

# Метод k-средних. Реализация алгоритма

# Цель

Научиться имплементировать алгоритм кластеризации k-средних с помощью numpy.

# Постановка задачи

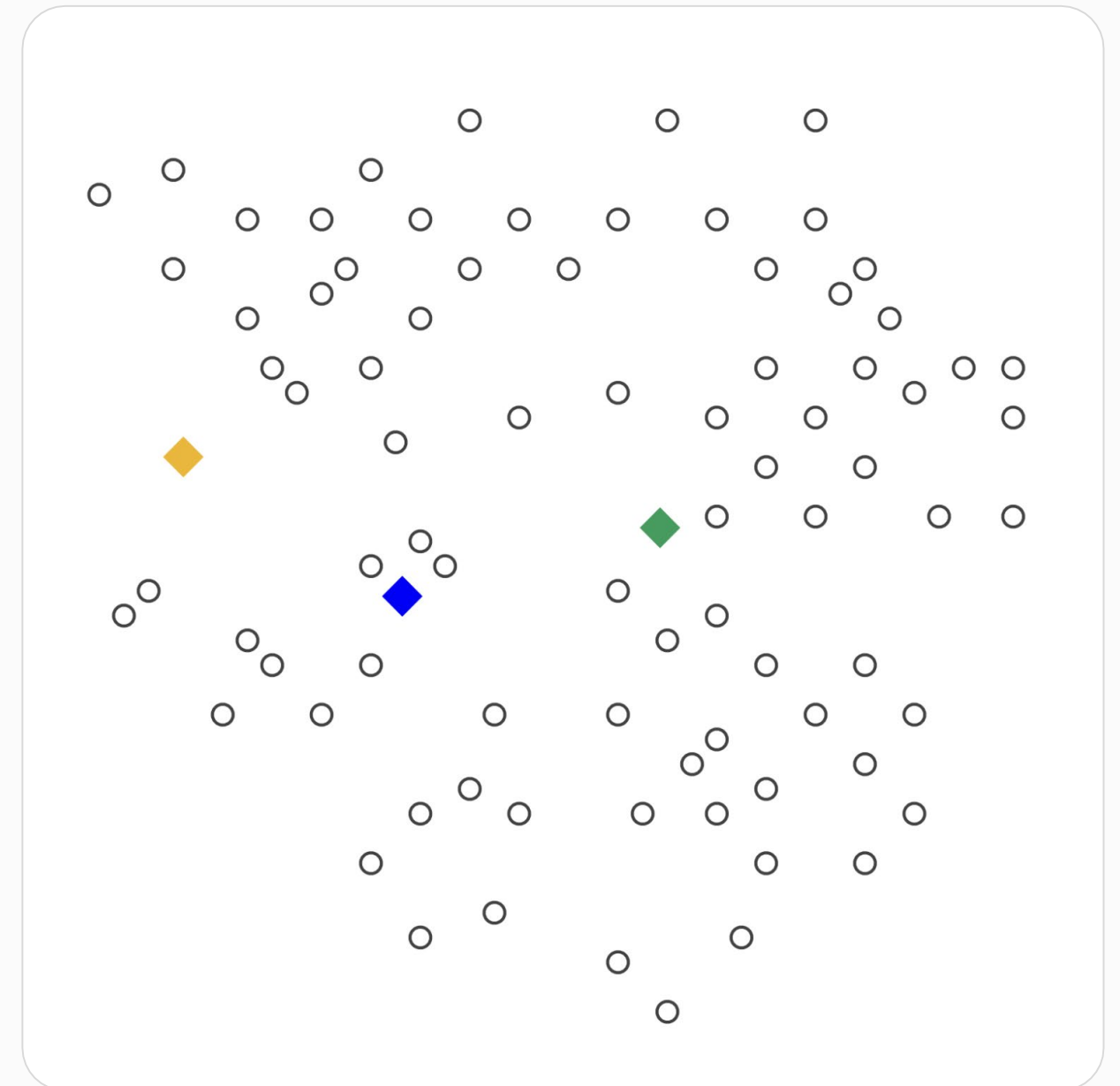
Вспомните процесс обучения алгоритма k-средних.

Процесс обучения алгоритма k-средних является итеративным, то есть циклическим.

Первый цикл обучения состоит из следующих шагов.

