

Họ tên: Lê Trung Hiếu

①

Mã HV: 20007920

BÀI KIỂM TRA HỌC MÁY

Bước 1:

Bước 1: Tính khoảng cách các điểm theo city-block: $D(x_i, y_j) = \sum_{k=1}^p |x_{ik} - x_{jk}|$

Điểm	(1)	(2)	(3)	(4)	(5)
(1)	0	4	1	3	4
(2)		0	5	7	2
(3)			0	2	3
(4)				0	5
(5)					0

Kết nhập (2), (4) thành (6) theo tiêu chuẩn đường kính 2 tập

Bước 2:

Điểm	(1)	(3)	(5)	(6)
(1)	0	1	4	4
(3)		0	3	5
(5)			0	5
(6)				0

$$d(6, \{1\}) = \max \{d(\{2\}, \{1\}), d(\{4\}, \{1\})\} = \max \{4, 3\} = 4$$

$$d(6, \{3\}) = \max \{d(\{2\}, \{3\}), d(\{4\}, \{3\})\} = \max \{5, 2\} = 5$$

$$d(6, \{5\}) = \max \{d(\{2\}, \{5\}), d(\{4\}, \{5\})\} = \max \{2, 5\} = 5$$

Kết nhập (5), (6) thành (7)

Bước 3: $d(7, \{1\}) = \max \{4, 5, 4\} = 4$

$$d(7, \{3\}) = \max \{5, 2, 3\} = 5$$

(2)

Điểm	(1)	(3)	(7)
(1)	0	1	4
(3)		0	5
(7)			0

Kết nhập (3), (7) thành (8)

Bước 4: $d(8, \{1\}) = \max\{4, 3, 4, 1\} = 4$.

Điểm	(1)	(8)
(1)	0	4
(8)		0

Kết nhập (1), (8) thành (9). Kết thúc phần này.

Bài 3:

Bước 1: Tính độ hỗ trợ cho từng phân tử:

$$\text{sup}(\{\text{cola}\}) = 2.$$

$$\text{sup}(\{\text{bánh mì}\}) = 3$$

$$\text{sup}(\{\text{martini}\}) = 3$$

$$\text{sup}(\{\text{bơ}\}) = 3$$

$$\text{sup}(\{\text{snack}\}) = 3$$

$$\text{sup}(\{\text{sữa}\}) = 1$$

$$\text{sup}(\{\text{xúc xích}\}) = 3$$

$$\text{sup}(\{\text{mật}\}) = 1.$$

Ngưỡng $s = 2$, tập mục thường xuyên:

$\{\text{cola}, \text{martini}, \text{snack}, \text{xúc xích}, \text{bánh mì}, \text{bơ}\}$

Bước 2:

$$\text{sup}(\{\text{cola}, \text{martini}\}) = 1$$

$$\text{sup}(\{\text{cola}, \text{snack}\}) = 1$$

...

$$\text{sup}(\{\text{martini}, \text{snack}\}) = 2.$$

$$\text{sup}(\{\text{martini}, \text{xúc xích}\}) = 2.$$

$$\text{sup}(\{\text{snack}, \text{xúc xích}\}) = 2.$$

Ngưỡng $s = 2$, chọn các tập mục: $\{\text{martini}, \text{snack}\}, \{\text{martini}, \text{xúc xích}\}, \{\text{snack}, \text{xúc xích}\}$

Bài 2:

	x_1	x_2	x_3	x_4
w1	1	0	1	0
w2	2	0	1	0
w3	1	1	1	0
w4	1	0	0	0
w5	4	1	2	0

Bước 1: Chọn ngẫu nhiên w1, w3 là tâm 2 cụm C1, C2.

Tính khoảng cách từ các điểm còn lại đến w1, w3

Điểm	w1	w3
w2	1	2
w4	1	2
w5	5	4

C1	w1
C2	w3

Phân w2, w4 vào cụm C1, w5 vào cụm C2

C1	w1, w2, w4
C2	w3, w5

Bước 2: Tính lại tâm cụm: gọi O1, O2 là tâm cụm 1 và 2.

	x_1	x_2	x_3	x_4
O1	1,33	—	0,67	—
O2	2,5	1	1,5	—

Tính lại khoảng cách đến tâm cụm:

	O1	O2
w1	0,67	3
w2	1	2
w3	1,67	2
w4	1	4
w5	5	2

C1	w1, w3, w2, w4
C2	w5

Tính lại tâm cụm sau bước 2.

	x_1	x_2	x_3	x_4
O1	1,25	0,25	0,75	—
O2	4	1	2	—

4

Tính lại khoảng cách từ các điểm đến tâm cụm:

	01	02
w1	0,75	5
w2	1,25	4
w3	1,25	4
w4	1,25	6
w5	4,75	-

c1	w1, w2, w3, w4
c2	w5

Đừng thuật toán vì tâm cụm không đổi.