

Баллодожималка. Вебинар 13.06.2026

Ссылка на архив с файлами к задачам ✨

Кластеризация (№27)

Задание 1

Фрагмент звёздного неба спроецирован на плоскость с декартовой системой координат. Учёный решил провести кластеризацию полученных точек, являющихся изображениями звёзд, то есть разбить их множество на N непересекающихся непустых подмножеств (кластеров), таких что точки каждого подмножества лежат внутри прямоугольника со сторонами длиной H и W , причём эти прямоугольники между собой не пересекаются. Стороны прямоугольников не обязательно параллельны координатным осям. Гарантируется, что такое разбиение существует и единственно для заданных размеров прямоугольников.

Будем называть центром кластера точку этого кластера, сумма расстояний от которой до всех остальных точек кластера минимальна. Для каждого кластера гарантируется единственность его центра. Расстояние между двумя точками на плоскости $A(x_1, y_1)$ и $B(x_2, y_2)$ вычисляется по формуле:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

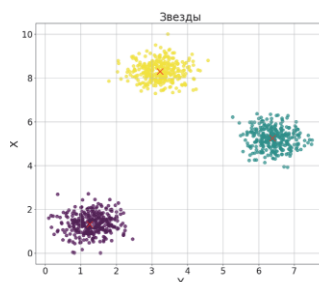
В файле A хранятся данные о звёздах двух кластеров, где $H = 6$, $W = 4,5$ для каждого кластера. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y . Значения даны в условных единицах. Известно, что количество звёзд не превышает 1000.

В файле B хранятся данные о звёздах трёх кластеров, где $H = 5$, $W = 6$ для каждого кластера. Известно, что количество звёзд не превышает 1000. Известно, что в файле B имеются координаты ровно трёх «лишних» точек, являющихся аномалиями, возникшими в результате помех при передаче данных. Эти три точки не относятся ни к одному из кластеров, их учитывать не нужно. Для файла A определите координаты центра каждого кластера, затем найдите два числа: P_x — максимальную из абсцисс центров кластеров, и P_y — максимальную из ординат центров кластеров. Для файла B определите координаты центра каждого кластера, затем найдите два числа: Q_x — разность абсцисс центров кластеров с минимальным и максимальным количеством точек, и Q_y — разность ординат центров кластеров с минимальным и максимальным количеством точек. Гарантируется, что во всех кластерах количество точек различно.

В ответе запишите четыре числа: в первой строке — сначала целую часть абсолютного значения произведения $P_x \times 10000$, затем целую часть абсолютного значения произведения $P_y \times 10000$; во второй строке — сначала целую часть абсолютного значения произведения $Q_x \times 10000$, затем целую часть абсолютного значения произведения $Q_y \times 10000$.

Возможные данные одного из файлов иллюстрированы графиком.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию. Для выполнения задания используйте данные из прилагаемого файла.



Задание 2

Фрагмент звёздного неба спроецирован на плоскость с декартовой системой координат. Учёный решил провести кластеризацию полученных точек, являющихся изображениями звёзд, то есть разбить их множество на N непересекающихся непустых подмножеств (кластеров), таких что точки каждого подмножества лежат внутри прямоугольника со сторонами длиной H и W , причём эти прямоугольники между собой не пересекаются. Стороны прямоугольников не обязательно параллельны координатным осям. Гарантируется, что такое разбиение существует и единственно для заданных размеров прямоугольников.

Будем называть центром кластера точку этого кластера, сумма расстояний от которой до всех остальных точек кластера минимальна. Для каждого кластера гарантируется единственность его центра. Расстояние между двумя точками на плоскости $A(x_1, y_1)$ и $B(x_2, y_2)$ вычисляется по формуле:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

В файле A хранятся данные о звёздах двух кластеров, где $H = 6$, $W = 5$ для каждого кластера. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y . Значения даны в условных единицах. Известно, что количество звёзд не превышает 1000.

