

Clustering merupakan

---

- ☐ A. Proses pengenalan sebuah data berdasarkan labelnya, dimana jumlah label terbatas.
  - ☐ B. Proses prediksi data ke dalam label kontinyu
  - ☒ C. Proses pengelompokan data berdasarkan kemiripannya.
  - ☐ D. Proses pengenalan sebuah data berdasarkan labelnya, dimana jumlah label tidak terbatas.
  - ☐ E. Tidak ada yang benar.
- 

Clustering ditujukan untuk menemukan kelompok alami dari data dengan karakteristik sebagai berikut (Jawaban bisa lebih dari satu):

---

- ☒ A. Pengelompokan data dilakukan berdasarkan kemiripan karakteristiknya.
- ☒ B. High intra-cluster similarity, yaitu data pada kelompok yang sama harus semirip mungkin.
- ☒ C. Low inter-cluster similarity, data pada kelompok yang berbeda memiliki jarak data sejauh mungkin.
- ☒ D. Metrik pengukuran jarak ditentukan dengan jelas.
- ☒ E. Metrik pengukuran disesuaikan dengan semantiknya pada domain yang bersesuaian.

Clustering dilakukan jika:

---

- ☐ A. Jumlah kelompok data tidak diketahui
- ☐ B. Label dari masing-masing data tidak diketahui
- ☐ C. Jumlah kelompok data bisa diketahui bisa tidak
- ☐ D. A dan B
- ☒ E. B dan C

Berapa jumlah minimum variabel atau fitur yang dibutuhkan untuk melakukan clustering?

- ☐ 0
- ☒ 1
- ☐ 2
- ☐ 3
- ☐ Bebas

Contoh kasus dimana teknik pembelajaran mesin dengan menggunakan clustering tepat dilakukan:

- ☐ A. Memprediksi jumlah penderita covid-19 pada waktu tertentu di wilayah tertentu.
- ☐ B. Mengidentifikasi e-mail fraud
- ☒ C. Mengelompokkan artikel berdasarkan topiknya dengan topik-topiknya belum diketahui sebelumnya.
- ☐ D. Mengenali kalimat-kalimat dengan sentimen negatif
- ☐ E. Mengelompokkan pengguna berdasarkan jenis kelaminnya.

Berikut ini tahap-tahap yang harus dilakukan dalam proses clustering: 1. Melakukan perhitungan kemiripan dengan menggunakan metrik jarak tertentu. 2. Mengumpulkan data dimana data belum ada labelnya. 3. Menganalisis kelompok data yang terbentuk dan memberikan label data jika memungkinkan; 4. Mengelompokkan data; 5. Menentukan fitur yang akan digunakan, melakukan seleksi dan/atau ekstraksi fitur. Urutan tahap yang benar adalah:

- ☐ A. 1-2-3-4-5
- ☐ B. 2-4-3-5-1
- ☐ C. 2-5-4-3-1
- ☒ D. 2-5-1-4-3
- ☐ E. 2-5-1-3-4

Berikut ini adalah pernyataan yang benar terkait algoritma K-Means, kecuali:

- ☒ A. Algoritma K-Means termasuk ke dalam teknik clustering berbasis densitas.
- ☐ B. Algoritma K-Means termasuk ke dalam teknik clustering berbasis partisi.
- ☐ C. Hasil pengelompokkan dengan algoritma K-Means sangat dipengaruhi oleh inisialisasi centroid.
- ☐ D. Algoritma K-Means menggunakan squared error clustering
- ☐ E. Prosesnya sensitif terhadap outliers.

Berikut ini adalah tahap-tahap yang dilakukan pada algoritma K-Means: 1. Menentukan jumlah kelompok yang akan digunakan; 2. Melakukan perhitungan jarak antara setiap data dengan setiap centroid menggunakan metrik jarak tertentu, misal Euclidian distance; 3. Menentukan centroid awal secara random sejumlah cluster yang akan dibangun; 4. Menentukan centroid baru yaitu hasil perhitungan means dari kelompok yang terbentuk; 5. Mengelompokkan data berdasarkan centroid ; 6. Lakukan proses pengelompokkan sampai konvergen. Urutan tahap yang benar adalah:

- ☐ A. 1-2-3-4-5-6
- ☐ B. 1-3-4-5-6-2
- ☒ C. 1-3-2-5-4-6
- ☐ D. 1-3-2-4-5-6
- ☐ E. Tidak ada yang benar.

