

Кластеризация на основе
плотности точек
(density based clustering)

План

1. Идея методов на основе плотности точек
2. Пример основных, граничных и шумовых точек
3. DBSCAN
4. Пример работы DBSCAN
5. Определение числа кластеров
6. Настройка параметров DBSCAN

Идея density-based методов

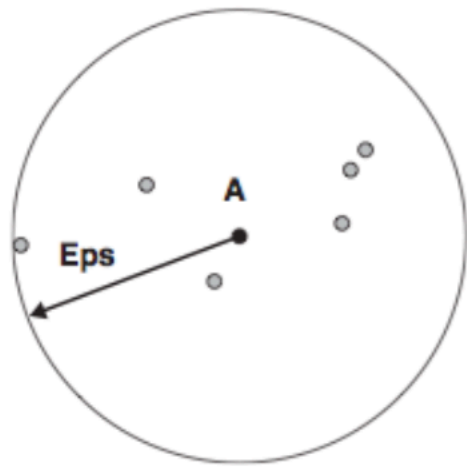


Figure 8.20. Center-based density.

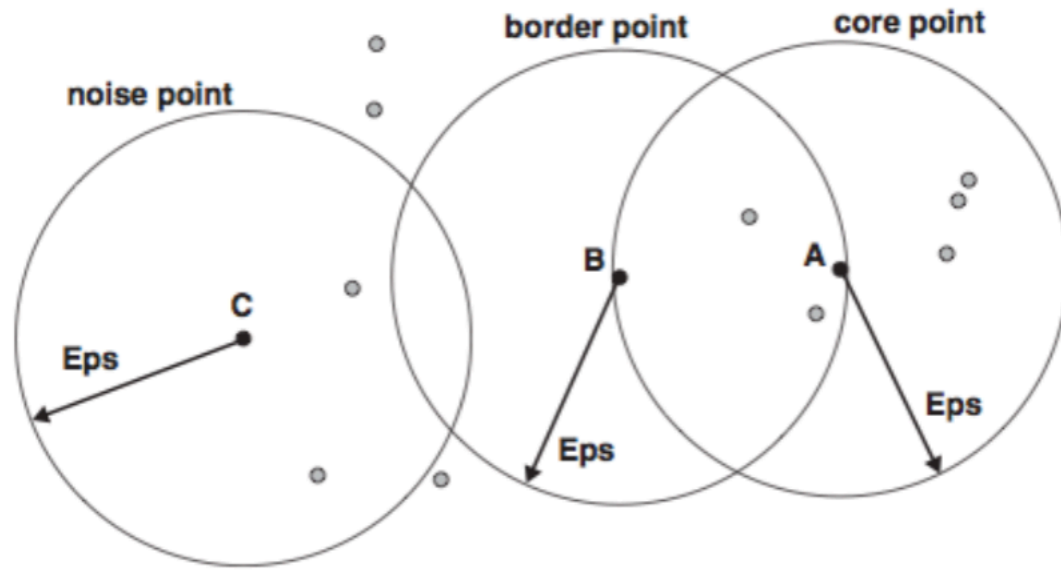
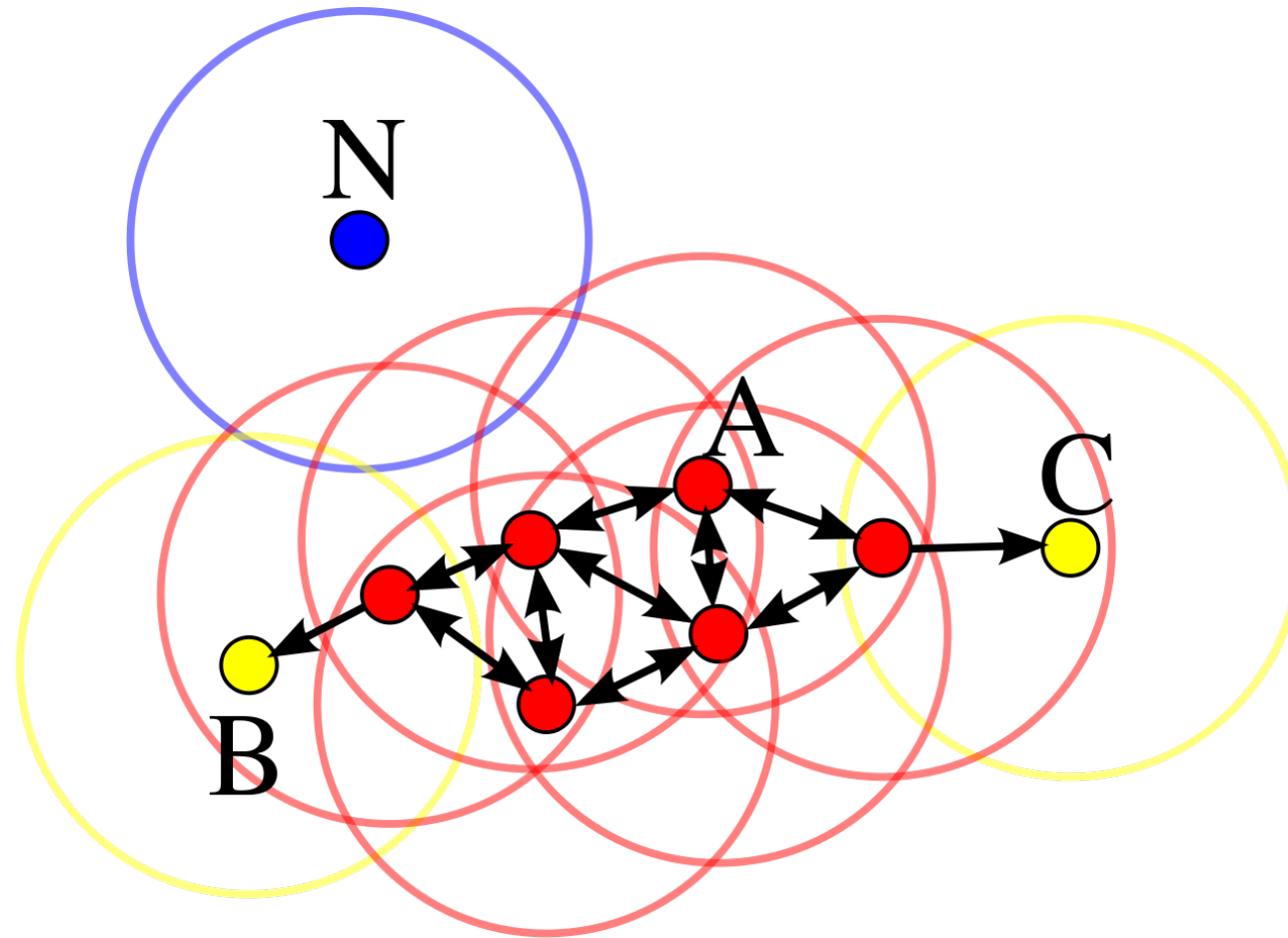