В файле B хранятся данные о звёздах трёх кластеров, где $H = 6$, $W = 5$ для каждого кластера. Известно, что количество звёзд не превышает 1000. Структура хранения информации о звёздах в файле B аналогична файлу A .

Известно, что в файле B имеются координаты ровно трёх «лишних» точек, являющихся аномалиями, возникшими в результате помех при передаче данных. Эти три точки не относятся ни к одному из кластеров, их учитывать не нужно.

Для файла A определите координаты центра каждого кластера, затем найдите два числа: P_x — сумму абсцисс центров кластеров, и P_y — сумму ординат центров кластеров.

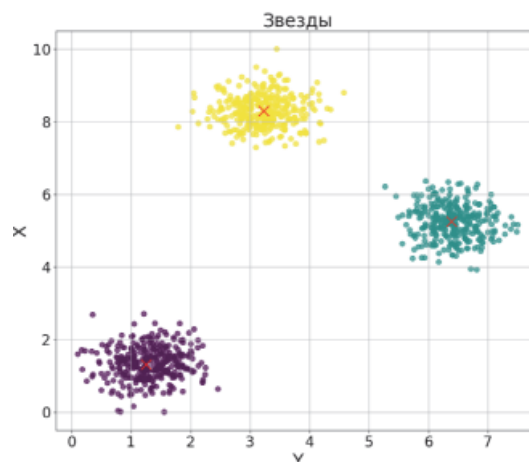
Для файла B определите координаты центра каждого кластера, затем найдите два числа: Q_1 — минимальное расстояние между центрами различных кластеров, и Q_2 — максимальное расстояние между центрами кластеров.

В ответе запишите четыре числа: в первой строке — сначала абсолютную величину целой части произведения $P_x \times 10000$, затем абсолютную величину целой части произведения $P_y \times 10000$; во второй строке — сначала абсолютную величину целой части произведения $Q_1 \times 10000$, затем абсолютную величину целой части произведения $Q_2 \times 10000$.

Возможные данные одного из файлов иллюстрированы графиком.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию.

Для выполнения задания используйте данные из прилагаемого файла.



Задание 3

Фрагмент звёздного неба спроецирован на плоскость с декартовой системой координат. Учёный решил провести кластеризацию полученных точек, являющихся изображениями звёзд, то есть разбить их множество на N непересекающихся непустых подмножеств (кластеров), таких что точки каждого подмножества лежат внутри прямоугольника со сторонами длиной H и W , причём эти прямоугольники между собой не пересекаются. Стороны прямоугольников не обязательно параллельны координатным осям. Гарантируется, что такое разбиение существует и единственно для заданных размеров прямоугольников.

Будем называть центром кластера точку этого кластера, сумма расстояний от которой до всех остальных точек кластера минимальна. Для каждого кластера гарантируется единственность его центра. Расстояние между двумя точками на плоскости $A(x_1, y_1)$ и $B(x_2, y_2)$ вычисляется по формуле:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

В файле A хранятся данные о звёздах двух кластеров, где $H = 6,5$, $W = 4,5$ для каждого кластера. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y . Значения даны в условных единицах. Известно, что количество звёзд не превышает 1000.

В файле B хранятся данные о звёздах трёх кластеров, где $H = 6,5$, $W = 5$ для каждого кластера. Известно, что количество звёзд не превышает 1000. Структура хранения информации о звёздах в файле B аналогична файлу A .

Для файла A определите координаты центра каждого кластера, затем найдите два числа: P_1 — минимальное расстояние от точки с координатами $(1, 0; 1, 0)$ до центра кластера, и P_2 — максимальное расстояние от этой же точки до центра кластера.

