**Принцип “дальнего соседа”**.

Исходные данные.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № п/п | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| x1 | 4 | 8 | 15 | 9 | 2 | 15 | 2 | 12 |
| x2 | 3 | 6 | 9 | 8 | 4 | 10 | 3 | 8 |

1. Воспользуемся агломеративным иерархическим алгоритмом классификации. В качестве расстояния между объектами примем обычное евклидовое расстояние. Тогда согласно формуле:

$$p(x\_{ij})$$

 = $$\sqrt{\sum\_{}^{}(x\_{il}-x\_{ji})^{2}}$$

где l - признаки; k - количество признаков

$$p(x\_{1,2})$$

$$=$$

$$\sqrt{(4-8)^{2} + (3-6)^{2}}$$

$$=$$

$$5$$

$$p(x\_{1,3})$$

$$=$$

$$\sqrt{(4-15)^{2} + (3-9)^{2}}$$

$$=$$

$$12.53$$

$$p(x\_{1,4})$$

$$=$$

$$\sqrt{(4-9)^{2} + (3-8)^{2}}$$

$$=$$

$$7.07$$

2. Полученные данные помещаем в таблицу (матрицу расстояний).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № п/п | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | 0 | 5 | 12.53 | 7.071 | 2.236 | 13.038 | 2 | 9.434 |
| 2 | 5 | 0 | 7.616 | 2.236 | 6.325 | 8.062 | 6.708 | 4.472 |
| 3 | 12.53 | 7.616 | 0 | 6.083 | 13.928 | 1 | 14.318 | 3.162 |
| 4 | 7.071 | 2.236 | 6.083 | 0 | 8.062 | 6.325 | 8.602 | 3 |
| 5 | 2.236 | 6.325 | 13.928 | 8.062 | 0 | 14.318 | 1 | 10.77 |
| 6 | 13.038 | 8.062 | 1 | 6.325 | 14.318 | 0 | 14.765 | 3.606 |
| 7 | 2 | 6.708 | 14.318 | 8.602 | 1 | 14.765 | 0 | 11.18 |
| 8 | 9.434 | 4.472 | 3.162 | 3 | 10.77 | 3.606 | 11.18 | 0 |

3. Поиск наименьшего расстояния.

Из матрицы расстояний следует, что объекты 3 и 6 наиболее близки P3;6 = 1 и поэтому объединяются в один кластер.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № п/п | 1 | 2 | **[3]** | 4 | 5 | **[6]** | 7 | 8 |
| 1 | 0 | 5 | 12.53 | 7.071 | 2.236 | 13.038 | 2 | 9.434 |
| 2 | 5 | 0 | 7.616 | 2.236 | 6.325 | 8.062 | 6.708 | 4.472 |
| **[3]** | 12.53 | 7.616 | 0 | 6.083 | 13.928 | 1 | 14.318 | 3.162 |
| 4 | 7.071 | 2.236 | 6.083 | 0 | 8.062 | 6.325 | 8.602 | 3 |
| 5 | 2.236 | 6.325 | 13.928 | 8.062 | 0 | 14.318 | 1 | 10.77 |
| **[6]** | 13.038 | 8.062 | 1 | 6.325 | 14.318 | 0 | 14.765 | 3.606 |
| 7 | 2 | 6.708 | 14.318 | 8.602 | 1 | 14.765 | 0 | 11.18 |
| 8 | 9.434 | 4.472 | 3.162 | 3 | 10.77 | 3.606 | 11.18 | 0 |

При формировании новой матрицы расстояний, выбираем наибольшее значение из значений объектов №3 и №6.