Figure 8.21. Core, border, and noise points.

Основные, шумовые и граничные точки

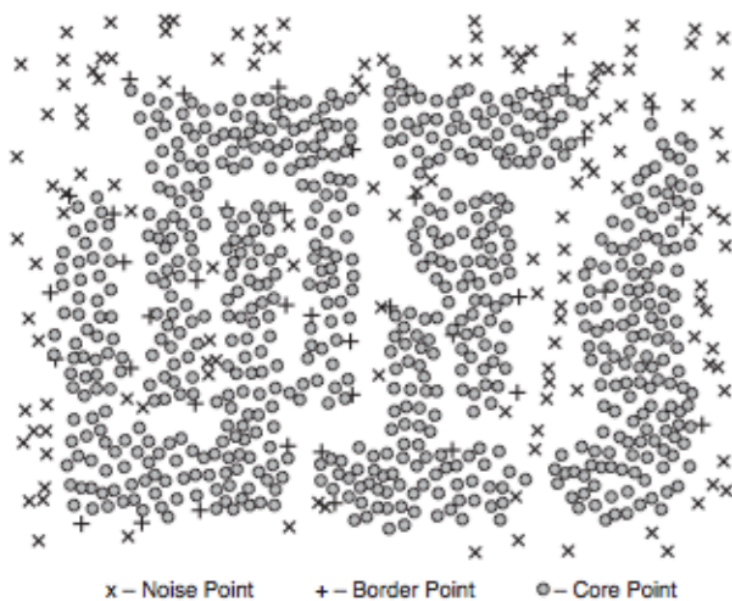


DBSCAN

1: Пометить все точки, как основные, пограничные или шумовые.



