

ANALISIS PERBANDINGAN METODE K-MEDOID DAN AGGLOMERATIVE HIERARCHICAL CLUSTERING PADA DATA KONSUMSI REMPAH-REMPAH DI KABUPATEN / KOTA

Lina Rhomaningtias¹, M. Nurhadyatullah Kusharyadi², Reagen Westerdam Sean Jatindra³, Trimono⁴, Muhammad Nasrudin⁵

^{1,2,3,4} Universitas Pembangunan Nasional 'Veteran' Jawa Timur; Jl. Rungkut Madya, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur 60294; Telp. (031) 8706369

Keywords:

Clustering, Spice Consumption, K-Medoid, Agglomerative Hierarchical Clustering, PCA.

Correspondent Email:

23083010077@student.upnja
tim.ac.id

Abstrak. Konsumsi rempah-rempah di Indonesia mencerminkan keragaman budaya, geografi, dan pola hidup masyarakat di berbagai daerah. Namun, kajian kuantitatif yang memetakan pola konsumsi rempah antar kabupaten/kota masih terbatas, meskipun data statistik tersedia secara terbuka. Penelitian ini bertujuan untuk mengeksplorasi dan membandingkan efektivitas dua metode klasterisasi K-Medoid dan Agglomerative Hierarchical Clustering (AHC) dalam mengelompokkan wilayah berdasarkan kesamaan konsumsi enam jenis rempah utama: bawang merah, bawang putih, bawang bombay, cabai merah, cabai hijau, dan cabai rawit. Data sekunder berasal dari Badan Pusat Statistik tahun 2024, dengan preprocessing berupa pembersihan data, standarisasi, serta reduksi dimensi menggunakan Principal Component Analysis (PCA). Evaluasi dilakukan menggunakan metrik validasi internal seperti Silhouette Score, Dunn Index, Davies-Bouldin Index, Calinski-Harabasz Index, dan Cophenetic Correlation. Hasil menunjukkan bahwa metode AHC dengan linkage ward dan lima klaster memberikan performa paling optimal dibandingkan K-Medoid. Segmentasi wilayah berdasarkan hasil klaster mengungkapkan struktur konsumsi rempah yang berbeda antara wilayah pedesaan dan perkotaan. Penelitian ini memberikan kontribusi penting dalam pemetaan konsumsi rempah berbasis data dan dapat dijadikan dasar perumusan kebijakan pangan dan pembangunan wilayah yang lebih tepat sasaran.



JITET is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Abstract. Spice consumption in Indonesia reflects the cultural, geographical, and lifestyle diversity of communities in various regions. However, quantitative studies mapping spice consumption patterns between districts/cities are still limited, even though statistical data is publicly available. This study aims to explore and compare the effectiveness of two clustering methods K-Medoid and Agglomerative Hierarchical Clustering (AHC) in grouping regions based on the similarity of consumption of six main types of spices: red onion, garlic, shallot, red chili, green chili, and bird's eye chili. Secondary data were obtained from the Central Statistics Agency in 2024, with preprocessing in the form of data cleaning, standardization, and dimension reduction using Principal Component Analysis (PCA). Evaluation was carried out using internal validation metrics such as Silhouette Score, Dunn Index, Davies-Bouldin Index, Calinski-Harabasz Index, and Cophenetic Correlation. The results show that the AHC method with Ward linkage and five clusters provides the most optimal performance compared to K-Medoid.

Regional segmentation based on cluster results reveals different spice consumption structures between rural and urban areas. This study makes an important contribution to data-based spice consumption mapping and can serve as a basis for formulating more targeted food and regional development policies.

1. PENDAHULUAN

Rempah-rempah merupakan salah satu komoditas penting dalam kehidupan masyarakat Indonesia. Bagian tanaman seperti buah, daun, akar, kulit, batang, dan biji sering dimanfaatkan sebagai bumbu masak, penyedap rasa, pemberi aroma, pengawet, serta pewarna alami [1]. Selain memberikan cita rasa khas dalam masakan, rempah-rempah juga memiliki berbagai khasiat kesehatan, seperti antioksidan, antibakteri, dan anti-inflamasi [2]. Dalam praktik sehari-hari, penggunaannya telah menjadi bagian yang tidak terpisahkan dari budaya kuliner dan pengobatan tradisional di berbagai daerah di Indonesia [3].

