

Klasterisasi Makanan Berdasarkan Kandungan Gizi Menggunakan Algoritma K-Medoids

Edward Rajagukguk¹, Desria Adventina Sinaga², Alfonsus Situmorang³

^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, 5 Juni 2026
Revised, 26 Juli 2026
Accepted, 30 Juli 2026

Keywords:

Data mining,
K-Medoids,
Klasterisasi,
Kandungan gizi,
Karakteristik klaster.

ABSTRAK

Informasi kandungan gizi makanan pada umumnya disajikan dalam bentuk data numerik tanpa pengelompokan sehingga menyulitkan pembaca mengenali makanan yang memiliki karakteristik gizi serupa. Penelitian ini bertujuan menerapkan algoritma K-Medoids untuk mengelompokkan makanan berdasarkan kemiripan kandungan gizinya, mengevaluasi hasil pengelompokan menggunakan koefisien silhouette, serta menganalisis karakteristik kandungan gizi pada setiap klaster. Data yang digunakan berupa 1.126 data makanan dari platform Kaggle dengan 21 atribut kandungan gizi. Tahapan penelitian meliputi pembersihan dan seleksi data, normalisasi menggunakan Min-Max Scaling, penentuan jumlah klaster optimal berdasarkan koefisien silhouette untuk dua sampai tujuh klaster, penerapan algoritma K-Medoids dengan jarak Euclidean, uji sensitivitas terhadap pemilihan medoid awal, serta interpretasi karakteristik klaster berdasarkan nilai rata-rata atribut pada data asli. Hasil penelitian menunjukkan bahwa jumlah klaster optimal adalah dua dengan koefisien silhouette sebesar 0,3947, yang termasuk kategori struktur lemah. Klaster 0 beranggotakan 807 makanan dan Klaster 1 beranggotakan 319 makanan. Klaster 1 memiliki rata-rata lebih tinggi pada energi, lemak, karbohidrat, serat pangan, serta sebagian besar vitamin dan mineral, sedangkan Klaster 0 memiliki rata-rata lebih tinggi pada kolesterol, vitamin A, dan natrium. Enam nilai random seed yang berbeda menghasilkan medoid akhir, biaya total, dan koefisien silhouette yang sama sehingga hasil pengelompokan tergolong konsisten. Temuan ini menunjukkan bahwa algoritma K-Medoids mampu membentuk dua kelompok makanan dengan karakteristik gizi yang berbeda, meskipun keterpisahan antarklaster masih lemah.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Koresponden:

Edward Rajagukguk,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: edw4rd.raja@gmail.com

1. PENDAHULUAN

Perkembangan teknologi informasi meningkatkan ketersediaan data mengenai makanan dan kandungan gizinya, baik dari lembaga resmi maupun platform digital. Informasi energi, protein, lemak, karbohidrat, vitamin, dan mineral berbagai jenis makanan kini mudah diakses. Namun, informasi tersebut umumnya masih disajikan dalam bentuk tabel berisi deretan angka sehingga kemiripan kandungan gizi antarmakanan sulit dikenali secara langsung, terutama ketika jumlah makanan dan atribut gizi yang dianalisis cukup banyak. Data kandungan gizi merupakan data multivariat dengan rentang nilai yang beragam, sehingga analisis secara manual menjadi tidak efisien

dan diperlukan pendekatan yang mampu mengelompokkan makanan berdasarkan kemiripan kandungan gizinya.

Salah satu teknik data mining yang sesuai untuk kebutuhan tersebut adalah klusterisasi (*clustering*), yaitu pengelompokan data berdasarkan tingkat kemiripan tanpa memerlukan label awal [1]. Teknik ini telah dimanfaatkan pada berbagai data multivariat, misalnya untuk memetakan pola persebaran penyakit demam berdarah antarwilayah [2] dan mengelompokkan puskesmas berdasarkan kualitas kesehatan keluarga dan gizi masyarakat [3].

