D(5; -3) до эталонных точек.

D(D,e1) = **√** (5 – (-0,5))2 +(-3 - 1)2 = **√** 46,25

D(D,e2) = **√** (5 - 6,5)2 +(-3 - 2,5)2 = √ 32,5

D(D,e3) = **√** (5 – 4)2 +(-3 – (-2,5))2 = **√ 1,25**

Минимальным является расстояние d(D,e3)

Сравниваем расстояние от точки E(6; 2) до эталонных точек.

D(E,e1) = **√** (6– (-0,5))2 +(2 - 1)2  = √ 43,25

D(E,e2) = **√** (6 - 6,5)2 +(2 - 2,5)2  = **√ 0,5**

D(E,e3) = **√** (6 – 4)2 +(2 – (-2,5))2 = 1,5

Минимальным является расстояние d(E,e2)

Сравниваем расстояние от точки F(3; -2) до эталонных точек.

D(F,e1) = **√** (3– (-0,5))2 +(-2 - 1)2 = √ 21,25

D(F,e2) = **√** (3 - 6,5)2 +(-2 - 2,5)2  = √ 32,5

D(F,e3) = **√** (3– 4)2 +(-2 – (-2,5))2 = **√ 1,25**

Минимальным является расстояние d(F,e3)

Выписываем в таблице принадлежность объектов кластерам

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Объект | A | B | C | D | E | F |
| Номер кластера | 1 | 1 | 2 | 3 | 2 | 3 |

|  |  |  |
| --- | --- | --- |
| e1 | e2 | e3 |
| A,B | C,E | D,F |

Пересчитываем значения для эталонных точек.

e1 = (-0,5; 1)

e2 = (6,5; 2,5)

e3 = (4; -2,5)  
Границы кластеров не изменились процесс разбиения на кластеры закончен.