Indonesia sendiri dikenal sebagai salah satu produsen rempah terbesar di dunia dengan keanekaragaman jenis rempah yang tersebar dari Sabang hingga Merauke [4]. Setiap daerah memiliki kekhasan dalam konsumsi rempah, yang dipengaruhi oleh faktor geografis, sosial, budaya, dan ekonomi [5]. Masyarakat perkotaan, misalnya, cenderung mengikuti pola konsumsi modern yang dipengaruhi oleh tren global dan informasi kesehatan, sedangkan masyarakat pedesaan lebih mempertahankan pola konsumsi tradisional yang berbasis ketersediaan lokal dan warisan budaya [6].

Meskipun data mengenai konsumsi rempah telah tersedia melalui lembaga resmi seperti Badan Pusat Statistik (BPS), penelitian yang secara spesifik menganalisis perbedaan pola konsumsi rempah-rempah antarwilayah di Indonesia masih terbatas. Sebagian besar studi sebelumnya lebih berfokus pada aspek agronomi, perdagangan, atau gizi rempah-rempah tanpa melakukan eksplorasi spasial terhadap pola konsumsinya. Di sisi lain, metode analisis data seperti klusterisasi (*clustering*) memiliki potensi besar dalam mengelompokkan wilayah berdasarkan kemiripan konsumsi, namun belum banyak dimanfaatkan dalam konteks ini. Metode kluster seperti K-Medoid dan Agglomerative Hierarchical Clustering merupakan dua pendekatan yang umum digunakan dalam segmentasi wilayah berbasis

data tanpa label, dan memiliki keunggulan masing-masing dalam menangkap struktur dan pola dalam data multivariat. Namun, belum ada penelitian yang secara langsung membandingkan efektivitas kedua metode tersebut dalam konteks konsumsi rempah-rempah di Indonesia berdasarkan data konsumsi per kapita di tingkat kabupaten/kota.

Oleh karena itu, penelitian ini memiliki kebaruan dalam hal pendekatan, yaitu dengan membandingkan dua metode klusterisasi yakni K-Medoid dan Agglomerative Hierarchical Clustering, untuk menganalisis pola konsumsi enam jenis rempah utama, yaitu bawang merah, bawang putih, bawang bombay, cabai merah, cabai hijau, dan cabai rawit, berdasarkan data BPS tahun 2024. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Indonesia berdasarkan kesamaan pola konsumsi rempah-rempah, membandingkan hasil klusterisasi yang dihasilkan oleh kedua metode tersebut, serta mengevaluasi metode mana yang paling efektif dalam merepresentasikan struktur konsumsi rempah di masing-masing wilayah, menggunakan matrik evaluasi seperti Silhouette Score, Dunn Index, hingga Cophenetic Correlation. Selain itu, Principal Component Analysis (PCA) digunakan untuk mereduksi dimensi data agar visualisasi dan analisis menjadi lebih efisien. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pemetaan pola konsumsi pangan nasional serta menjadi dasar bagi kebijakan pangan, kesehatan, dan pengembangan wilayah berbasis data.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Widodo et al. melakukan pengelompokan tingkat kemiskinan di seluruh provinsi Indonesia menggunakan metode Agglomerative Hierarchical Clustering (AHC). Mereka menggunakan tujuh indikator sosial seperti indeks keparahan dan kedalaman kemiskinan, rata-rata lama sekolah, tingkat

pengangguran terbuka, dan persentase penduduk miskin. Penelitian tersebut menunjukkan bahwa metode AHC mampu mengelompokkan provinsi menjadi tiga klaster utama (rendah, sedang, dan tinggi) dan menemukan bahwa metode *ward linkage* menghasilkan nilai koefisien agglomerative tertinggi sebesar 0.985, yang menunjukkan struktur klaster yang kuat [7].

Sementara itu, Zahra et al. menggunakan algoritma K-Medoids untuk mengelompokkan tingkat kemiskinan di 34 provinsi Indonesia berdasarkan data BPS 2015–2022. Penelitian tersebut membentuk tiga klaster sangat miskin, miskin, dan rentan miskin dengan evaluasi menggunakan *Silhouette Score* sebesar 0,4 yang menunjukkan kualitas klaster cukup baik. Metode K-Medoids dipilih karena ketahanannya terhadap outlier dan kestabilan hasil meskipun data diacak [8].

