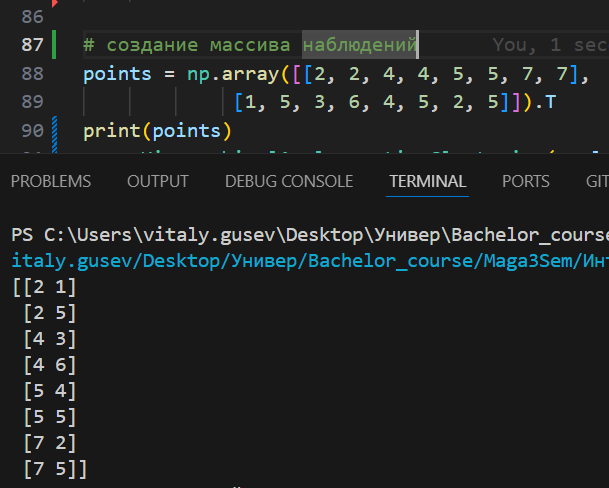
Гусев Виталий Евгеньевич

09-335

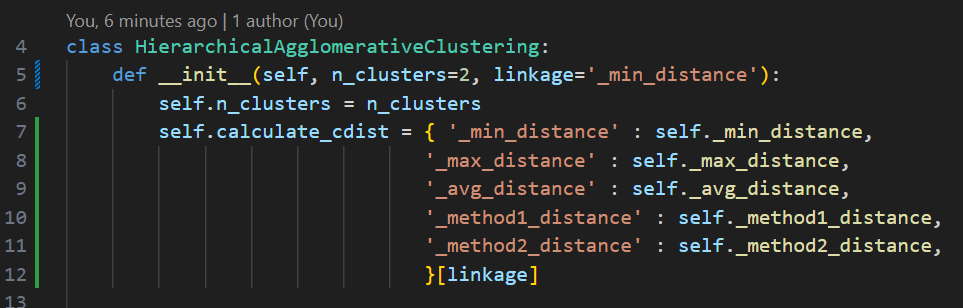
Интеллектуальный анализ данных

Работа 2. Иерархическая кластеризация

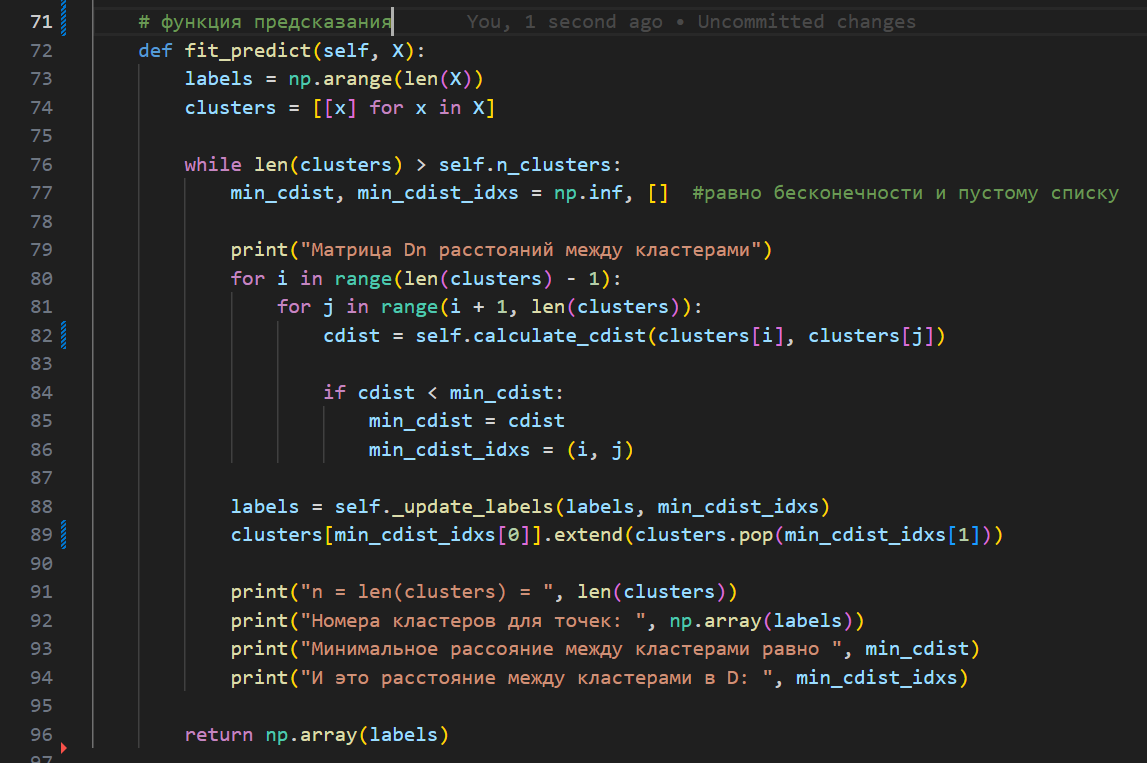
1. Был создан массив наблюдений и выведен в консоли.



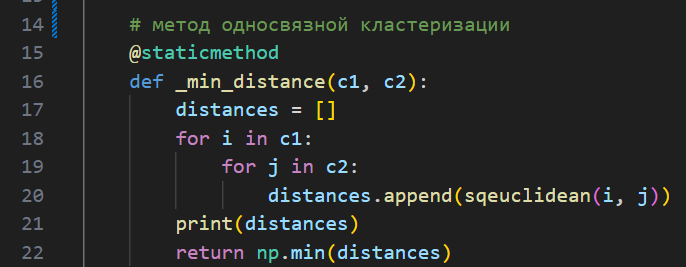
2. Был создан класс иерархической кластеризации, которая на вход получает количество кластеров для вычисления и метод вычисления расстояния между кластерами.