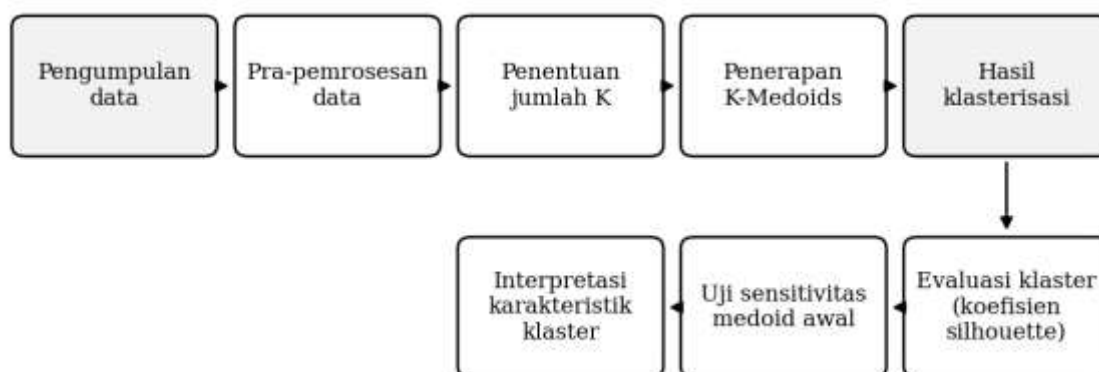
Pada data makanan, klusterisasi telah diterapkan dengan beberapa metode [4] menerapkan K-Means pada 359 data makanan berdasarkan kebutuhan kalori harian dan memperoleh dua klaster dengan koefisien silhouette 0,61 [5] mengelompokkan 1.346 data makanan Indonesia berdasarkan kalori, protein, lemak, dan karbohidrat menggunakan K-Means dan memperoleh tiga klaster [6] menggunakan *agglomerative hierarchical clustering* pada empat variabel gizi utama dan memperoleh dua klaster dengan koefisien silhouette 0,7098, sedangkan [7] menganalisis 48 variabel gizi untuk mengungkap pola tren asupan gizi global. Dua studi pengelompokan makanan yang disebut terakhir hanya melibatkan empat variabel gizi makro, sehingga informasi vitamin dan mineral belum ikut menentukan kemiripan antarmakanan.

Algoritma K-Medoids merupakan metode klusterisasi partisional yang menggunakan objek data sebenarnya (medoid) sebagai pusat klaster sehingga lebih tahan terhadap pencilan dibandingkan metode berbasis rata-rata seperti K-Means [8]. Algoritma ini telah digunakan untuk mengelompokkan data penyakit alergi pada anak [9], status gizi balita [10], produksi perikanan tangkap [11], usaha mikro, kecil, dan menengah bidang kuliner [12], serta tingkat kemiskinan wilayah [13][14], dengan hasil yang dinilai stabil dan representatif [15][16]. Pusat klaster yang berupa objek nyata juga memudahkan interpretasi, karena setiap klaster dapat diwakili oleh makanan yang benar-benar ada pada data. Pada studi yang ditelaah dalam penelitian ini, penerapan K-Medoids pada data makanan dengan atribut gizi yang mencakup vitamin dan mineral belum ditemukan.

Penelitian ini menerapkan algoritma K-Medoids untuk mengelompokkan 1.126 data makanan berdasarkan 21 atribut kandungan gizi yang mencakup energi, zat gizi makro, vitamin, dan mineral. Nilai baru penelitian ini terletak pada (1) penggunaan atribut gizi yang lebih lengkap, (2) penentuan jumlah klaster optimal sekaligus evaluasi hasil menggunakan koefisien silhouette, (3) uji sensitivitas terhadap pemilihan medoid awal untuk menilai konsistensi hasil, dan (4) interpretasi karakteristik klaster berdasarkan rata-rata atribut pada data asli. Tujuan penelitian adalah menerapkan K-Medoids pada data makanan, mengevaluasi kualitas klaster yang terbentuk, dan menginterpretasikan karakteristik kandungan gizi pada setiap klaster.

2. METODE PENELITIAN

Penelitian ini berbasis data sekunder dan dilakukan tanpa pengumpulan data langsung di lapangan. Tahapan penelitian ditunjukkan pada Gambar 1, mulai dari pengumpulan data hingga interpretasi karakteristik klaster. Seluruh proses dijalankan dengan bahasa pemrograman Python menggunakan pustaka *pandas*, *NumPy*, dan *scikit-learn-extra*.



