

K-Means Clustering:

1. Pertama, centroid (pusat cluster) akan diinisialisasi sesuai dengan parameter yang diberikan. Jika random, centroid akan diambil secara acak juga. Jika menggunakan `kmeans++`, centroid pertama dipilih secara acak, selanjutnya pilih centroid yang paling jauh jaraknya (dengan Euclidean distance) dari centroid yang sudah ada. Kemudian, untuk setiap titik dalam dataset, hitung jaraknya dari setiap centroid, tiap titik akan ditugaskan ke centroid terdekat sehingga terbentuk cluster, tiap centroid akan diperbarui dengan menghitung rata-rata dari semua titik data yang telah ditugaskan ke suatu cluster. Jika tidak ada perubahan centroid, algoritma dianggap telah konvergen dan berhenti.
2. Hasil *clustering* dari K-Mean Clustering *scratch* dengan menggunakan `kmeans++` sebagai inisialisasinya memberikan hasil yang hampir menyerupai model K-Mean Clustering bawaan dari scikit learn sehingga dapat disimpulkan model yang dibuat sudah benar.