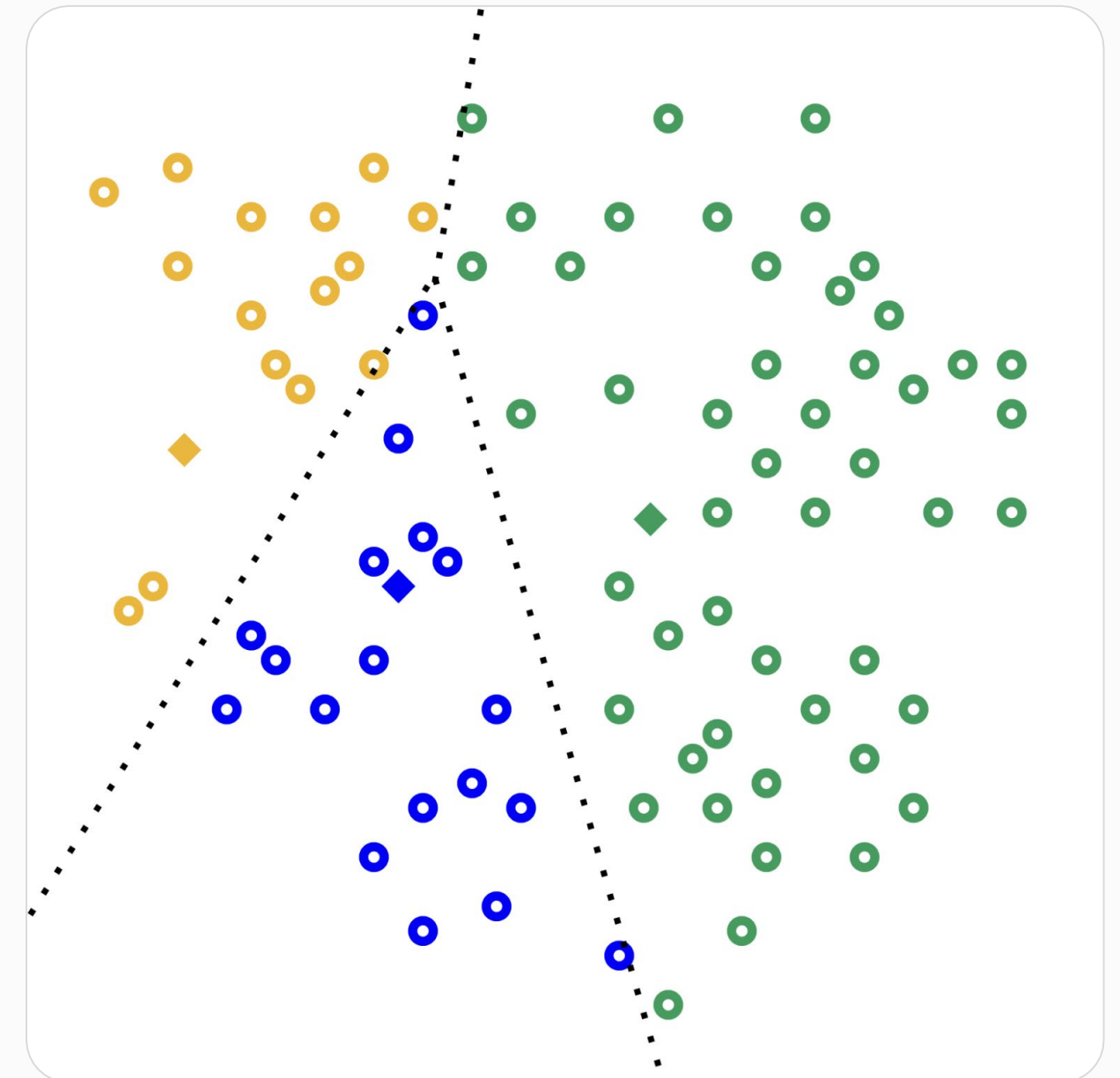
# Процесс обучения алгоритма

- 1 На первом шаге алгоритм случайным образом создаёт  $k$  центроид, то есть центры будущих кластеров