Gambar 1. Tahapan penelitian

2.1. Pengumpulan Data

Data diperoleh dari platform Kaggle berupa dataset publik berjudul “Data Nutrisi Makanan” dalam format CSV [17]. Dataset terdiri atas 1.126 data makanan dengan 23 atribut, yaitu dua atribut identitas (*Unnamed* dan *Menu*) dan 21 atribut kandungan gizi bertipe numerik dengan rentang nilai yang berbeda-beda. Atribut kandungan gizi yang digunakan dikelompokkan pada Tabel 1.

Tabel 1. Atribut kandungan gizi yang digunakan

Kelompok	Atribut	Jumlah
Energi dan zat gizi makro	Energy, Protein, Fat, Carbohydrates, Dietary Fiber, PUFA, Cholesterol	7
Vitamin	Vitamin A, Vitamin E, Vitamin B1, Vitamin B2, Vitamin B6, Total Folic Acid, Vitamin C	7
Mineral	Sodium, Potassium, Calcium, Magnesium, Phosphorus, Iron, Zinc	7

2.2. Pra-pemrosesan Data

Pra-pemrosesan terdiri atas tiga langkah. Pertama, pembersihan data dilakukan dengan memeriksa nilai kosong dan data ganda menggunakan pustaka *pandas*. Kedua, seleksi data dilakukan dengan memilih atribut yang relevan dengan tujuan klasterisasi [18], yaitu 21 atribut kandungan gizi sebagai variabel perhitungan jarak, sedangkan atribut *Unnamed* dan *Menu* hanya dipertahankan sebagai identitas data untuk menampilkan nama makanan pada hasil akhir. Ketiga, normalisasi dilakukan dengan Min-Max Scaling agar seluruh atribut berada pada interval 0 sampai 1 dan tidak ada atribut berskala besar yang mendominasi perhitungan jarak [19], sesuai Persamaan (1).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

Dalam persamaan tersebut, x adalah nilai asli suatu atribut, sedangkan $x_{\min}$ dan $x_{\max}$ masing-masing adalah nilai terkecil dan terbesar atribut tersebut. Min-Max Scaling sensitif terhadap nilai ekstrem. Meskipun demikian, nilai ekstrem pada dataset, misalnya kandungan lemak kemiri yang secara alami jauh lebih tinggi daripada bahan pangan lain, dipandang sebagai karakteristik asli data sehingga tidak dihapus. Keputusan ini sejalan dengan sifat K-Medoids yang lebih tahan terhadap nilai ekstrem dibandingkan algoritma berbasis centroid [8][20].

2.3. Algoritma K-Medoids

K-Medoids, atau *Partitioning Around Medoids* (PAM), membagi data ke dalam k klaster dengan medoid sebagai pusat klaster, yaitu objek data yang paling mewakili kelompoknya dan meminimalkan total ketidaksamaan terhadap anggota klasternya [8]. Langkah-langkahnya mengikuti [21]. (1) Sebanyak k medoid awal dipilih secara acak; pada penelitian ini digunakan *random seed* bernilai 42 agar hasil dapat direproduksi. (2) Setiap data ditempatkan pada klaster dengan medoid terdekat berdasarkan jarak Euclidean pada Persamaan (2). (3) Satu objek non-medoid pada setiap klaster dipilih sebagai kandidat medoid baru. (4) Total simpangan S dihitung sebagai selisih biaya total baru dan biaya total lama; jika $S < 0$, kandidat menggantikan medoid lama. (5) Langkah (2) sampai (4) diulang sampai medoid tidak lagi berubah.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Pada Persamaan (2), $d(x, y)$ adalah jarak antara data x dan y , x_i dan y_i adalah nilai atribut ke- i kedua data, dan n adalah jumlah atribut, yaitu 21.

2.4. Penentuan Jumlah Klaster dan Evaluasi

Jumlah klaster optimal dipilih dari $K = 2$ sampai $K = 7$ berdasarkan koefisien silhouette tertinggi [22], sebagaimana koefisien ini juga digunakan untuk menilai hasil K-Medoids pada pengelompokan data penyakit [23]. Untuk setiap objek i yang berada pada klaster A , nilai $a(i)$ adalah rata-rata jarak objek tersebut ke objek lain dalam klaster yang sama (Persamaan 3), sedangkan $b(i)$

adalah rata-rata jarak terkecil objek tersebut ke seluruh objek pada kluster lain (Persamaan 4). Nilai silhouette objek dihitung dengan Persamaan (5).