Kedua penelitian tersebut menekankan pentingnya metode klasterisasi dalam menganalisis data sosial berskala wilayah, namun belum ada studi yang membandingkan langsung metode AHC dan K-Medoids dalam konteks konsumsi rempah-rempah di tingkat kabupaten/kota. Oleh karena itu, penelitian ini mengisi celah tersebut dengan membandingkan kedua metode tersebut menggunakan berbagai matrik evaluasi seperti *Silhouette Score*, *Davies-Bouldin Index*, *Calinski-Harabasz Index*, *Dunn Index*, *Cophenetic Correlation*, *Average Connectivity*, dan *Cluster Connectivity*, serta memanfaatkan *Principal Component Analysis (PCA)* sebagai tahap reduksi dimensi.

2.2. Rempah-Rempah

Rempah-rempah adalah bagian tanaman aromatik seperti biji, kulit, dan akar yang digunakan terutama sebagai bumbu atau obat [9].

2.3. Clustering

Klasterisasi (*clustering*) adalah teknik unsupervised learning yang bertujuan untuk mengelompokkan data ke dalam grup yang homogen di dalam dan heterogen antar grup. Metode ini digunakan secara luas dalam eksplorasi data, terutama ketika data tidak memiliki label kelas. Dalam konteks geospasial atau sosial, *clustering* digunakan untuk

menemukan pola tersembunyi dalam data wilayah atau populasi [10].

2.4. K-Medoid

K-Medoid adalah metode klasterisasi partisi yang mirip dengan K-Means, namun lebih tahan terhadap outlier karena menggunakan medoid sebagai pusat klaster, bukan mean. Medoid adalah titik data aktual yang meminimalkan total jarak ke titik lain dalam klasternya. K-Medoid bekerja dengan iterasi pemilihan medoid awal, pembentukan klaster, dan perhitungan ulang medoid hingga mencapai hasil optimal [11].

2.5. Agglomerative Hierarchical Clustering

Agglomerative Hierarchical Clustering (AHC) merupakan metode klasterisasi hierarkis yang dimulai dari tiap data sebagai satu klaster terpisah, kemudian menggabungkannya secara bertahap berdasarkan kedekatan atau kesamaan, hingga menjadi satu klaster besar [12]. Proses penggabungan dilakukan menggunakan linkage method, seperti *single*, *complete*, *average*, atau *ward linkage* [13].

3. METODE PENELITIAN

3.1. Jenis dan Pendekatan Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan eksploratif komparatif. Penelitian dilakukan untuk mengeksplorasi pola konsumsi rempah-rempah antar wilayah kabupaten/kota di Indonesia dan membandingkan hasil klasterisasi yang diperoleh dari dua metode berbeda, yaitu K-Medoid dan Agglomerative Hierarchical Clustering (AHC). Pendekatan ini memungkinkan identifikasi segmentasi wilayah berdasarkan kesamaan karakteristik konsumsi rempah yang bersifat multivariat.

3.2. Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) tahun 2024, khususnya data rata-rata konsumsi per kapita per minggu (dalam gram) untuk enam jenis rempah: bawang merah, bawang putih, bawang bombay, cabai merah, cabai hijau, dan cabai rawit. Unit analisis pada penelitian ini adalah kabupaten/kota di Indonesia.

3.3. Variabel Penelitian

Variabel-variabel dalam penelitian ini berupa enam jenis rempah-rempah yang masing-masing variabel diukur dalam satuan gram per kapita per minggu, yaitu:

Tabel 1. Daftar Variabel Penelitian

Kode Variabel	Nama Variabel
X_1	Konsumsi bawang merah
X_2	Konsumsi bawang putih
X_3	Konsumsi bawang bombay
X_4	Konsumsi cabai merah
X_5	Konsumsi cabai hijau
X_6	Konsumsi cabai rawit

3.4. Pre-Processing Data

3.4.1. Pembersihan Data

Data yang digunakan terlebih dahulu diperiksa untuk mengatasi duplikasi, nilai kosong, atau inkonsistensi. Pembersihan ini bertujuan meningkatkan kualitas data agar tidak mengganggu proses klasterisasi. Jika terdapat nilai kosong, dilakukan imputasi atau penghapusan baris sesuai konteks data [14].

