

1. Algoritma K-Means adalah metode clustering (unsupervised learning) yang berfungsi untuk mengelompokkan data ke dalam cluster-cluster berdasarkan kesamaan karakteristik. K-Means bekerja dengan melakukan dua langkah utama: assignment data points ke cluster terdekat dan update centroid cluster. Cara kerja
  - a. Inisialisasi Parameter
    - i. `n_clusters`: Jumlah cluster yang diinginkan (default: 3)
    - ii. `max_iters`: Maksimum iterasi untuk konvergensi (default: 100)
    - iii. `init_method`: Metode inisialisasi centroid ('random' atau 'kmeans++')
  - b. Inisialisasi Centroid (`_initialize_centroids`)
    - i. Random: Memilih sampel acak dari data sebagai centroid awal
    - ii. K-means++: Metode yang menyebar centroid awal untuk hasil yang lebih baik dengan memastikan bahwa titik-titik pusat awal (centroid) yang dipilih saling berjauhan satu sama lain
  - c. Training Loop (fit)
    - i. Inisialisasi centroid awal
    - ii. Ulangi hingga konvergen atau mencapai max iterations:
      1. Assign data points ke cluster terdekat
      2. Update centroid berdasarkan mean data points dalam cluster
      3. Berhenti jika centroid tidak berubah
  - d. Cluster Assignment (`_assign_clusters`)
    - i. Hitung jarak setiap data point ke semua centroid
    - ii. Assign data point ke cluster dengan centroid terdekat
  - e. Centroid Update (`_update_centroids`)
    - i. Hitung mean dari semua data points dalam setiap cluster
    - ii. Gunakan mean baru sebagai centroid yang diperbarui
2. Berdasarkan hasil evaluasi, model dari Sklearn memiliki nilai Silhouette 0.2058, sementara untuk model yang saya buat sendiri mendapatkan nilai Silhouette 0.2058. Hal ini menunjukkan performa model yang sudah sama persis.