(a) Clusters found by DBSCAN.



(b) Core, border, and noise points.

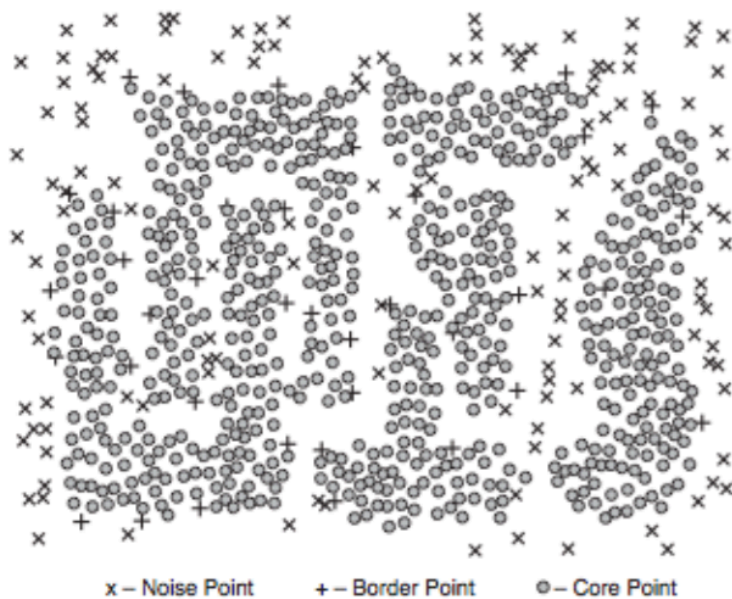
DBSCAN

1: Пометить все точки, как основные, пограничные или шумовые.

2: Отбросить точки шума.



(a) Clusters found by DBSCAN.



(b) Core, border, and noise points.

DBSCAN

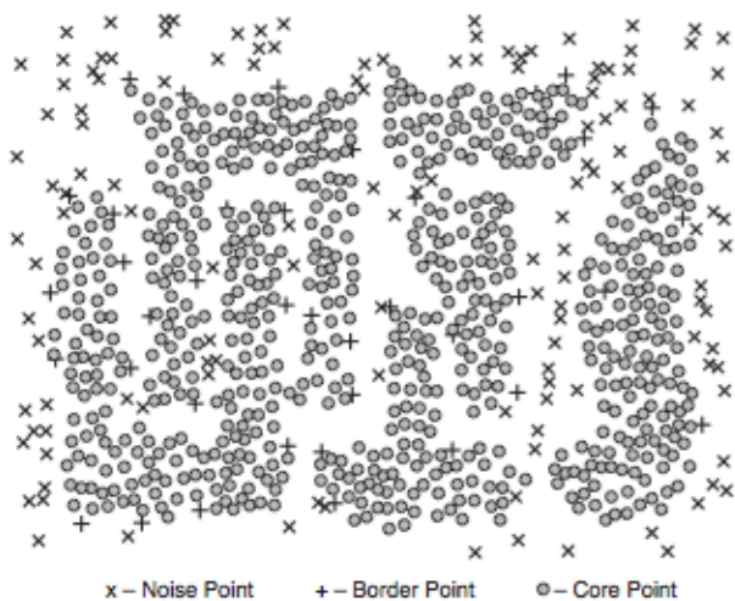


(a) Clusters found by DBSCAN.

1: Пометить все точки, как основные, пограничные или шумовые.

2: Отбросить точки шума.

3: Соединить все основные точки, находящиеся на расстоянии Eps радиуса одна от другой.