3.4.2. Transformasi Data Kategorikal

Beberapa atribut dalam dataset memiliki tipe kategorikal yang tidak dapat langsung diproses oleh algoritma berbasis jarak. Oleh karena itu, dilakukan transformasi menggunakan Label Encoding, yaitu metode yang mengonversi setiap label kategori menjadi representasi bilangan bulat unik. Langkah ini memungkinkan semua fitur diproses secara matematis [15].

3.4.3. Standarisasi Data

Standarisasi data adalah transformasi yang menghasilkan data dengan distribusi standar, yaitu rata-rata nol dan simpangan baku satu. Dengan demikian, setiap fitur akan memiliki kontribusi yang seimbang dalam perhitungan jarak antar data, sehingga mencegah dominasi fitur tertentu dalam proses klasterisasi [16]. Berikut rumus dari standarisasi data:

$$z = \frac{x - \mu}{\sigma}$$

dengan:

x = nilai asli dari suatu fitur.

μ = rata-rata fitur tersebut.

σ = simpangan bakunya.

3.4.4. Reduksi Dimensi dengan PCA

Jika dataset memiliki dimensi yang tinggi, maka dilakukan reduksi dimensi menggunakan Principal Component Analysis (PCA). Melalui pendekatan ini, PCA dapat mereduksi kompleksitas data tanpa menghilangkan struktur informasi yang penting. Hal ini tidak hanya meningkatkan efisiensi komputasi, tetapi juga mempermudah proses visualisasi hasil klasterisasi secara lebih terstruktur dan informatif [17].

Secara matematis, transformasi PCA dinyatakan sebagai:

$$Z = XW$$

dengan:

X = matriks data yang telah dinormalisasi.

W = matriks eigenvector yang diperoleh dari matriks kovarians X .

Z = hasil proyeksi data ke ruang berdimensi lebih rendah.

3.5. Metode Clustering

3.5.1 Agglomerative Clustering

Pada Agglomerative Clustering itu sendiri, terdapat berbagai metode pengelompokan. metode tersebut terdiri sebagai berikut:

a. Single Linkage

Metode Single Linkage adalah salah satu metode dalam klasterisasi hierarkis, di mana jarak antara dua klaster didefinisikan sebagai jarak terpendek antara sepasang titik data dari masing-masing klaster [18].

Secara matematis, single linkage dinyatakan sebagai:

$$d(A, B) = \min_{a \in A, b \in B} [d(a, b)]$$

dengan:

$d(A, B)$ = jarak rata-rata antara kedua klaster.

a dan b = elemen klaster A dan B .

$d(a, b)$ = menunjukkan jarak antara dua elemen data a dan b .

$\min_{a \in A, b \in B}$ = nilai minimum dari seluruh jarak antara titik a dan b yang berasal dari klaster A dan B .

b. Complete Linkage

Metode Complete Linkage adalah salah satu teknik dalam klasterisasi hierarkis yang menghitung jarak antar klaster berdasarkan jarak maksimum antara sepasang titik data dari masing-

masing kluster. Dua kluster akan digabung jika jarak maksimum antar anggotanya adalah yang paling kecil dibandingkan pasangan kluster lainnya [18].

Secara matematis, complete linkage dinyatakan sebagai:

$$d(A, B) = \max_{a \in A, b \in B} [d(a, b)]$$

dengan:

$\max_{a \in A, b \in B}$ = nilai maksimum dari seluruh jarak antara titik a dan b yang berasal dari kluster A dan B .

c. Average Linkage

Metode Average Linkage adalah salah satu metode dalam klusterisasi hierarkis, di mana jarak antara dua kluster dihitung sebagai rata-rata jarak antara semua pasangan titik data, dengan satu titik berasal dari masing-masing kluster [18].

Secara matematis, average linkage dinyatakan sebagai:

$$d(A, B) = \frac{T_{ab}}{(N_a N_b)}$$

dengan:

T_{ab} = jumlah seluruh jarak antara pasangan titik data dari kluster A dan B .

N_a = jumlah objek dalam kluster A .

N_b = jumlah objek dalam kluster B .

d. Ward

Metode Ward adalah metode klusterisasi hierarkis yang menggabungkan dua kluster dengan peningkatan terkecil pada nilai *Sum of Squared Error* (SSE). Tujuannya adalah meminimalkan keragaman dalam kluster, sehingga menghasilkan kluster yang kompak dan homogen [18].

