

1. Algoritma DBSCAN adalah metode clustering berbasis kepadatan yang dapat mengidentifikasi cluster dengan bentuk yang fleksibel dan secara otomatis mendeteksi outliers/noise. Berbeda dengan algoritma K-Means yang berbasis centroid, DBSCAN bekerja dengan mengelompokkan titik-titik data yang memiliki kepadatan tinggi yang dipisahkan oleh area kepadatan rendah. Cara kerja
  - a. Inisialisasi Parameter
    - i.  $\epsilon$  ( $\epsilon$ ): Jarak maksimum antara dua titik untuk dianggap bertetangga
    - ii. min\_samples: Minimum titik yang diperlukan untuk membentuk core point
    - iii. metric: Metode perhitungan jarak ('euclidean', 'manhattan', 'minkowski')
    - iv. p: Parameter untuk metric Minkowski
  - b. Inisialisasi State (fit)
    - i. Inisialisasi semua label sebagai unvisited (0)
    - ii. Siapkan counter untuk cluster ID
  - c. Main Clustering Loop
    - i. Untuk setiap titik data:
      1. Jika sudah dikunjungi, skip
      2. Cari semua tetangga dalam radius  $\epsilon$
      3. Jika jumlah tetangga < min\_samples: label sebagai noise (-1)
      4. Jika cukup tetangga: expand cluster
  - d. Cluster Expansion (\_expand\_cluster)
    - i. Label titik sebagai cluster tertentu
    - ii. Iterasi melalui semua tetangga:
      1. Jika tetangga adalah noise, ubah menjadi border point
      2. Jika tetangga unvisited, tambahkan ke cluster dan cari tetangganya
      3. Jika tetangga adalah core point, tambahkan tetangganya ke queue
2. Berdasarkan hasil evaluasi, model dari Sklearn memiliki nilai Silhouette 0.3578, sementara untuk model yang saya buat sendiri mendapatkan nilai Silhouette 0.3578. Hal ini menunjukkan performa model yang sudah sama persis.