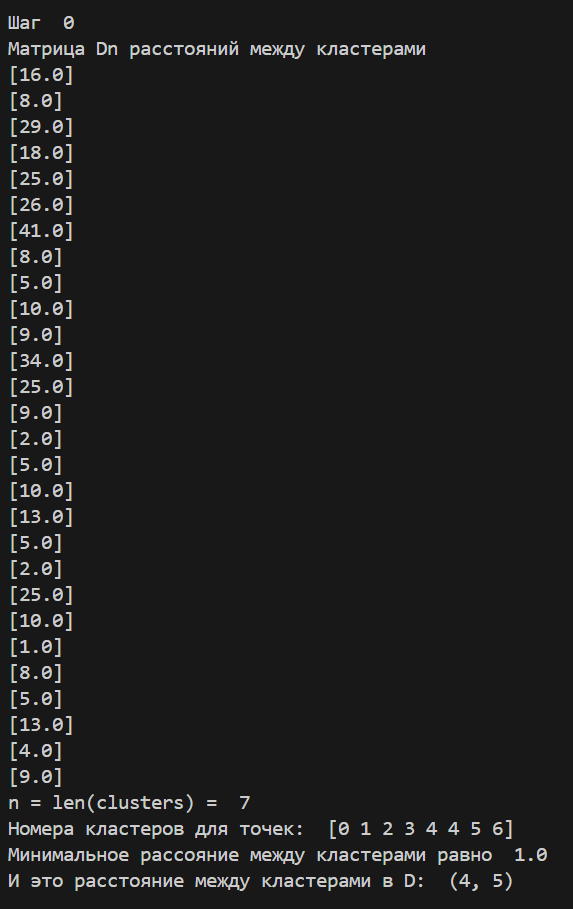
3. В данном классе была реализована функция вычисления кластеров, которая выводит матрицу Dn расстояний между кластерами, количество кластеров, номера кластеров для точек, равенство минимального расстояния между кластерами в D.

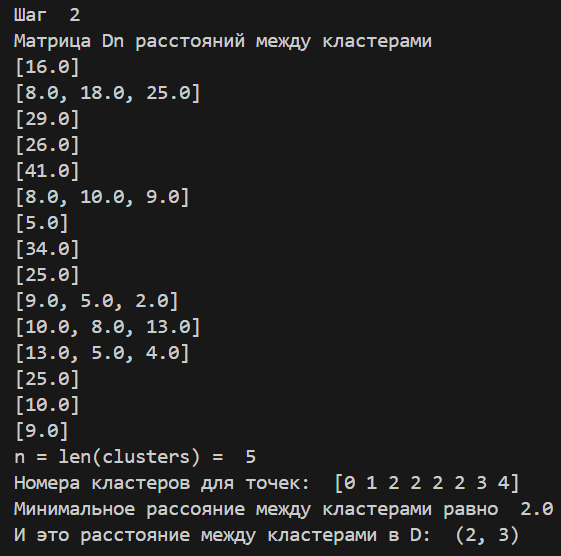


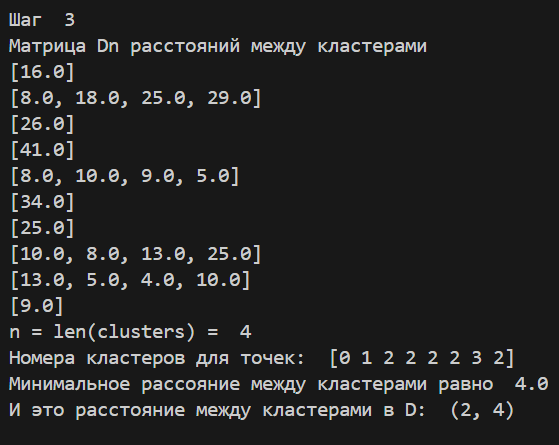
4. Был реализован метод односвязной кластеризации.

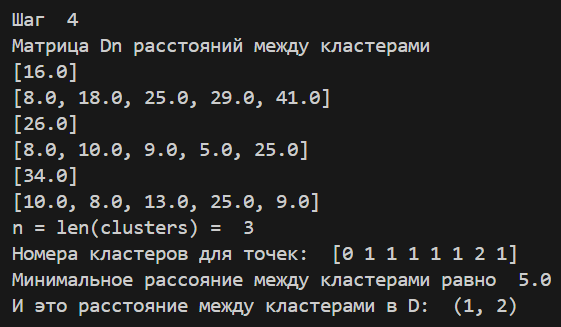


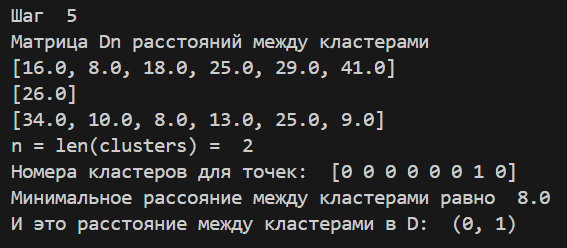
В результате вычисления 2-ух кластеров данным методом были получены следующие результаты:







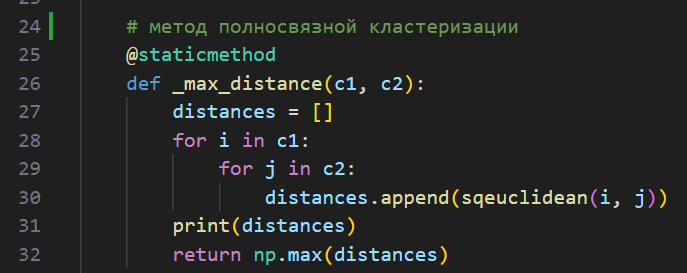




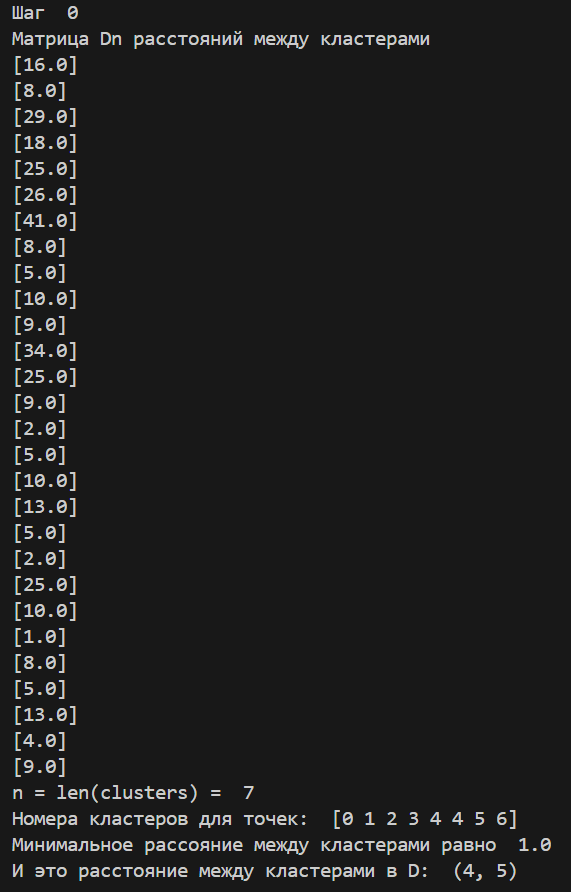
Результат предсказания методом односвязной кластеризации:

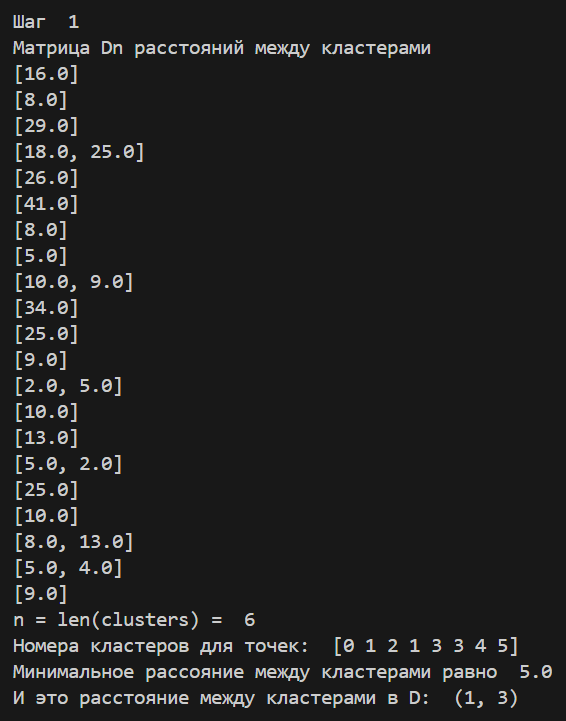


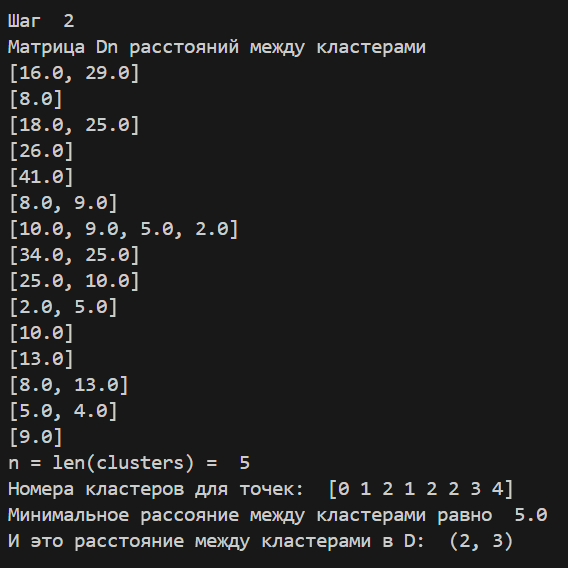
5. Был реализован метод полносвязной кластеризации.

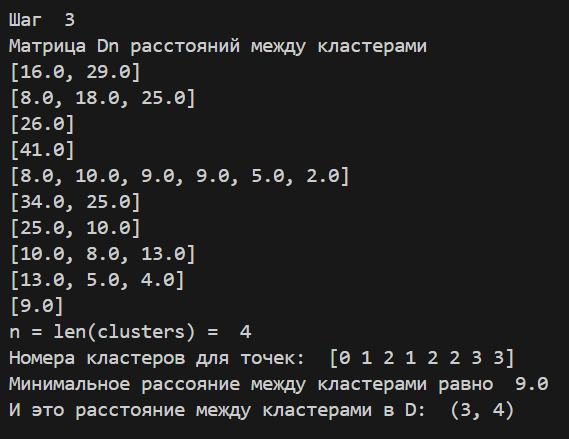


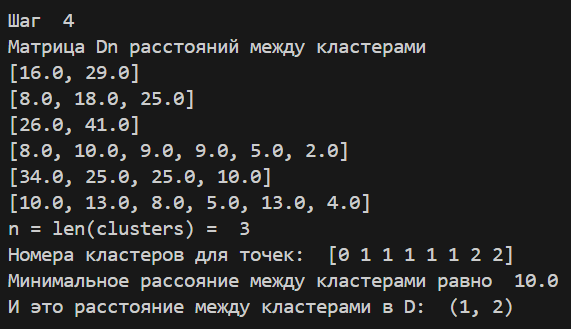
В результате вычисления 2-ух кластеров данным методом были получены следующие результаты:

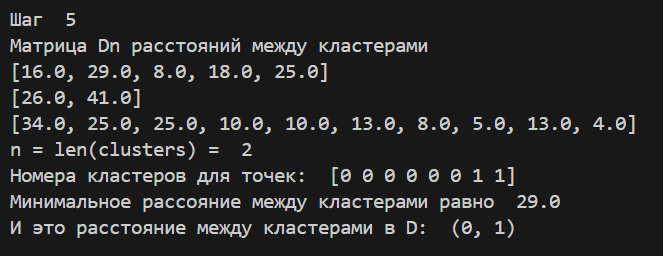








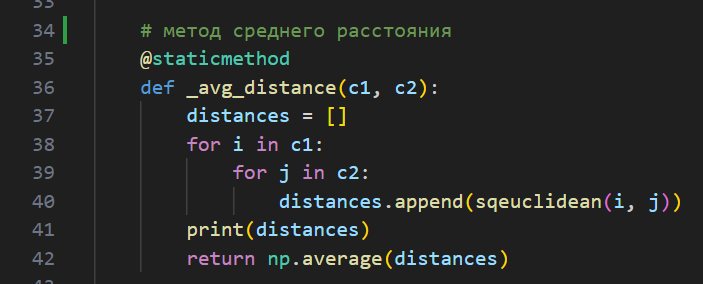




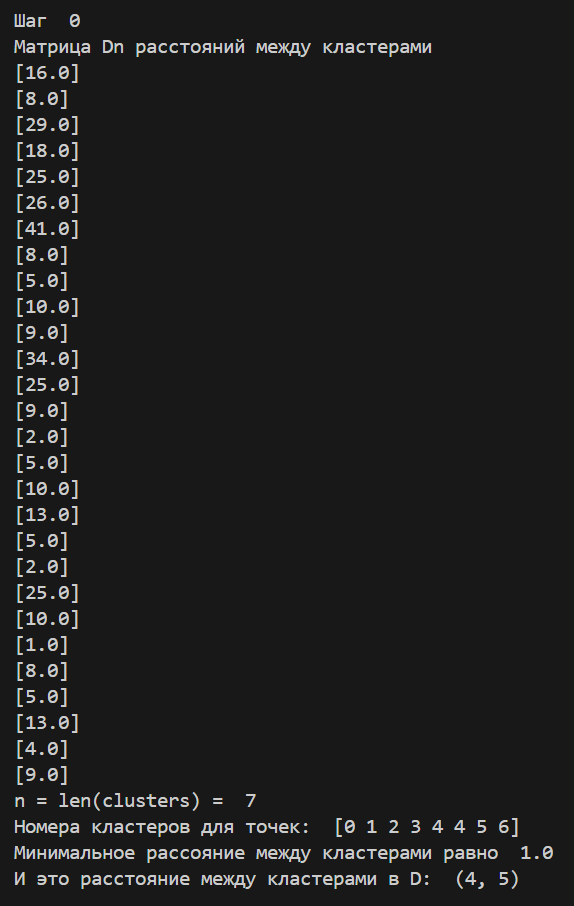
Результат предсказания методом односвязной кластеризации:

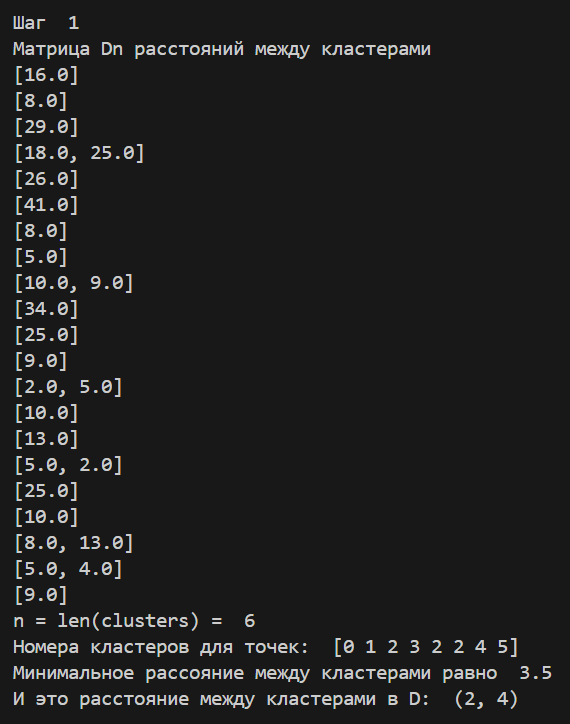


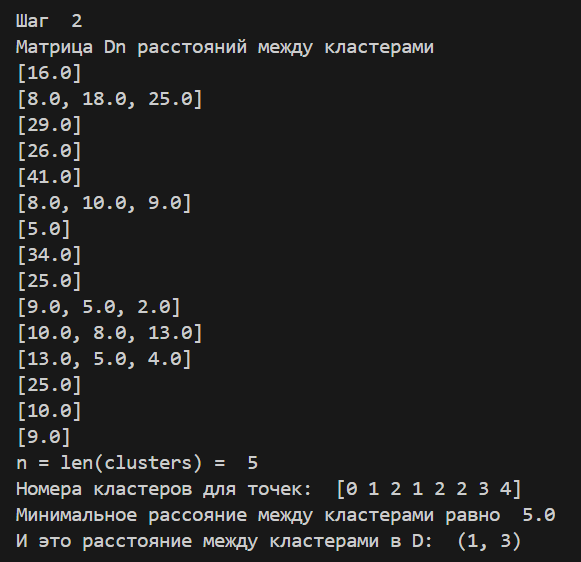
6. Был реализован метод среднего расстояния.

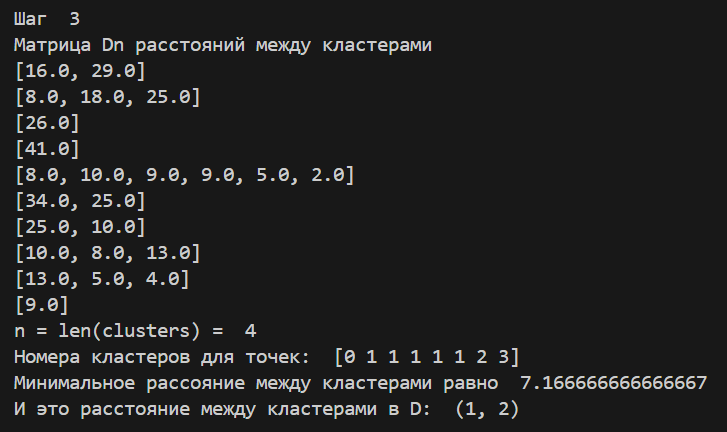


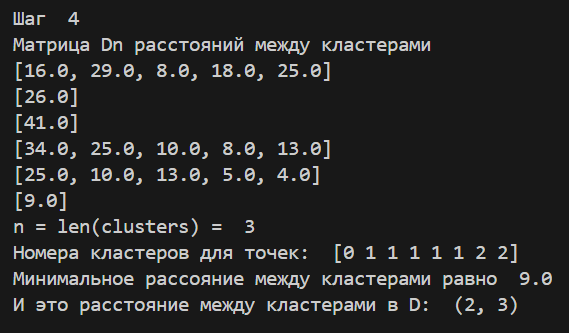
В результате вычисления 2-ух кластеров данным методом были получены следующие результаты:

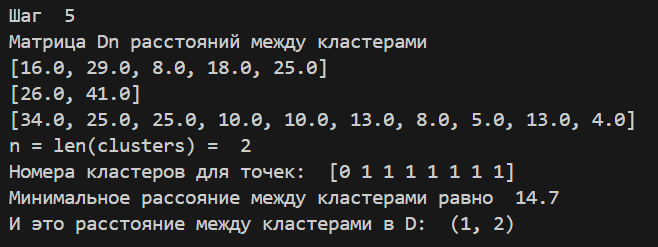








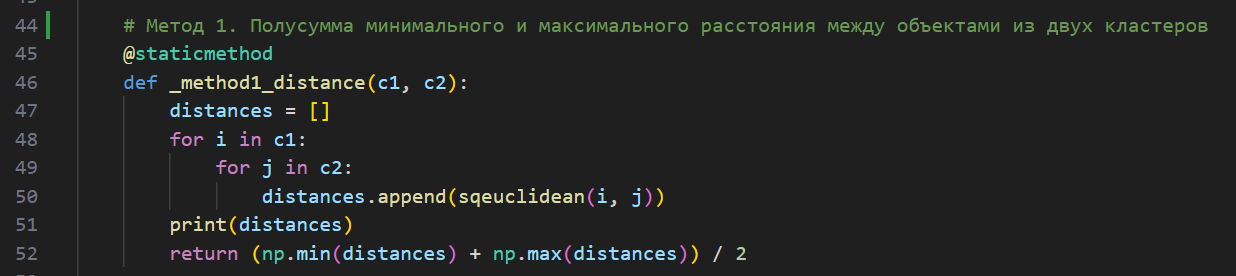




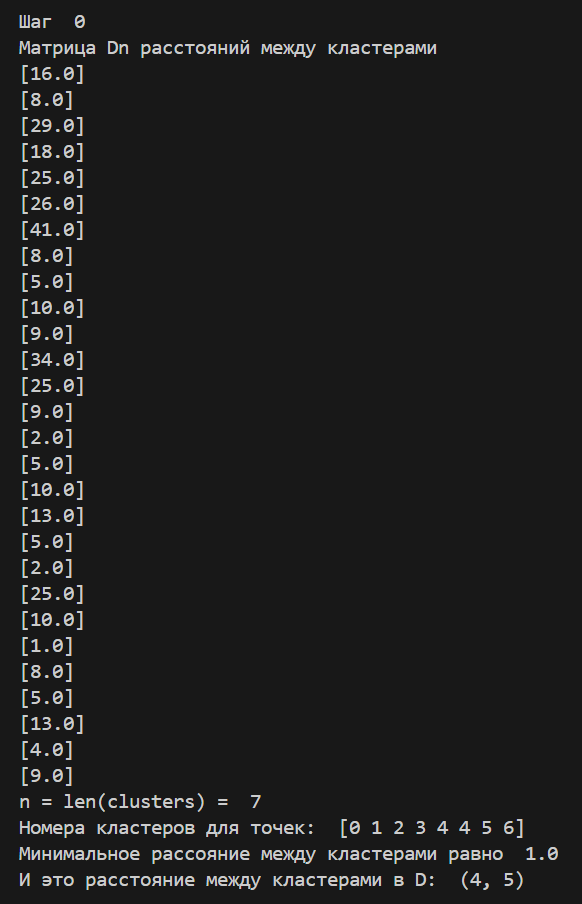
Результат предсказания методом односвязной кластеризации:

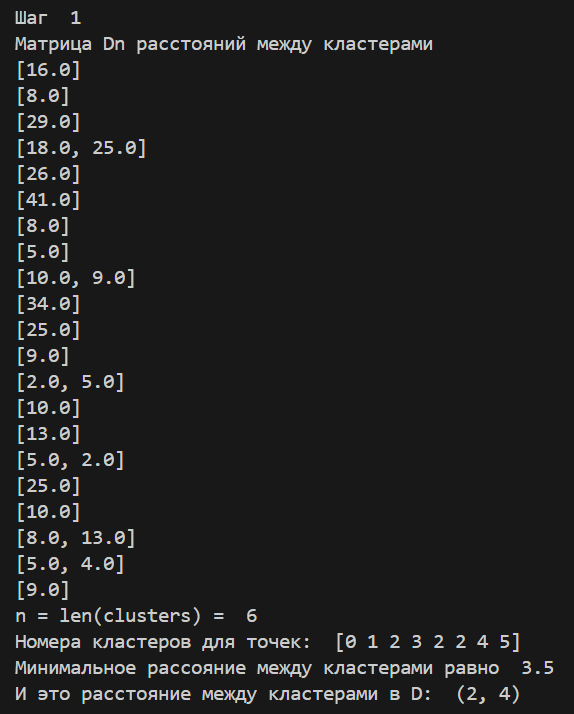


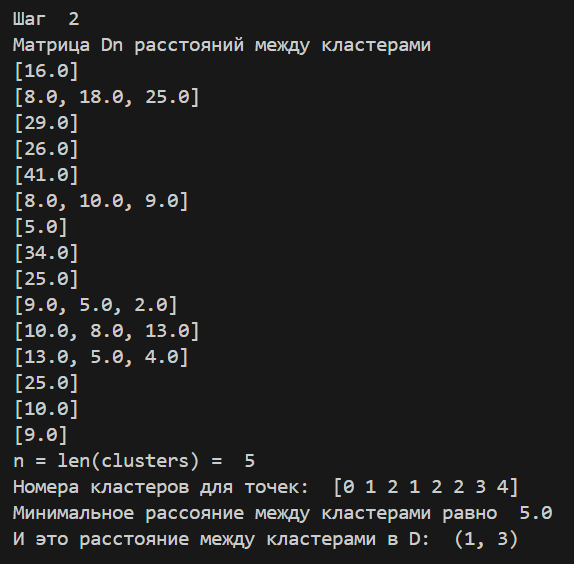
7. Был реализован «Метод 1. Полусумма минимального и максимального расстояния между объектами из двух кластеров».

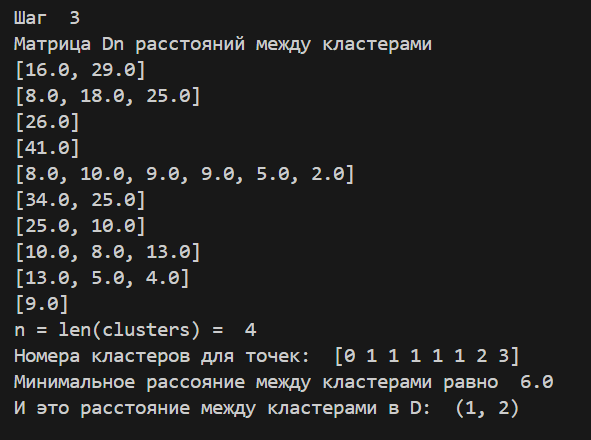


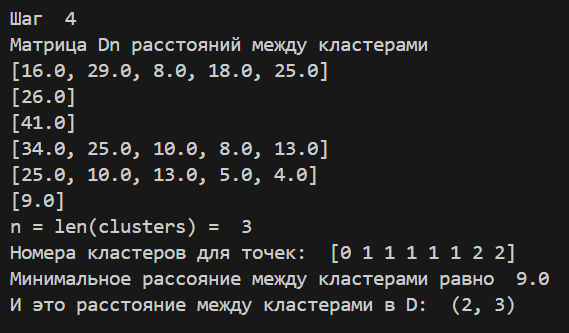
В результате вычисления 2-ух кластеров данным методом были получены следующие результаты:

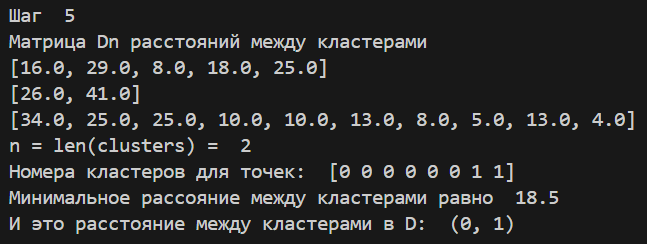








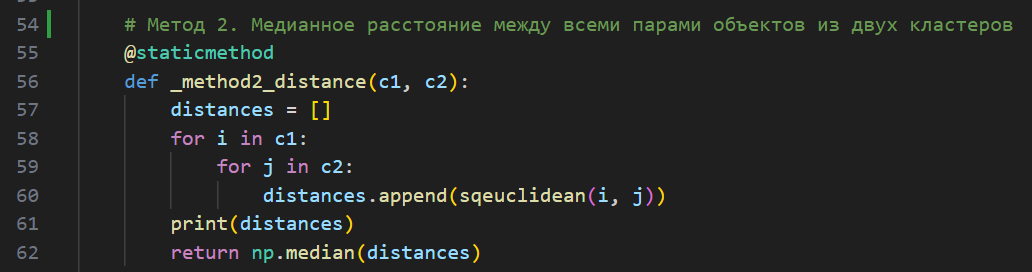




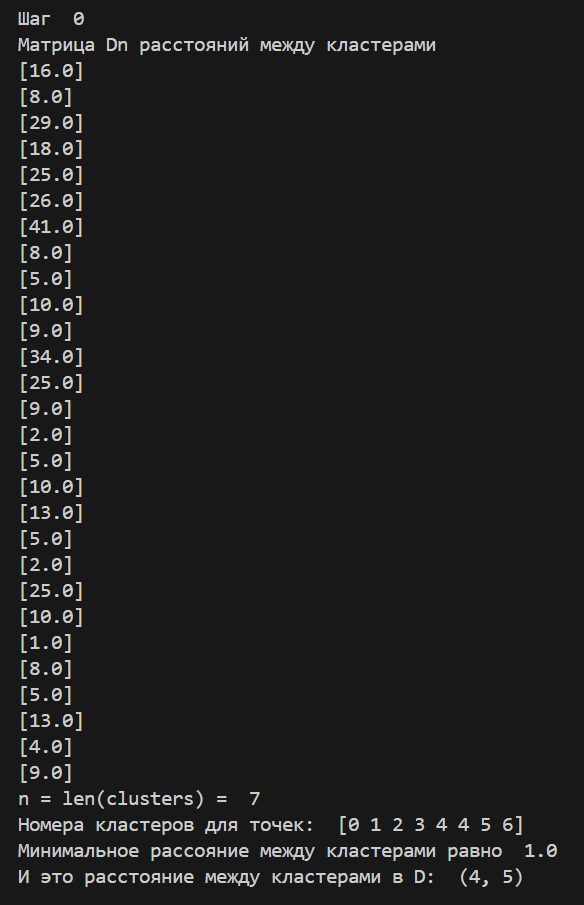
Результат предсказания методом односвязной кластеризации:

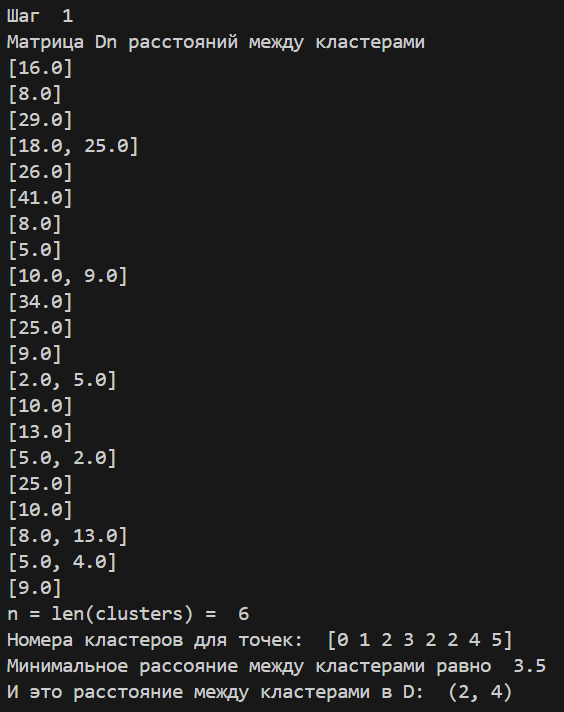


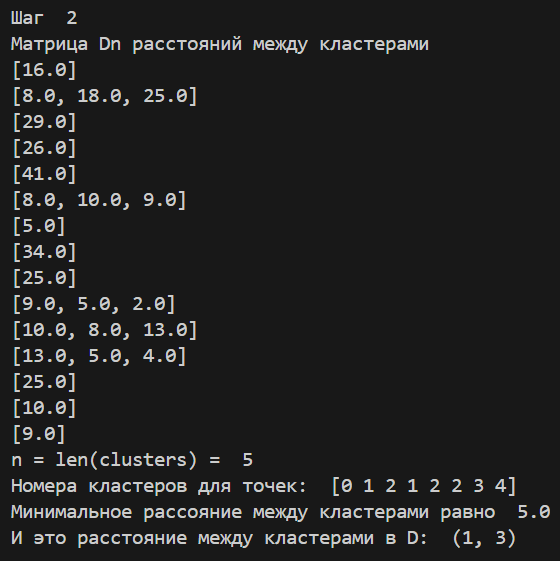
8. Был реализован «Метод 2. Медианное расстояние между всеми парами объектов из двух кластеров».