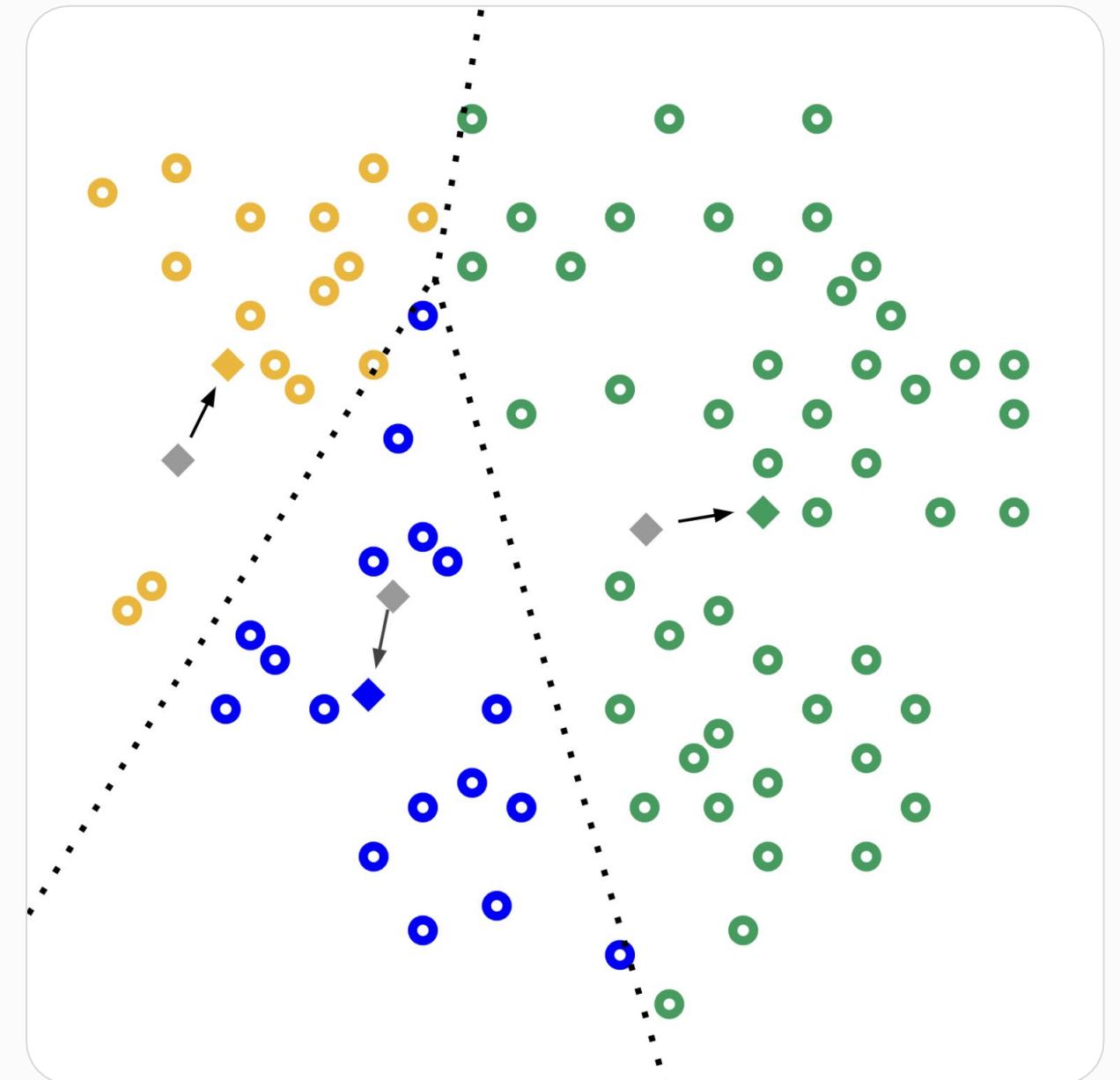
# Процесс обучения алгоритма

- 2 Затем алгоритм приписывает каждую точку данных к ближайшей центроиде, формируя тем самым k кластеров