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (3)$$

$$b(i) = \min_{C \neq A} \frac{1}{|C|} \sum_{j \in C} d(i, j) \quad (4)$$

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (5)$$

Rata-rata $s(i)$ seluruh objek menjadi koefisien silhouette keseluruhan yang bernilai antara -1 dan 1. Nilai yang mendekati 1 menunjukkan objek berada pada kluster yang tepat dan terpisah jelas dari kluster lain, sedangkan nilai yang mendekati 0 menunjukkan kluster saling tumpang tindih. Interpretasi nilai mengikuti Kaufman dan Rousseeuw [24] pada Tabel 2. Evaluasi internal ini dipilih karena dataset tidak memiliki label kelas (*ground truth*) sehingga validasi eksternal tidak dapat diterapkan.

Tabel 2. Interpretasi koefisien silhouette menurut Kaufman dan Rousseeuw [24]

Nilai silhouette	Interpretasi
0,71 – 1,00	Struktur kuat
0,51 – 0,70	Struktur baik
0,26 – 0,50	Struktur lemah
$\leq 0,25$	Tidak ditemukan struktur

2.5. Uji Sensitivitas dan Interpretasi Kluster

Uji sensitivitas dilakukan dengan menjalankan K-Medoids pada jumlah kluster optimal menggunakan enam nilai *random seed* (1, 16, 28, 42, 900, dan 1102), kemudian membandingkan medoid akhir, biaya total, dan koefisien silhouette yang dihasilkan. Interpretasi karakteristik kluster dilakukan dengan menghitung rata-rata setiap atribut pada data asli (sebelum normalisasi) untuk masing-masing kluster agar nilainya tetap dalam satuan kandungan gizi sebenarnya. Perbandingan antarkluster dianalisis secara deskriptif. Uji signifikansi statistik tidak dilakukan karena penelitian ini tidak bertujuan menguji hipotesis tentang perbedaan rata-rata antarkelompok.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Pra-pemrosesan Data

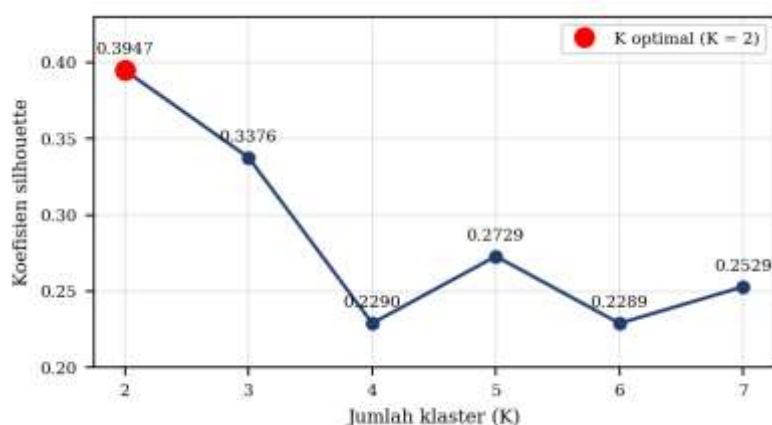
Pemeriksaan menunjukkan tidak ada nilai kosong maupun data ganda, sehingga seluruh 1.126 data digunakan tanpa imputasi atau penghapusan. Sebagian makanan memiliki nilai 0 pada beberapa atribut. Nilai tersebut valid karena menunjukkan makanan memang tidak mengandung zat gizi tertentu, bukan data yang hilang. Setelah normalisasi, seluruh atribut berada pada rentang 0 sampai 1. Sebagai contoh, energi Breastmilk sebesar 288 pada rentang 0–4.506 menjadi 0,0639, dan protein sebesar 1,13 g pada rentang 0–103,3 g menjadi 0,0109.

3.2. Penentuan Jumlah Kluster Optimal

Sebagai ilustrasi proses perhitungan, koefisien silhouette dihitung secara manual terlebih dahulu pada 15 data sampel acak dari keseluruhan data. Hasilnya adalah 0,3380 untuk $K = 2$, 0,2685 untuk $K = 3$, dan 0,3042 untuk $K = 4$, sehingga $K = 2$ terbaik pada sampel. Perhitungan pada seluruh data untuk $K = 2$ sampai 7 disajikan pada Tabel 3 dan Gambar 2.