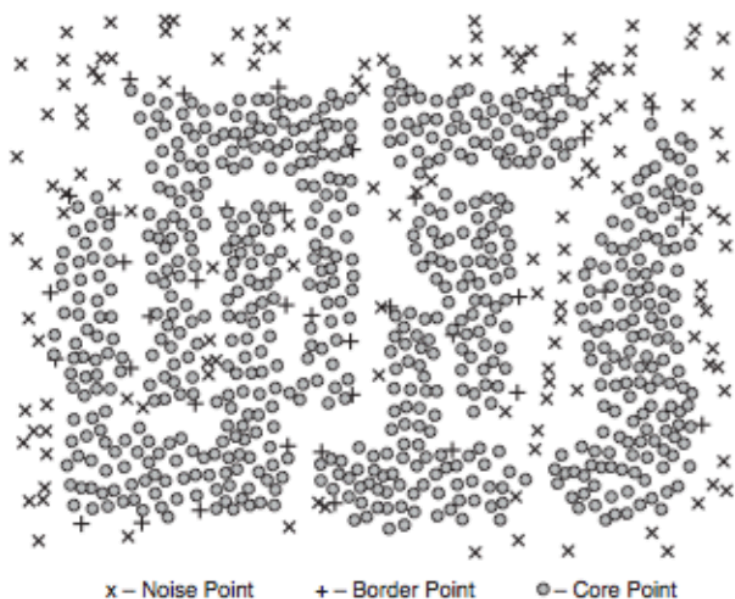


(b) Core, border, and noise points.

DBSCAN



(a) Clusters found by DBSCAN.



(b) Core, border, and noise points.

1: Пометить все точки, как основные, пограничные или шумовые.

2: Отбросить точки шума.

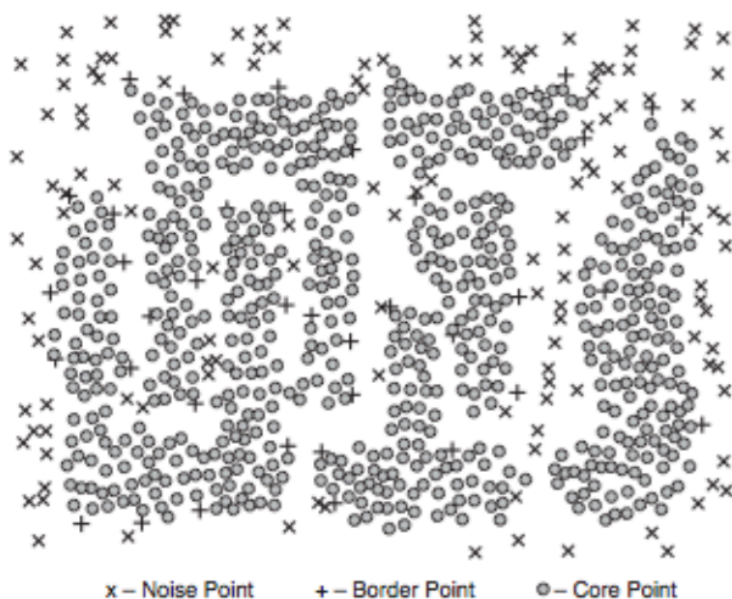
3: Соединить все основные точки, находящиеся на расстоянии Eps радиуса одна от другой.

4: Объединить каждую группу соединенных основных точек в отдельный кластер.

DBSCAN



(a) Clusters found by DBSCAN.



(b) Core, border, and noise points.

1: Пометить все точки, как основные, пограничные или шумовые.

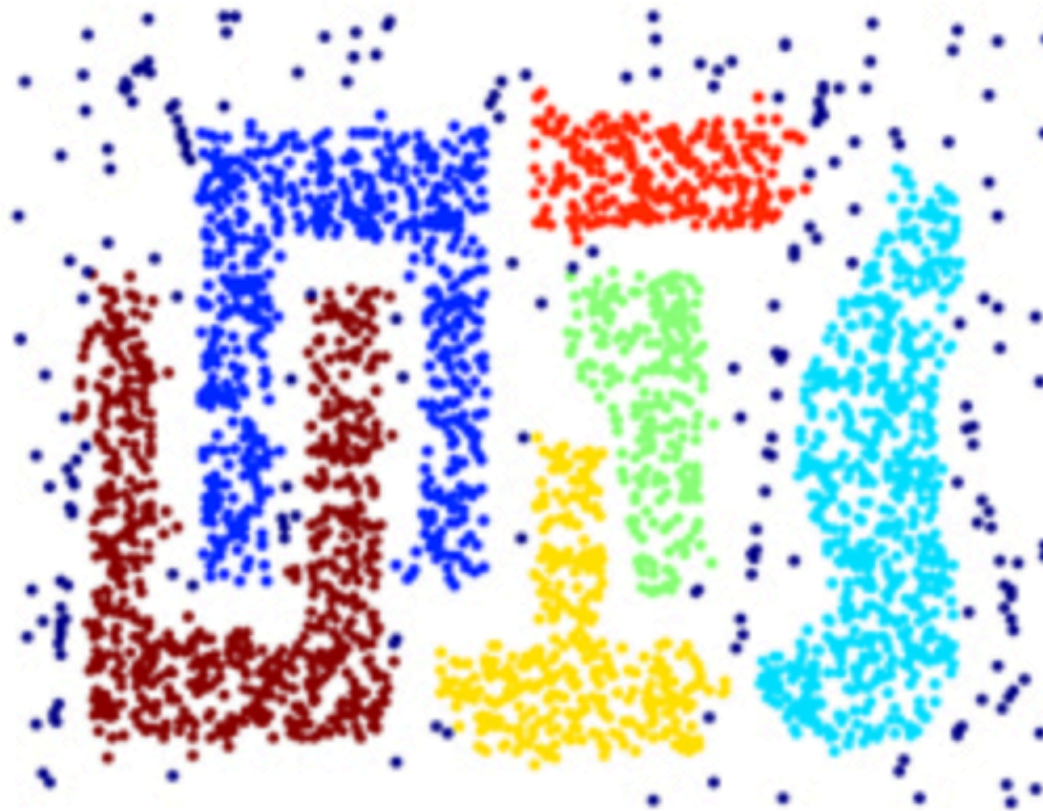
2: Отбросить точки шума.