Secara matematis, SSE dihitung dengan rumus:

$$SSE = \sum_{i=1}^n (x_j - \underline{x})(x_j - \underline{x})$$

dengan:

x_j = data objek ke- j .

$\underline{x}$ = rata-rata objek dalam kluster.

n = jumlah objek dalam kluster.

3.5.2 K - Medoids

Tahapan pembentukan kluster menggunakan algoritma K-Medoids adalah sebagai berikut:

1. Menentukan jumlah kluster yang diinginkan (k) sebagai parameter awal.
2. Memilih k objek secara acak dari data sebagai medoid awal.
3. Mengelompokkan setiap data ke medoid terdekat berdasarkan jarak Euclidean.
4. Memilih secara acak objek dalam kluster sebagai kandidat medoid pengganti.
5. Menghitung ulang jarak total seluruh anggota kluster terhadap kandidat medoid.
6. Menentukan perubahan biaya (S) dari selisih total jarak baru dan lama; jika $S < 0$, lakukan pertukaran medoid.
7. Mengulangi proses evaluasi dan pertukaran hingga medoid tidak berubah, yang menandakan proses klusterisasi selesai [19].

3.6. Evaluasi Hasil

Evaluasi performa klusterisasi dilakukan menggunakan berbagai metrik validasi internal untuk mengukur kualitas hasil pengelompokan.

3.6.1. Silhouette Score

Silhouette Score digunakan untuk menilai seberapa baik setiap objek terkelompok dalam kluster yang tepat. Nilai skor mendekati +1 menunjukkan bahwa objek berada dalam kluster yang sangat sesuai, sedangkan nilai mendekati -1 menunjukkan kemungkinan salah kluster [20].

3.6.2. Dunn Index

Dunn Index mengukur rasio antara jarak minimum antar kluster dan diameter maksimum dalam kluster. Semakin tinggi nilai indeks, semakin baik kualitas pemisahan antar kluster [21].

3.6.3. Davies Bouldin Index

Indeks ini mengukur rasio antara jarak antar pusat kluster dan sebaran internal kluster. Nilai yang lebih kecil menunjukkan kluster yang lebih kompak dan terpisah dengan baik dari yang lainnya [21].

3.6.4. Calinski Harabasz Index

Calinski Harabasz Index mengukur rasio antara dispersi antar kluster dan dalam kluster. Semakin tinggi nilainya, semakin baik struktur kluster yang terbentuk [22].

3.6.5. Xie Beni Index

Xie Beni Index menggabungkan ukuran kepadatan internal kluster dengan jarak antar pusat kluster. Nilai yang lebih rendah menunjukkan hasil klusterisasi yang optimal dalam hal pemisahan dan kepadatan [23].

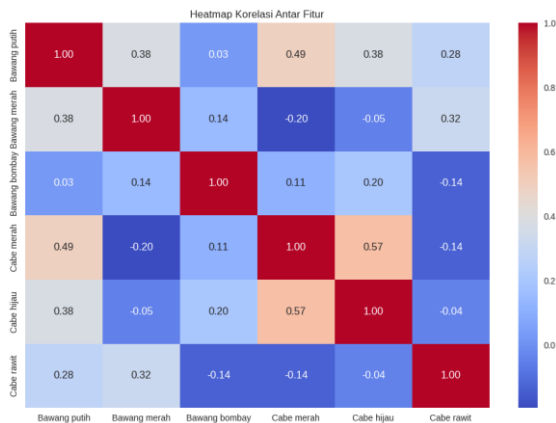
3.6.6. Connectivity Coefficient

Cluster Connectivity digunakan untuk menilai sejauh mana data dalam kluster terhubung dengan baik terhadap tetangga terdekatnya. Semakin rendah nilai konektivitas, semakin baik kualitas klusterisasi karena menandakan data tidak tersebar ke kluster lain [24].

4. HASIL DAN PEMBAHASAN

4.1 Pengecekan Asumsi

Dalam analisis *cluster*, terdapat beberapa asumsi yang harus terpenuhi terlebih dahulu guna menentukan metode yang tepat dan mencegah kesalahan serta menjamin validitas hasil dari metode. Asumsi pertama yang akan kita uji adalah asumsi non-multikolinearitas.