Tabel 3. Koefisien silhouette pada seluruh data untuk $K = 2$ sampai 7

K	Koefisien silhouette	Kategori (Tabel 2)
2	0,3947	Struktur lemah
3	0,3376	Struktur lemah
4	0,2290	Tidak ditemukan struktur
5	0,2729	Struktur lemah
6	0,2289	Tidak ditemukan struktur
7	0,2529	Struktur lemah

Gambar 2. Koefisien silhouette untuk $K = 2$ sampai 7

Koefisien silhouette tertinggi diperoleh pada $K = 2$ sebesar 0,3947. Nilainya turun pada $K = 3$ (0,3376) dan $K = 4$ (0,2290), naik tipis pada $K = 5$ (0,2729), turun kembali pada $K = 6$ (0,2289), dan naik pada $K = 7$ (0,2529) tanpa melampaui $K = 2$. Dengan demikian, dua kluster dipilih sebagai jumlah kluster optimal. Seluruh nilai berada pada kategori struktur lemah atau lebih rendah menurut Tabel 2, sehingga pemilihan $K = 2$ lebih tepat dibaca sebagai konfigurasi terbaik di antara alternatif yang diuji, bukan sebagai pemisahan yang tegas.

3.3. Hasil Klasterisasi K-Medoids

Penerapan K-Medoids dengan $K = 2$ menghasilkan Klaster 0 sebanyak 807 makanan (71,67%) dan Klaster 1 sebanyak 319 makanan (28,33%). Pada perhitungan manual lima iterasi, biaya total terkecil diperoleh pada iterasi ke-4, yaitu 412,2338. Implementasi dengan Python menghasilkan biaya akhir yang sama dengan medoid akhir pada data berindeks 406 dan 733. Contoh anggota kedua kluster disajikan pada Tabel 4.

Tabel 4. Contoh anggota setiap kluster

Klaster	Jumlah	Contoh anggota
0	807	Nasi putih, ubi jalar merah, kentang, jagung kuning segar, tempe kedele murni, tahu, apel, mangga harum manis, pisang ambon, jeruk manis, sup daging, <i>puree</i> wortel
1	319	Tepung terigu, tepung tapioka, beras putih giling, mie kering, kemiri, kacang tanah tanpa kulit, minyak kelapa, margarin, gula pasir, madu, biskuit marie susu, kerupuk udang, susu dancow, roti tawar

Berdasarkan daftar anggotanya, Klaster 0 sebagian besar berisi makanan segar atau olahan basah, seperti buah-buahan, umbi-umbian, nasi dan bubur, tempe dan tahu, serta sup dan *puree*. Klaster 1 sebagian besar berisi bahan kering atau pekat, seperti tepung, beras giling, biskuit, kerupuk dan makanan ringan, kacang-kacangan, minyak dan lemak, gula, serta susu bubuk. Proporsi anggota yang timpang menunjukkan bahwa sebagian besar makanan dalam dataset memiliki kandungan gizi per takaran yang relatif rendah, sedangkan makanan padat gizi dan energi hanya mencakup sebagian kecil data.

3.4. Evaluasi Klaster

Koefisien silhouette hasil klasterisasi akhir sebesar 0,3947. Menurut Tabel 2, nilai ini termasuk kategori struktur lemah, artinya pemisahan antarklaster belum terbentuk secara kuat meskipun objek dalam satu klaster masih relatif mirip. Kondisi ini kemungkinan dipengaruhi oleh keragaman data, karena dataset memuat banyak jenis makanan dengan kandungan energi, zat gizi makro, vitamin, dan mineral yang sangat beragam sehingga batas antarklaster tidak selalu tegas. Banyaknya nilai 0 yang valid pada sejumlah atribut dan kemiripan pola gizi pada sebagian makanan juga dapat mengurangi keterpisahan klaster.

Nilai tersebut lebih rendah daripada koefisien silhouette pada pengelompokan makanan oleh Nugroho dkk. [4] sebesar 0,61 dan Hanifah dan Herlina [6] sebesar 0,7098. Namun, studi Hanifah dan Herlina hanya memakai empat variabel gizi, sedangkan penelitian ini memakai 21 atribut sehingga jarak antarmakanan dihitung pada ruang yang lebih kompleks. Nilai penelitian ini berada dekat dengan hasil K-Medoids pada pengelompokan tingkat kemiskinan provinsi sebesar 0,4 [14] dan lebih tinggi daripada hasil pada tingkat kemiskinan kabupaten/kota sebesar 0,224 [13]. Perbandingan ini bersifat indikatif karena dataset, atribut, dan jumlah data setiap studi berbeda. Hasil pengelompokan sebaiknya dibaca sebagai kecenderungan umum, bukan batas yang tegas antarkelompok makanan.