3: Соединить все основные точки, находящиеся на расстоянии Eps радиуса одна от другой.

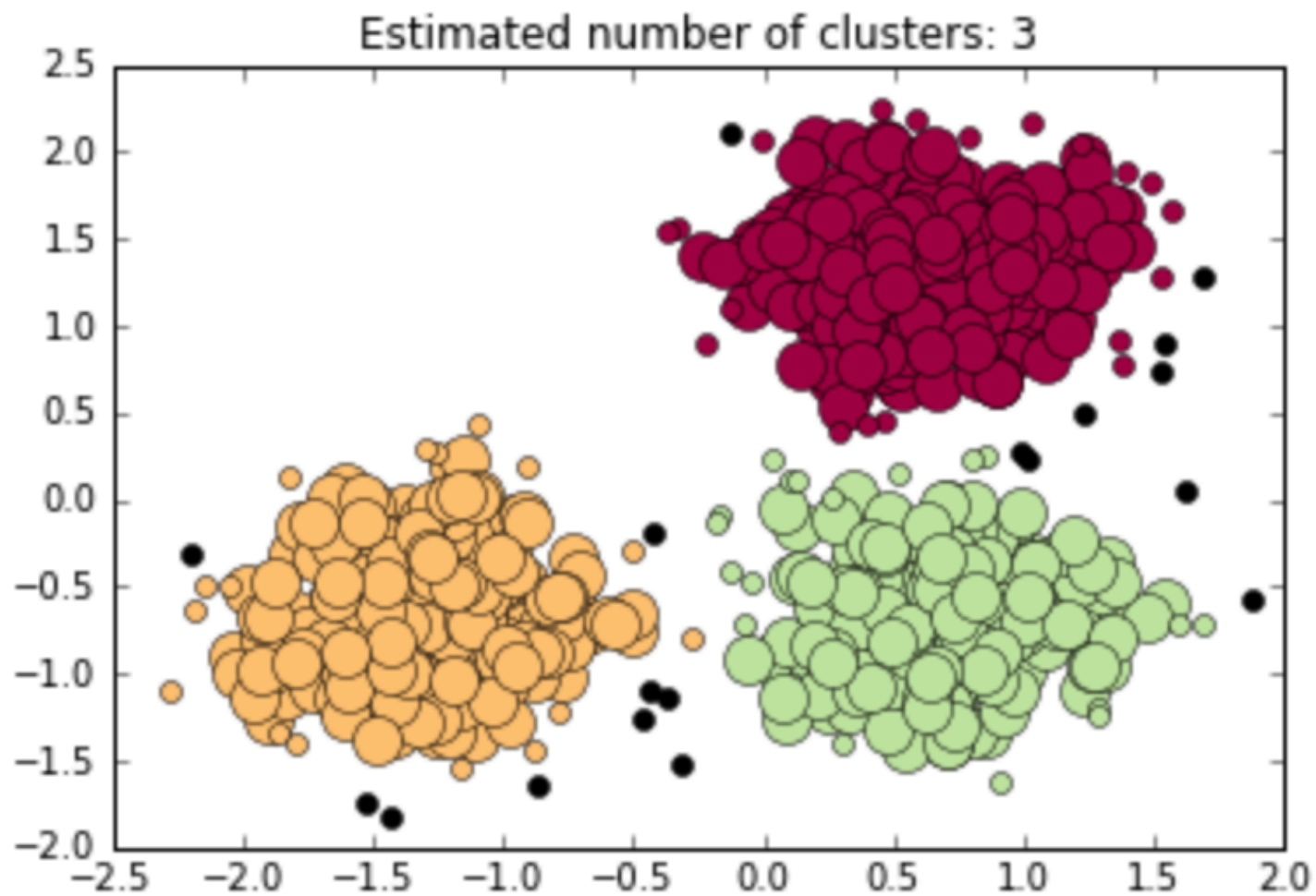
4: Объединить каждую группу соединенных основных точек в отдельный кластер.