Hal-hal apa saja yang dapat digunakan sebagai kondisi berhenti pada K-Means? (Jawaban bisa lebih dari 1)

- ☒ A. Jumlah iterasi yang telah ditetapkan (biasanya jika data sangat banyak dan waktu untuk mencapai konvergen lama)
- ☒ B. Penetapan observasi ke kluster tidak berubah antara iterasi, kecuali untuk kasus dengan minimum lokal (kondisi sudah dinyatakan konvergen)
- ☒ C. Centroid tidak berubah antara iterasi yang berurutan
- ☒ D. Perubahan centroid sudah di bawah threshold perubahan yang ditetapkan
- ☐ E. Akurasi pengelompokkan sudah mencapai 100%

Anda ingin mengelompokkan 7 pengamatan ke dalam 3 kelompok menggunakan algoritma K-Means. Setelah cluster iterasi pertama, C1, C2, C3 memiliki pengamatan berikut: C1: {(2,2), (4,4), (6,6)}, C2: {(0,4), (4,0)}, dan C3: {(5,5), (9,9)}. Tentukan centroid cluster jika Anda ingin melanjutkan pengelompokkan sampai iterasi kedua?

- ☐ A. C1: (6,6), C2: (4,4), C3: (9,9)
- ☒ B. C1: (4,4), C2: (2,2), C3: (7,7)
- ☐ C. C1: (4,4), C2: (7,7), C3: (2,2)
- ☐ D. C1: (2,2), C2: (0,0), C3: (5,5)
- ☐ E. Tidak ada yang benar.

Algoritma mana yang paling sensitif terhadap keberadaan outliers?

---

- ☒ A. K-Means clustering
- ☐ B. K-Medoids clustering
- ☐ C. K-Medians clustering
- ☐ D. Fuzzy C-Means
- ☐ E. Tidak ada yang benar

Berikut ini yang merupakan kelemahan dari algoritma K-Means, kecuali:

---

- ☐ A. Tidak ada panduan penentuan nilai k yang baik
- ☐ B. Hasil sangat dipengaruhi oleh inisialisasi centroid, bisa jadi berhenti pada optimum lokal
- ☒ C. Hanya dapat digunakan untuk atribut non-numerik
- ☐ D. Algoritma tidak scalable
- ☐ E. Prosesnya sensitif terhadap outliers.

13. Berikut ini merupakan kelebihan dari algoritma K-Modes, kecuali:

---

- ☒ A. Prosesnya lebih cepat dibanding K-Means
- ☐ B. Lebih robust karena dapat menangani atribut non-numerik
- ☐ C. Tidak sensitif terhadap outliers
- ☐ D. Perhitungan jarak menggunakan kriteria error absolut
- ☐ E. Tidak ada yang benar.

Pada PAM (Partitioning Around Medoids), pertukaran medoid dilakukan jika:

---

- ☐ A. Anggota cluster memiliki jarak yang lebih dekat ke medoid yang baru
- ☐ B. Total absolut error akibat pertukaran lebih besar daripada error sebelumnya
- ☒ C. Total absolut error akibat pertukaran lebih kecil daripada error sebelumnya
- ☐ D. Semua data sudah pernah menjadi kandidat medoid
- ☐ E. Tidak ada yang benar

Berikut ini pernyataan yang benar terkait algoritma Graph-Theoretic clustering:

---

- ☐ A. Menggunakan minimal spanning tree (MST)
- ☐ B. Cluster yang dibuat dengan memotong hubungan yang paling dekat antar cluster secara berurutan hingga jumlah cluster yang diinginkan terbentuk.
- ☐ C. Cluster yang dibuat dengan memotong hubungan yang paling jauh antar cluster secara berurutan hingga jumlah cluster yang diinginkan terbentuk.
- ☒ D. A dan C
- ☐ E. A dan B