3.5. Uji Sensitivitas Pemilihan Medoid Awal

Enam nilai *random seed*, yaitu 1, 16, 28, 42, 900, dan 1102, menghasilkan medoid akhir yang sama (data berindeks 406 dan 733), biaya total yang sama sebesar 412,2338, dan koefisien silhouette yang sama sebesar 0,3947. Biaya tersebut juga sama dengan biaya pada iterasi ke-4 perhitungan manual. Perubahan medoid awal pada pengujian ini tidak memengaruhi hasil akhir, sehingga hasil klasterisasi pada dataset ini tergolong konsisten dan stabil terhadap variasi inisialisasi yang diuji.

3.6. Interpretasi Karakteristik Klaster

Rata-rata setiap atribut kandungan gizi pada data asli untuk kedua klaster disajikan pada Tabel 5.

Tabel 5. Rata-rata atribut kandungan gizi pada setiap klaster (data asli)

Atribut	Klaster 0	Klaster 1
Energy	491,68	1.775,30
Protein (g)	8,29	9,13
Fat (g)	4,40	17,98
Carbohydrates (g)	12,03	60,03
Dietary Fiber	1,23	4,19
PUFA	0,74	5,85
Cholesterol	55,95	10,55
Vitamin A	281,94	104,76
Vitamin E	0,57	1,92
Vitamin B1	0,07	0,15
Vitamin B2	0,13	0,18
Vitamin B6	0,17	0,22
Total Folic Acid	30,29	44,76
Vitamin C	8,32	10,76
Sodium	235,70	92,76
Potassium	242,37	410,51
Calcium	51,74	166,19
Magnesium	26,51	59,94
Phosphorus	106,54	199,38
Iron	1,56	3,99
Zinc	0,82	1,40

Satuan atribut mengikuti dataset sumber. Cetak tebal: nilai rata-rata tertinggi di antara kedua klaster.

Klaster 1 memiliki rata-rata lebih tinggi pada 18 dari 21 atribut. Rata-rata energi pada Klaster 1 sebesar 1.775,30 dibandingkan 491,68 pada Klaster 0, lemak 17,98 g dibandingkan 4,40 g, karbohidrat 60,03 g dibandingkan 12,03 g, serat pangan 4,19 g dibandingkan 1,23 g, dan *polyunsaturated fatty acid* (PUFA) 5,85 g dibandingkan 0,74 g. Klaster 1 juga lebih tinggi pada vitamin E, B1, B2, B6, total asam folat, dan vitamin C, serta pada mineral kalium, kalsium, magnesium, fosfor, zat besi, dan seng. Rata-rata protein kedua klaster relatif sebanding, yaitu 9,13 g pada Klaster 1 dan 8,29 g pada Klaster 0.

Klaster 0 memiliki rata-rata lebih tinggi pada tiga atribut, yaitu kolesterol (55,95 dibandingkan 10,55), vitamin A (281,94 dibandingkan 104,76), dan natrium (235,70 dibandingkan 92,76). Dengan demikian, Klaster 1 dapat dicirikan sebagai kelompok makanan padat energi dengan kandungan zat gizi makro, serat, sebagian besar vitamin, dan sebagian besar mineral yang lebih tinggi, sedangkan Klaster 0 dicirikan sebagai kelompok makanan dengan energi dan zat gizi makro yang lebih rendah tetapi kolesterol, vitamin A, dan natrium yang lebih tinggi. Perlu dicatat bahwa rata-rata yang lebih tinggi tidak berarti makanan tersebut lebih baik; energi dan lemak yang tinggi pada Klaster 1, misalnya, mencerminkan bahan pekat seperti minyak, tepung, dan biskuit.