5: Назначить каждую пограничную точку одному из кластеров, ассоциированных с ней основных точек.

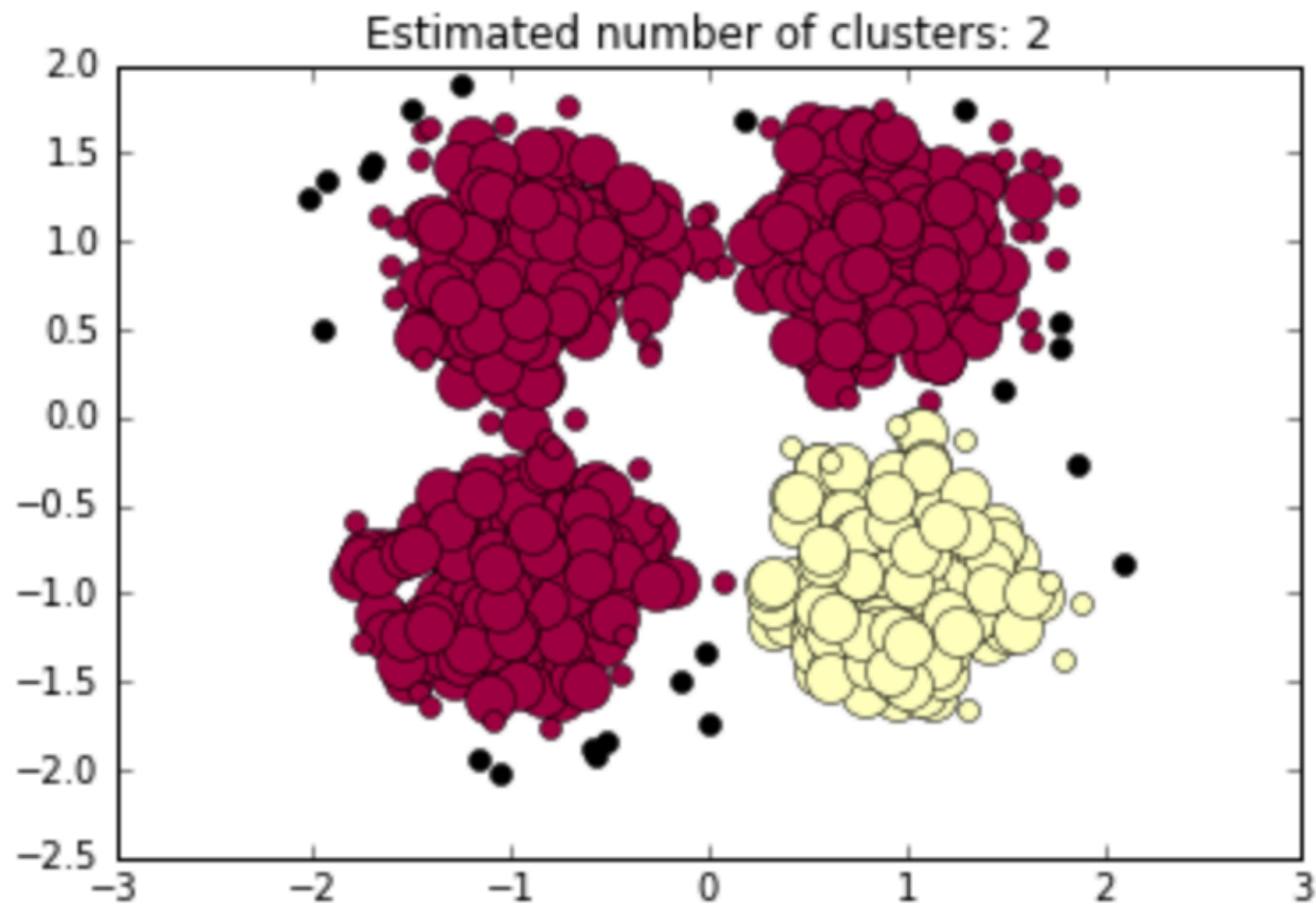
DBSCAN: результаты работы