В результате имеем 7 кластера: S(1), S(2), S(3,6), S(4), S(5), S(7), S(8)

Из матрицы расстояний следует, что объекты 5 и 7 наиболее близки P5;7 = 1 и поэтому объединяются в один кластер.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № п/п | 1 | 2 | 3,6 | 4 | **[5]** | **[7]** | 8 |
| 1 | 0 | 5 | 13.038 | 7.071 | 2.236 | 2 | 9.434 |
| 2 | 5 | 0 | 8.062 | 2.236 | 6.325 | 6.708 | 4.472 |
| 3,6 | 13.038 | 8.062 | 0 | 6.325 | 14.318 | 14.765 | 3.606 |
| 4 | 7.071 | 2.236 | 6.325 | 0 | 8.062 | 8.602 | 3 |
| **[5]** | 2.236 | 6.325 | 14.318 | 8.062 | 0 | 1 | 10.77 |
| **[7]** | 2 | 6.708 | 14.765 | 8.602 | 1 | 0 | 11.18 |
| 8 | 9.434 | 4.472 | 3.606 | 3 | 10.77 | 11.18 | 0 |

При формировании новой матрицы расстояний, выбираем наибольшее значение из значений объектов №5 и №7.

В результате имеем 6 кластера: S(1), S(2), S(3,6), S(4), S(5,7), S(8)

Из матрицы расстояний следует, что объекты 1 и 5,7 наиболее близки P1;5,7 = 2.24 и поэтому объединяются в один кластер.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № п/п | **[1]** | 2 | 3,6 | 4 | **[5,7]** | 8 |
| **[1]** | 0 | 5 | 13.038 | 7.071 | 2.236 | 9.434 |
| 2 | 5 | 0 | 8.062 | 2.236 | 6.708 | 4.472 |
| 3,6 | 13.038 | 8.062 | 0 | 6.325 | 14.765 | 3.606 |
| 4 | 7.071 | 2.236 | 6.325 | 0 | 8.602 | 3 |
| **[5,7]** | 2.236 | 6.708 | 14.765 | 8.602 | 0 | 11.18 |
| 8 | 9.434 | 4.472 | 3.606 | 3 | 11.18 | 0 |

При формировании новой матрицы расстояний, выбираем наибольшее значение из значений объектов №1 и №5,7.

В результате имеем 5 кластера: S(1,5,7), S(2), S(3,6), S(4), S(8)

Из матрицы расстояний следует, что объекты 2 и 4 наиболее близки P2;4 = 2.24 и поэтому объединяются в один кластер.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № п/п | 1,5,7 | **[2]** | 3,6 | **[4]** | 8 |
| 1,5,7 | 0 | 6.708 | 14.765 | 8.602 | 11.18 |
| **[2]** | 6.708 | 0 | 8.062 | 2.236 | 4.472 |
| 3,6 | 14.765 | 8.062 | 0 | 6.325 | 3.606 |
| **[4]** | 8.602 | 2.236 | 6.325 | 0 | 3 |
| 8 | 11.18 | 4.472 | 3.606 | 3 | 0 |

При формировании новой матрицы расстояний, выбираем наибольшее значение из значений объектов №2 и №4.

В результате имеем 4 кластера: S(1,5,7), S(2,4), S(3,6), S(8)

Из матрицы расстояний следует, что объекты 3,6 и 8 наиболее близки P3,6;8 = 3.61 и поэтому объединяются в один кластер.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № п/п | 1,5,7 | 2,4 | **[3,6]** | **[8]** |
| 1,5,7 | 0 | 8.602 | 14.765 | 11.18 |
| 2,4 | 8.602 | 0 | 8.062 | 4.472 |
| **[3,6]** | 14.765 | 8.062 | 0 | 3.606 |
| **[8]** | 11.18 | 4.472 | 3.606 | 0 |

При формировании новой матрицы расстояний, выбираем наибольшее значение из значений объектов №3,6 и №8.

В результате имеем 3 кластера: S(1,5,7), S(2,4), S(3,6,8)

|  |  |  |  |
| --- | --- | --- | --- |
| № п/п | 1,5,7 | 2,4 | 3,6,8 |
| 1,5,7 | 0 | 8.602 | 14.765 |
| 2,4 | 8.602 | 0 | 8.062 |
| 3,6,8 | 14.765 | 8.062 | 0 |

Результаты иерархической классификации объектов представлены на рис. в виде дендрограммы.

Решение было получено и оформлено с помощью сервиса:

[Кластерный анализ](https://axd.semestr.ru/upr/cluster.php)

С этой задачей также решают:

[Метод К-средних](https://axd.semestr.ru/upr/average.php)

[Математические методы в психологии](https://math.semestr.ru/group/mathematical-psychology.php)