3.7. Implikasi dan Keterbatasan

Hasil pengelompokan dapat membantu pembaca mengenali dan membandingkan makanan berdasarkan kandungan gizinya, misalnya energi, protein, lemak, karbohidrat, serat, dan natrium. Namun, hasil ini tidak digunakan untuk menentukan pola makan atau diet tertentu, terutama bagi individu dengan kondisi kesehatan khusus; hasil hanya berfungsi sebagai informasi pendukung dan tetap perlu disesuaikan dengan kebutuhan individu serta anjuran tenaga kesehatan. Keterbatasan penelitian ini adalah struktur klaster yang lemah, penggunaan satu dataset dari satu sumber, penggunaan satu metrik validasi internal, sensitivitas Min-Max Scaling terhadap nilai ekstrem, dan analisis karakteristik yang bersifat deskriptif tanpa uji signifikansi statistik.

4. KESIMPULAN

Penelitian ini menerapkan algoritma K-Medoids untuk mengelompokkan 1.126 data makanan berdasarkan 21 atribut kandungan gizi. Jumlah klaster optimal berdasarkan koefisien silhouette adalah dua, dengan Klaster 0 beranggotakan 807 makanan dan Klaster 1 beranggotakan 319 makanan. Koefisien silhouette sebesar 0,3947 menunjukkan struktur klaster yang lemah menurut Kaufman dan Rousseeuw, tetapi pengelompokan masih menghasilkan dua kelompok dengan karakteristik gizi yang berbeda. Klaster 1 memiliki rata-rata lebih tinggi pada energi, lemak, karbohidrat, serat, dan sebagian besar vitamin serta mineral, sedangkan Klaster 0 memiliki rata-rata lebih tinggi pada kolesterol, vitamin A, dan natrium. Hasil yang sama pada enam *random seed* menunjukkan bahwa pengelompokan konsisten terhadap pemilihan medoid awal. Temuan ini sesuai dengan tujuan penelitian yang dinyatakan pada bagian Pendahuluan.

Penelitian selanjutnya disarankan menggunakan dataset yang lebih beragam dan lebih besar, membandingkan K-Medoids dengan algoritma lain seperti K-Means, *hierarchical clustering*, DBSCAN, atau *Fuzzy C-Means*, menganalisis pencilaan serta membandingkan metode normalisasi lain seperti *Z-Score Standardization* dan *Robust Scaling*, memanfaatkan visualisasi atau analisis statistik untuk karakterisasi klaster yang lebih mendalam, dan melibatkan ahli gizi untuk memvalidasi interpretasi klaster.

REFERENSI

- [1] D. F. Pramesti, M. T. Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan / Lahan Berdasarkan Persebaran Titik Panas (Hotspot)," vol. 1, no. 9, pp. 723–732, 2017.
- [2] A. N. Thamrin, "Penerapan Algoritma Clustering K-Medoids pada Penyebaran Penyakit Demam Berdarah di Kota Palopo," vol. 4, no. 3, pp. 1225–1232, 2024.
- [3] B. H. Prakoso, E. Rachmawati, D. Rachmatta, and P. Mudiono, "Klasterisasi Puskesmas dengan K-Means Berdasarkan Data Kualitas Kesehatan Keluarga dan Gizi Masyarakat," vol. 14, no. April, pp. 60–68, 2023.