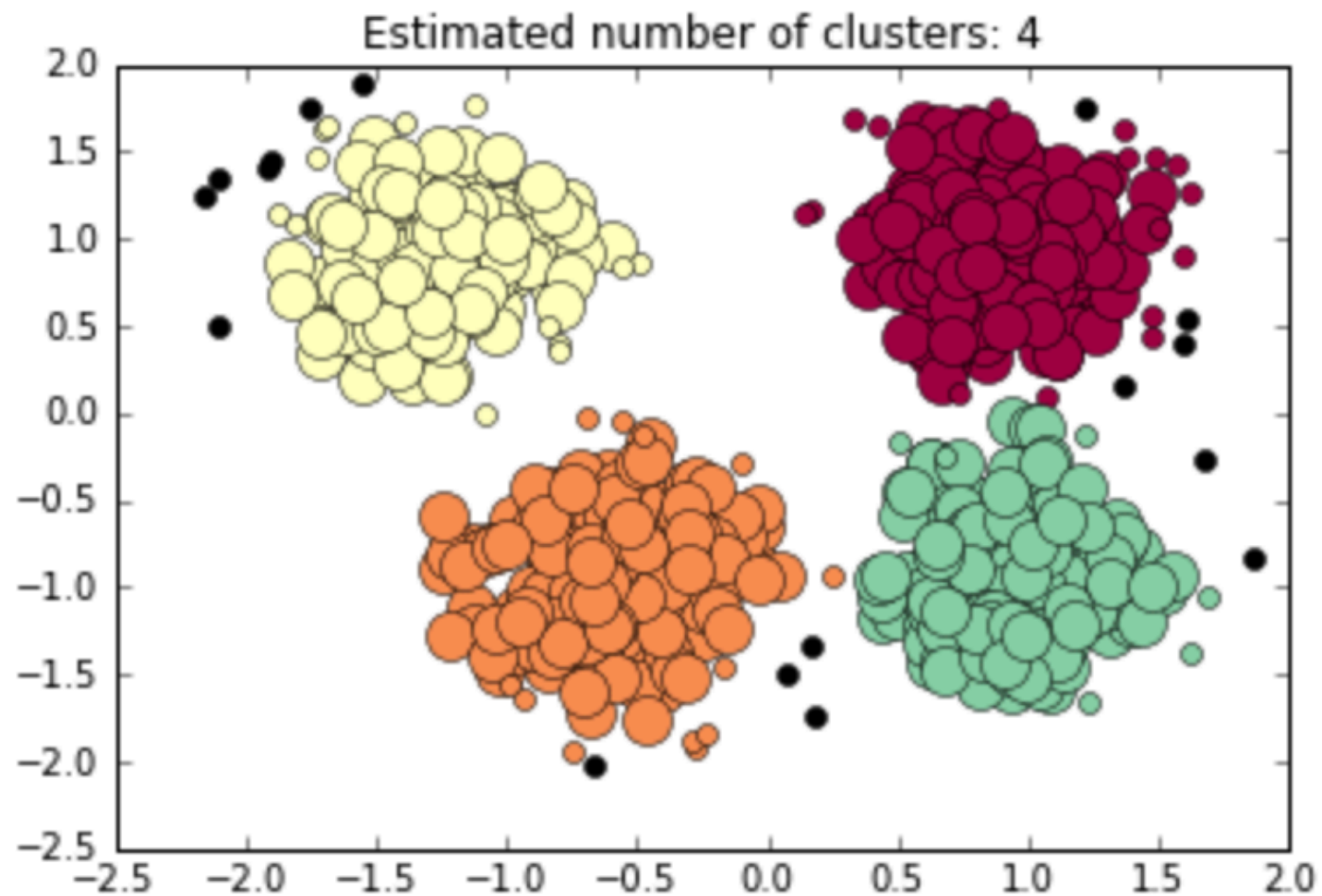
Определение числа кластеров



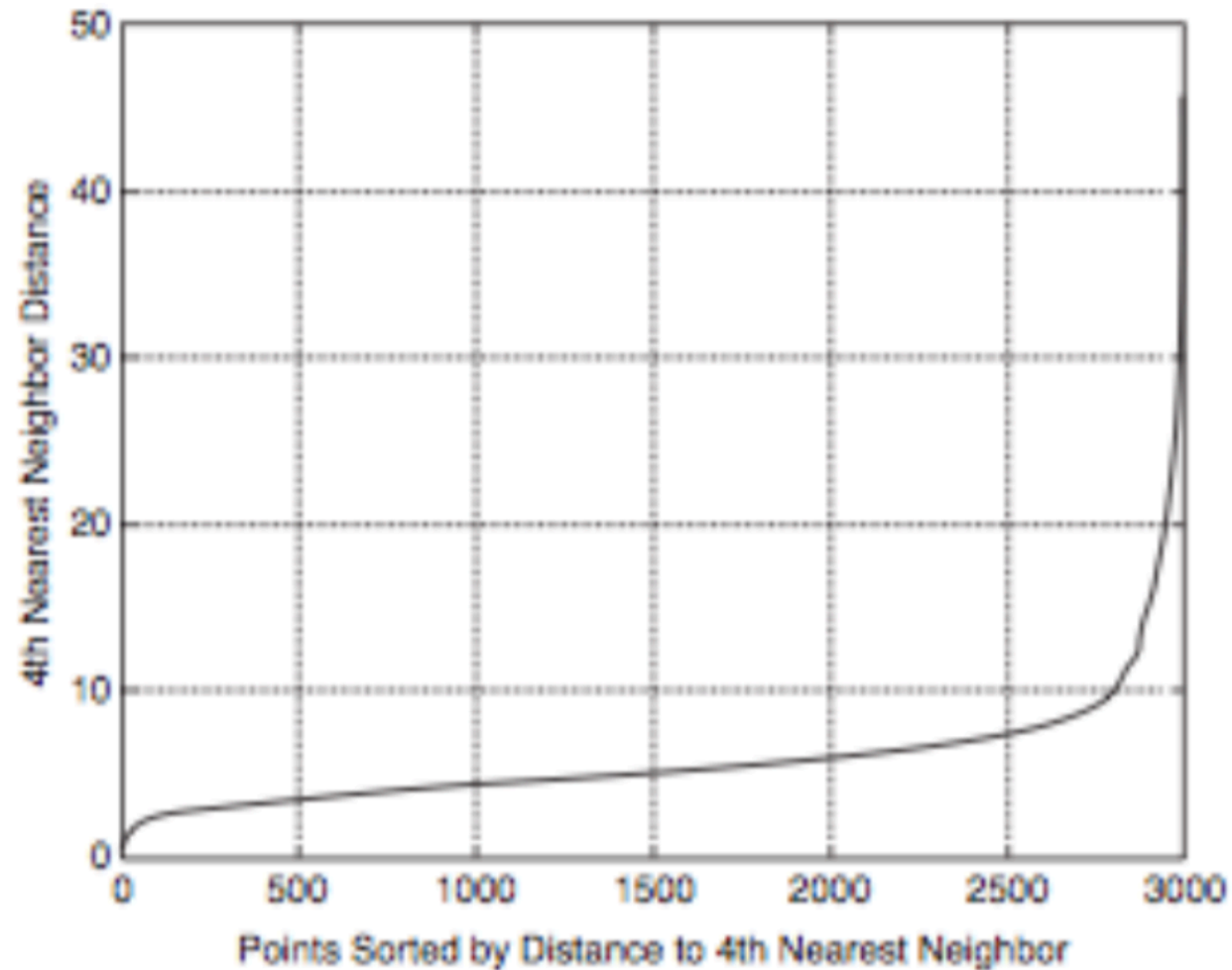
Определение числа кластеров



Определение числа кластеров



DBSCAN: подбор параметров



Резюме

1. Идея методов на основе плотности точек
2. Пример основных, граничных и шумовых точек
3. DBSCAN
4. Пример работы DBSCAN
5. Определение числа кластеров
6. Настройка параметров DBSCAN