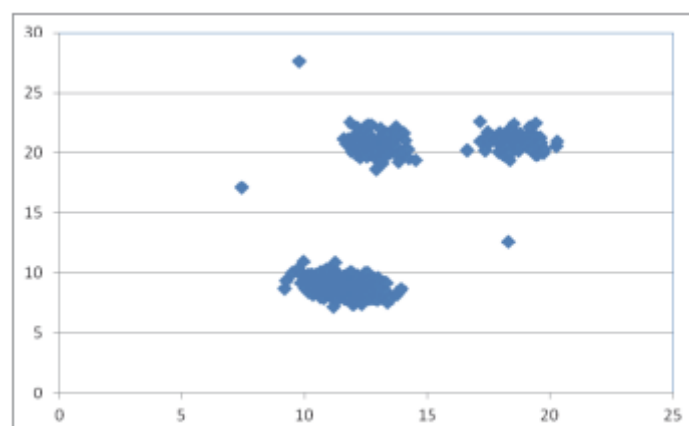
Для файла B определите координаты центра каждого кластера, затем найдите два числа: Q_1 — в кластере с наибольшим количеством точек число таких точек, которые находятся на расстоянии не более 1,2 от центра кластера, и Q_2 — в кластере с наибольшим количеством точек число таких точек, которые находятся на расстоянии не более 0,75 от центра кластера. Гарантируется, что во всех кластерах количество точек различно.

В ответе запишите четыре числа: в первой строке — сначала целую часть произведения $P_1 \times 10000$, затем целую часть произведения $P_2 \times 10000$; во второй строке — сначала Q_1 , затем Q_2 .

Возможные данные одного из файлов иллюстрированы графиком.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию.

Для выполнения задания используйте данные из прилагаемого файла.



Задание 4

Фрагмент звёздного неба спроецирован на плоскость с декартовой системой координат. Учёный решил провести кластеризацию полученных точек, являющихся изображениями звёзд, то есть разбить их множество на N непересекающихся непустых подмножеств (кластеров), таких что точки каждого подмножества лежат внутри прямоугольника со сторонами длиной H и W , причём эти прямоугольники между собой не пересекаются. Стороны прямоугольников не обязательно параллельны координатным осям. Гарантируется, что такое разбиение существует и единственно для заданных размеров прямоугольников.

Для каждой звезды дана характеристика: тип цвета, тип светимости и её размер в соответствии с искомой таблицей.

Обозначение	Цвет	Обозначение	Размер
G	белый	I	сверхгигант
J	зелёный	II	яркий гигант
L	синий	III	гигант
N	оранжевый	IV	субгигант
Y	красный	V	карлик
S	голубой	VI	субкарлик
Z	жёлтый	VII	белый карлик

Полученные значения записаны в характеристике слитно: обозначение цвета, светимость (обозначается арабской цифрой) и обозначение размера планеты.

Будем называть центром кластера точку этого кластера, сумма расстояний от которой до всех остальных точек кластера минимальна. Для каждого кластера гарантируется единственность его центра. Расстояние между двумя точками на плоскости $A(x_1, y_1)$ и $B(x_2, y_2)$ вычисляется по формуле:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

В файле A хранятся данные о звёздах двух кластеров, где $H = 6,5$, $W = 4,5$ для каждого кластера. В каждой строке записана информация о расположении на карте одной звезды: сначала координата x , затем координата y , а затем характеристика звезды. Значения даны в условных единицах. Известно, что количество звёзд не превышает 1000.

В файле B хранятся данные о звёздах трёх кластеров, где $H = 6,5$, $W = 5$ для каждого кластера. Известно, что количество звёзд не превышает 1000. Структура хранения информации о звёздах в файле B аналогична файлу A .

Для файла A определите координаты центра каждого кластера, затем найдите два числа: A_1 — минимальное расстояние от центра кластера с наименьшим количеством точек до красного гиганта, и A_2 — максимальное расстояние от центра кластера с наименьшим количеством точек до красного гиганта.