- [4] T. Nugroho, U. Enri, I. Maulana, P. S. Informatika, U. S. Karawang, and T. Timur, "CLUSTERING MENU MAKANAN BERDASARKAN KEBUTUHAN KALORI HARIAN MENGGUNAKAN ALGORITME K-MEANS," vol. 8, no. 4, pp. 4411–4417, 2024.
- [5] P. A. Vredizon, H. Firmansyah, N. S. Salsabila, and W. E. Nugroho, "Penerapan Algoritma K- Means untuk Mengelompokkan Makanan Berdasarkan Nilai Nutrisi," vol. 5, no. 2, 2024, doi: 10.37802/joti.v5i2.577.
- [6] F. S. Hanifah and M. Herlina, "Pengelompokan Makanan Berdasarkan Komposisi Zat Gizi Menggunakan Metode Agglomerative Hierarchical Clustering".
- [7] A. Matousek, T. H. Leung, and H. Pang, "Multivariate time series clustering analysis of the Global Dietary Database to uncover patterns in dietary trends (1990 – 2018)," no. 8, 2025, doi: 10.1017/S136898002500059X.
- [8] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R.H.Zer, and D. Hartama, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 166–173, 2020, doi: 10.36294/jurti.v4i1.1296.
- [9] H. Ningrum, E. Irawan, and M. R. Lubis, "Implementasi Metode K-Medoids Clustering Dalam Pengelompokan Data Penyakit Alergi Pada Anak," vol. 6, pp. 130–139, 2021.
- [10] Susliansyah, H. Sumarno, H. Priyono, and L. Maulida, "Penerapan Klastering pada Data Mining dalam Menentukan Status Gizi Anak Balita dengan," vol. 10, no. 1, pp. 64–72, 2026.
- [11] Fajriana, "Analisis Algoritma K-Medoids pada Sistem Klasterisasi Produksi Perikanan Tangkap," vol. 7, no. 2, pp. 263–269, 2021.
- [12] U. Linarti, A. Rahmawati, A. Hendri, S. Jones, and L. Zahrotun, "Penerapan Metode K-Medoids Guna Pengelompokan Data Usaha Mikro , Kecil dan Menengah (UMKM) Bidang Kuliner Di Kota Yogyakarta," vol. 7, pp. 37–45, 2024.
- [13] G. Risky Pratiwi, D. Wahiddin, E. Elvira Awal, and A. Fauzi, "Klasterisasi Tingkat Kemiskinan Kabupaten / Kota di Indonesia," pp. 197–208, 2024, doi: 10.33364/algoritma/v.21-2.1788.
- [14] F. Zahra, A. Khalif, B. N. Sari, U. S. Karawang, T. Timur, and J. Barat, "PROVINSI INDONESIA MENGGUNAKAN ALGORITMA K-MEDOIDS," vol. 12, no. 2, 2024.
- [15] E. I. Sari, A. Syahrani, and P. Mandarani, "KLAUSTERISASI PENYAKIT MENGGUNAKAN ALGORITMA K-MEDOIDS PADA DINAS KESEHATAN KABUPATEN AGAM Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI | 138," vol. 10, pp. 137–146, 2021.
- [16] P. N. Safitri, R. Aristawidya, and S. B. Faradilla, "KLAUSTERISASI FAKTOR-FAKTOR KEMISKINAN DI PROVINSI JAWA BARAT MENGGUNAKAN K-MEDOIDS CLUSTERING," vol. 4, no. 2, pp. 75–80, 2021.
- [17] nauvalalmas, "Data makanan stunting," *Kaggle*. [Online]. Available: <https://www.kaggle.com/datasets/nauvalalmas/data-makanan-stunting>
- [18] Y. A. Rohim, T. Z. Fatikhah, R. Q. Hardiansyah, and D. Hartanti, "Clustering Fasilitas Kesehatan Berdasarkan Kecamatan di Jakarta dengan Algoritma K-Medoids," pp. 737–742, 2023.
- [19] D. Kurmiati, M. Z. Fauzi, and A. Falegas, "Clustering of Earthquake Prone Areas in Indonesia Using K-Medoids Algorithm Klasterisasi Daerah Rawan Gempa Bumi di Indonesia," vol. 1, no. April, pp. 47–57, 2021.
- [20] A. O. Hermadi, W. Priatna, and A. D. Alexander, "Implementasi Algoritma K-Medoids Clustering Untuk Mencari Keuntungan Sementara Dalam Laporan Keuangan," vol. 6, no. April, pp. 6–11, 2023.
- [21] H. Sariangsah and R. Rosnelly, "Penentuan Kelas Menggunakan Algoritma K Medoids Untuk Clustering Siswa Tunagrahita," vol. 5, pp. 83–89, 2021, doi: 10.30865/mib.v5i1.2547.

-
- [22] Y. Kustiyahningsih, A. Khozaimi, and J. Purnama, “Pengelompokan UMKM Batik Madura Menggunakan Metode Grouping of Madurese Batik MSMEs Using the K-Means and Silhouette Coefficient Method,” vol. 13, no. 2, pp. 192–198, 2024, doi: 10.34148/teknika.v13i2.779.
 - [23] S. Nurlaela, A. Primajaya, and T. Nur Padilah, “ALGORITMA K-MEDOIDS UNTUK CLUSTERING PENYAKIT,” vol. 12, no. 2, pp. 56–62, 2020.
 - [24] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. New York, NY, USA: Wiley, 1990.