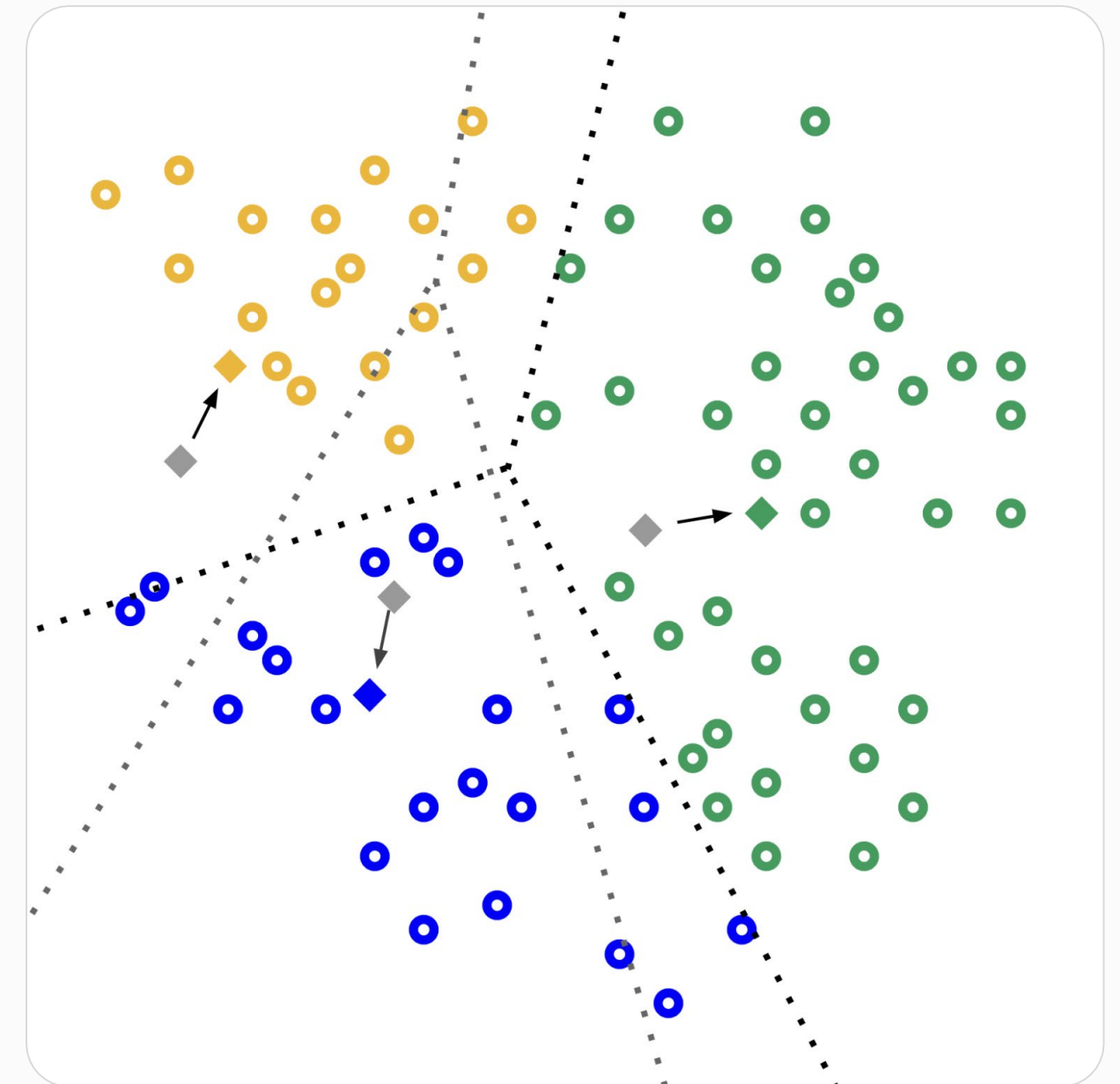
# Процесс обучения алгоритма

- 3 Для каждого кластера алгоритм вычисляет новую центроиду как среднее значение всех точек в этом кластере



# Процесс обучения алгоритма

- 4 Поскольку местоположение центроид изменилось, алгоритм заново приписывает каждую точку уже к новой ближайшей центроиде



# Процесс обучения алгоритма

После этого шаги 2–4 повторяются до тех пор, пока на какой-либо итерации на четвёртом шаге все точки будут приписаны тем же центроидам, что и на предыдущем этапе.

Напишите для этого алгоритма код, чтобы разобраться, как он работает внутри.



# Вывод

Изучили, как имплементировать алгоритм k-средних с нуля с помощью numpy, и разобрали, как он работает под капотом.