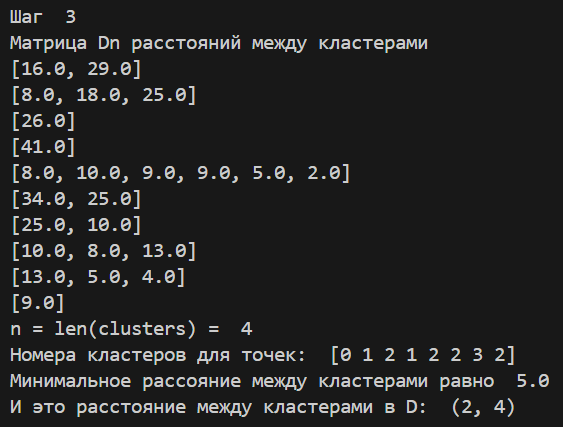


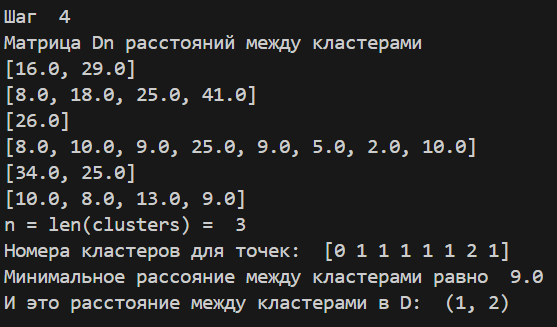
В результате вычисления 2-ух кластеров данным методом были получены следующие результаты:

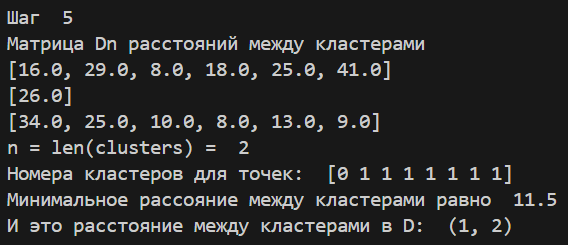








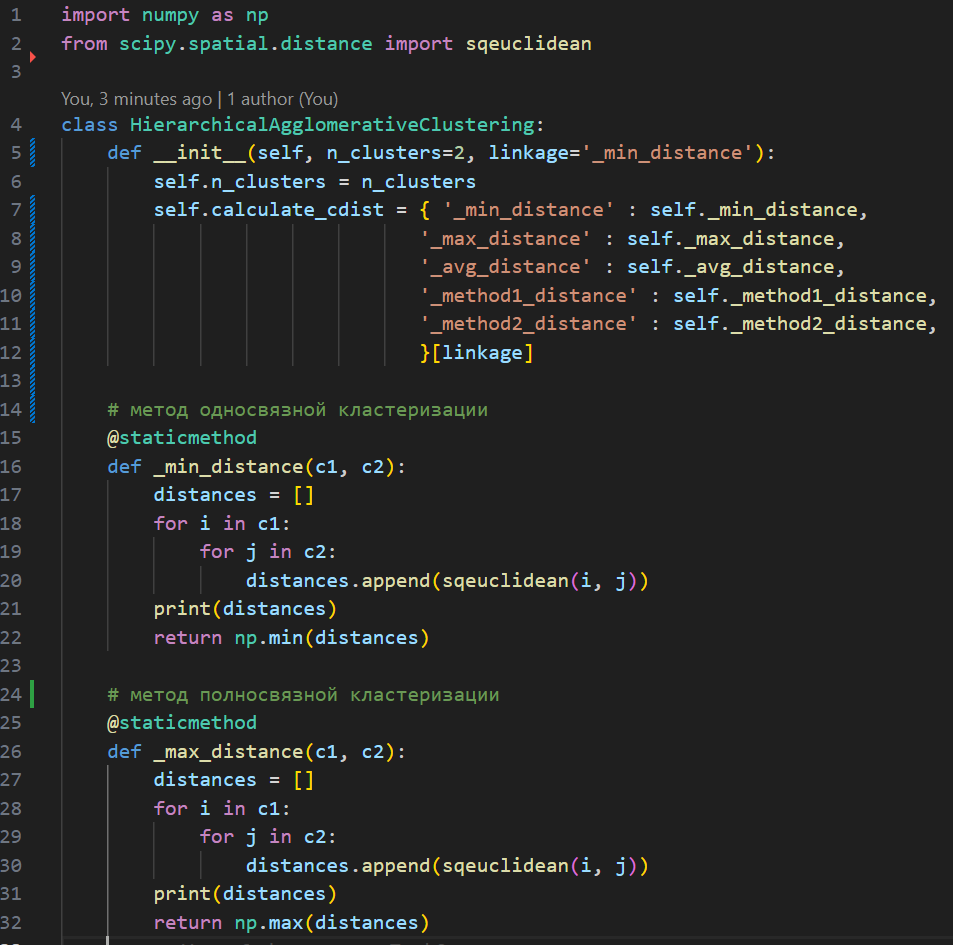




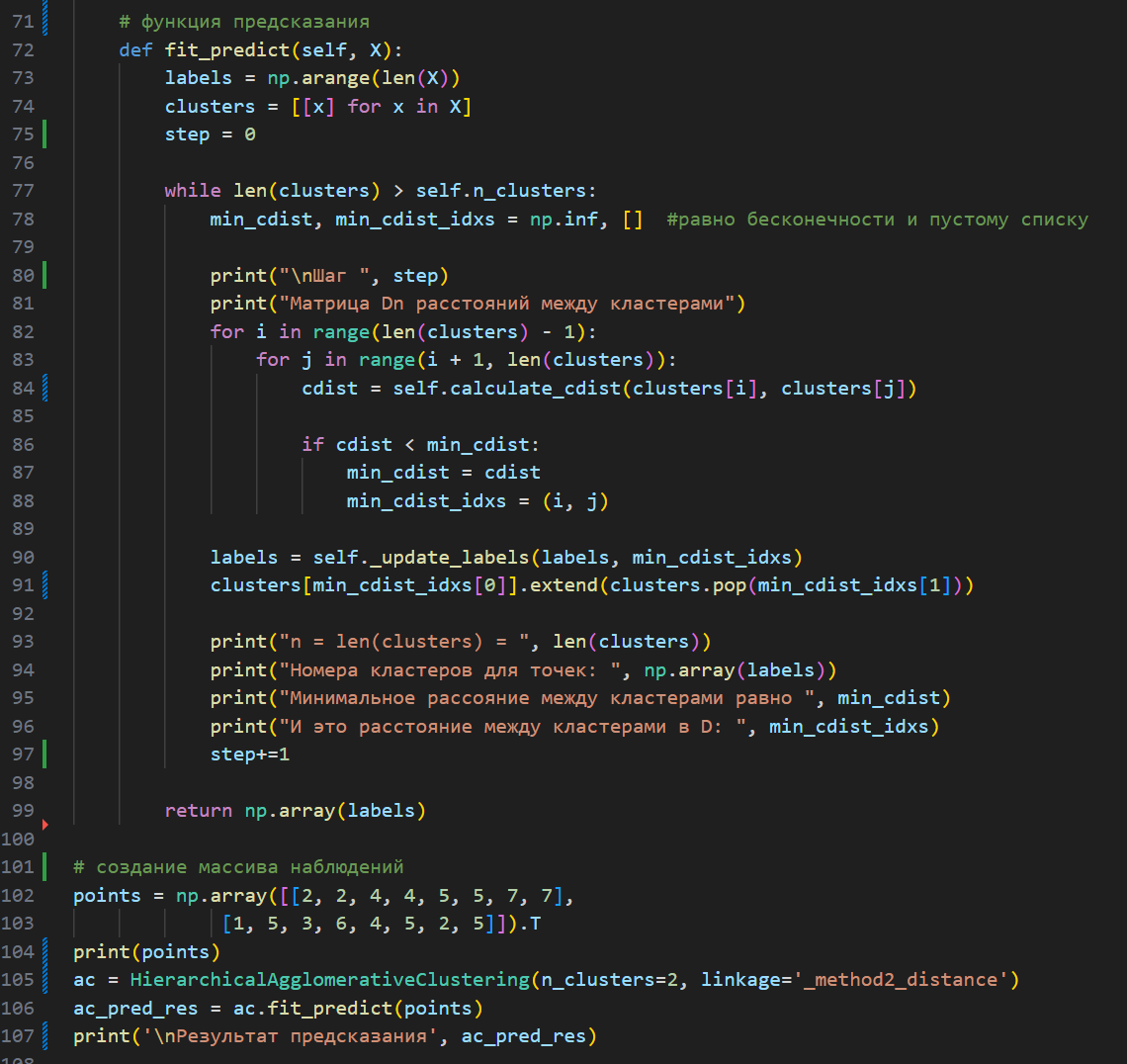
Результат предсказания методом односвязной кластеризации:



Полный код программы GusevVETask2.py.







[[2 1]

[2 5]

[4 3]

[4 6]

[5 4]

[5 5]

[7 2]

[7 5]]

Манхэттенское расстояние

[[0. 4. 4. 7. 6. 7. 6. 9.]

[4. 0. 4. 3. 4. 3. 8. 5.]

[4. 4. 0. 3. 2. 3. 4. 5.]

[7. 3. 3. 0. 3. 2. 7. 4.]

[6. 4. 2. 3. 0. 1. 4. 3.]

[7. 3. 3. 2. 1. 0. 5. 2.]

[6. 8. 4. 7. 4. 5. 0. 3.]

[9. 5. 5. 4. 3. 2. 3. 0.]]

Шаг 0

[[0, array([2, 1])]]

[[1, array([2, 5])]]

[[2, array([4, 3])]]

[[3, array([4, 6])]]

[[5, array([5, 5])], [4, array([5, 4])]]

[[6, array([7, 2])]]

[[7, array([7, 5])]]

n = len(clusters) = 7

Номера кластеров для точек: [0 1 2 3 4 4 5 6]

Рассояние между кластерами равно равно 1.0

И это расстояние между кластерами в D: (5, 4)

Матрица Dn расстояний между кластерами

[[0. 4. 4. 7. 7. 6. 9.]

[4. 0. 4. 3. 4. 8. 5.]

[4. 4. 0. 3. 3. 4. 5.]

[7. 3. 3. 0. 3. 7. 4.]

[7. 4. 3. 3. 0. 5. 3.]

[6. 8. 4. 7. 5. 0. 3.]

[9. 5. 5. 4. 3. 3. 0.]]

Шаг 1

[[0, array([2, 1])]]

[[2, array([4, 3])]]

[[3, array([4, 6])], [1, array([2, 5])]]

[[5, array([5, 5])], [4, array([5, 4])]]

[[6, array([7, 2])]]

[[7, array([7, 5])]]

n = len(clusters) = 6

Номера кластеров для точек: [0 2 1 2 3 3 4 5]

Рассояние между кластерами равно равно 3.0

И это расстояние между кластерами в D: (3, 1)

Матрица Dn расстояний между кластерами

[[0. 4. 7. 7. 6. 9.]

[4. 0. 4. 3. 4. 5.]

[7. 4. 0. 4. 8. 5.]

[7. 3. 4. 0. 5. 3.]

[6. 4. 8. 5. 0. 3.]

[9. 5. 5. 3. 3. 0.]]

Шаг 2

[[0, array([2, 1])]]

[[3, array([4, 6])], [1, array([2, 5])]]

[[5, array([5, 5])], [4, array([5, 4])], [2, array([4, 3])]]

[[6, array([7, 2])]]

[[7, array([7, 5])]]

n = len(clusters) = 5

Номера кластеров для точек: [0 1 2 1 2 2 3 4]

Рассояние между кластерами равно равно 3.0

И это расстояние между кластерами в D: (3, 1)

Матрица Dn расстояний между кластерами

[[0. 7. 7. 6. 9.]

[7. 0. 4. 8. 5.]

[7. 4. 0. 5. 5.]

[6. 8. 5. 0. 3.]

[9. 5. 5. 3. 0.]]

Шаг 3

[[0, array([2, 1])]]

[[3, array([4, 6])], [1, array([2, 5])]]

[[5, array([5, 5])], [4, array([5, 4])], [2, array([4, 3])]]

[[7, array([7, 5])], [6, array([7, 2])]]

n = len(clusters) = 4

Номера кластеров для точек: [0 1 2 1 2 2 3 3]

Рассояние между кластерами равно равно 3.0

И это расстояние между кластерами в D: (4, 3)

Матрица Dn расстояний между кластерами

[[0. 7. 7. 9.]

[7. 0. 4. 8.]

[7. 4. 0. 5.]

[9. 8. 5. 0.]]

Шаг 4

[[0, array([2, 1])]]

[[5, array([5, 5])], [4, array([5, 4])], [2, array([4, 3])], [3, array([4, 6])], [1, array([2, 5])]]

[[7, array([7, 5])], [6, array([7, 2])]]

n = len(clusters) = 3

Номера кластеров для точек: [0 1 1 1 1 1 2 2]

Рассояние между кластерами равно равно 4.0

И это расстояние между кластерами в D: (2, 1)

Матрица Dn расстояний между кластерами

[[0. 7. 9.]

[7. 0. 8.]

[9. 8. 0.]]

Шаг 5

[[0, array([2, 1])], [5, array([5, 5])], [4, array([5, 4])], [2, array([4, 3])], [3, array([4, 6])], [1, array([2, 5])]]

[[7, array([7, 5])], [6, array([7, 2])]]

n = len(clusters) = 2

Номера кластеров для точек: [0 0 0 0 0 0 1 1]

Рассояние между кластерами равно равно 7.0

И это расстояние между кластерами в D: (0, 1)

Матрица Dn расстояний между кластерами

[[0. 9.]

[9. 0.]]

Результат предсказания [0 0 0 0 0 0 1 1]