Для файла B определите координаты центра каждого кластера, затем найдите два числа: B_1 — минимальное расстояние между двумя различными жёлтыми сверхгигантами, расположенными в одном и том же кластере, и B_2 — расстояние между центрами кластеров с минимальным и максимальным количеством жёлтых сверхгигантов.

В ответе запишите четыре числа: в первой строке — сначала целую часть произведения $A_1 \times 10000$, затем целую часть произведения $A_2 \times 10000$; во второй строке — сначала целую часть абсолютного значения произведения $B_1 \times 10000$, затем целую часть абсолютного значения произведения $B_2 \times 10000$.

Внимание! График приведён в иллюстративных целях для произвольных значений, не имеющих отношения к заданию. Для выполнения задания используйте данные из прилагаемого файла.

Задание 5

Фрагмент звёздного неба спроецирован на плоскость с декартовой системой координат. Учёный решил провести кластеризацию полученных точек, являющихся изображениями звёзд, то есть разбить их множество на N непересекающихся непустых подмножеств (кластеров), таких что точки каждого подмножества лежат внутри прямоугольника со сторонами длиной H и W , причём эти прямоугольники между собой не пересекаются. Стороны прямоугольников не обязательно параллельны координатным осям.

Гарантируется, что такое разбиение существует и единственно для заданных размеров прямоугольников.

Гарантируется, что во всех кластерах количество точек различно.

Для каждой звезды задана характеристика – набор символов, включающий спектральный класс (обозначается латинской буквой), подкласс светимости (обозначается арабской цифрой) и класс светимости (обозначается римской цифрой).

Спектральный класс звезды определяется в соответствии с таблицей:

Спектральный класс звезды						
О	В	А	Ф	Г	К	М
Голубая	Бело-голубая	Белая	Жёлто-белая	Жёлтая	Оранжевая	Красная

Класс светимости звезды определяется в соответствии с таблицей:

Класс светимости звезды						
I	II	III	IV	V	VI	VII
Сверхгиганты	Яркие гиганты	Гиганты	Субгиганты	Карлики	Субкарлики	Белые карлики

Будем называть центром кластера точку этого кластера, сумма расстояний от которой до всех остальных точек кластера минимальна. Для каждого кластера гарантируется единственность его центра. Расстояние между двумя точками на плоскости $A(x_1, y_1)$ и $B(x_2, y_2)$ вычисляется по формуле:

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

В файле А хранятся данные о звёздах двух кластеров, где $H = 5$ и $W = 4$ для каждого кластера. В каждой строке записана информация одной звезды: координата x , координата y и её характеристика. Значения даны в условных единицах. Известно, что количество точек не превышает 1000.

В файле Б хранятся данные о звёздах трёх кластеров, где $H = 3$ и $W = 5$ для каждого кластера. Известно, что количество точек не превышает 10 000. Структура хранения информации о звёздах в файле Б аналогична структуре в файле А.

Для файла А определите координаты центра каждого кластера, затем найдите два числа: A_1 – абсциссу ближайшего белого гиганта к центру кластера с наименьшим количеством точек, и A_2 – ординату ближайшего белого гиганта к центру кластера с наименьшим количеством точек.

Для файла Б определите координаты центра каждого кластера, затем найдите два числа: B_1 – расстояние между центрами кластеров с наибольшим и наименьшим количеством красных сверхгигантов, и B_2 – максимальное расстояние между двумя жёлтыми карликами, находящимися в одном кластере.

В ответе запишите четыре числа: в первой строке – сначала целую часть произведения $A_1 \times 10\,000$, затем целую часть произведения $A_2 \times 10\,000$; во второй строке – сначала целую часть произведения $B_1 \times 10\,000$, затем целую часть произведения $B_2 \times 10\,000$.

Ответы

1. 69663 192156
867 161306;
2. 107002 323741
58778 151839;
3. 58605 128643
358 203;
4. 4940 74302
1035 125591.
5. 390205 182699
60306 16460.