Gambar 1. Hasil Analisis Korelasi

Berdasarkan pengujian korelasi pada Gambar 1, diketahui bahwa tidak ada satupun pasangan yang menunjukkan nilai di atas 0.85, dimana nilai korelasi tertinggi itu sendiri terdapat pada cabe merah dengan cabe hijau yaitu pada nilai 0.57. Berdasarkan hasil tersebut, tidak diperlukan penanganan khusus pada data.

Kemudian kita akan melakukan analisis pada PCA untuk menilai jumlah komponen berapa yang cocok pada data tersebut. Tahapan pertama yaitu pembentukan komponen utama PCA yang terdiri dari kombinasi linear variabel dalam data. Penentuan jumlah komponen utama

didasarkan pada *eigenvalue*, dan nilai proporsi kumulatif varians.

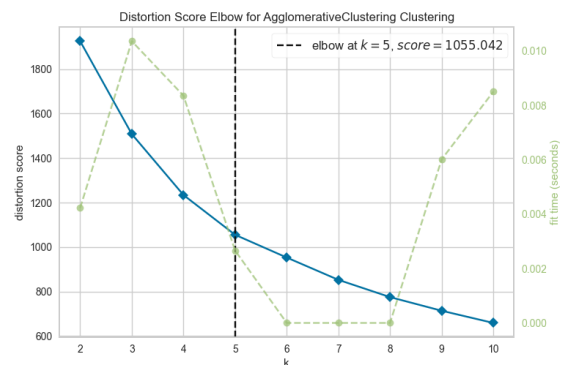
Komponen Utama	Eigen Value	Variance	Cumulative Variance
1	2.02	33.57	33.57
2	1.59	26.42	59.99
3	1.07	17.85	77.84
4	0.63	10.45	88.28
5	0.46	7.70	95.98
6	0.24	4.02	100.00

Tabel 2. Hasil Analisis komponen PCA

Berdasarkan Tabel 2, terlihat bahwa terdapat tiga komponen yang memiliki nilai *eigenvalue* diatas satu. Namun, jika dilihat dari nilai proporsi kumulatif varians ketiga komponen tersebut tidak ada satupun komponen dengan varians kumulatif yang mencapai 80%, sehingga perlu ditambahkan beberapa komponen lagi sehingga syarat terpenuhi. Varians kumulatif komponen tersebut telah mencapai 80% ketika jumlah komponen utama ada di empat yaitu pada nilai 88.28%. Dengan demikian penelitian ini menggunakan lima komponen utama untuk analisis selanjutnya

4.2 Agglomerative Hierarchical Clustering

Penggabungan kluster dalam metode ini bergantung pada jenis linkage yang digunakan, yaitu *single*, *complete*, *average*, dan *ward*. Langkah pertama yang akan dilakukan adalah menentukan jumlah kluster yang paling optimal untuk model Agglomerative Hierarchical Clustering. Dengan menggunakan metode *elbow*, kita dapat mengetahui jika jumlah kluster yang paling optimal terletak pada lima kluster.



Gambar 2. Hasil Metode Elbow Agglomerative Clustering

Setelah kita mendapatkan jumlah kluster yang optimal, maka langkah selanjutnya adalah menentukan jenis linkage yang akan digunakan untuk mendapatkan hasil yang optimal. Hasil penentuan linkage tersebut dapat dilihat dari tabel Tiga di bawah berikut

Tabel 3. Hasil Evaluasi Agglomerative

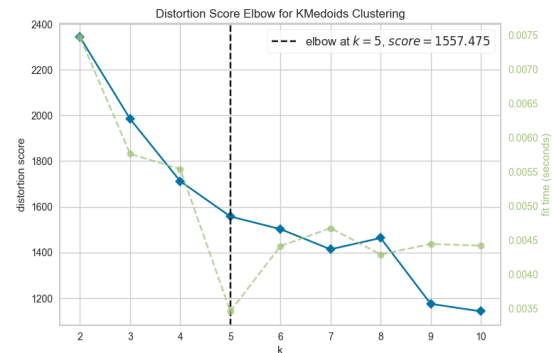
Linkage	SC	DI	Cophe netic	AC	CC
Single	0.35	0.209	0.79	0.78	19.4
Complete	0.38	0.245	0.84	0.81	18.1
Average	0.39	0.271	0.87	0.84	17.6
Ward	0.43	0.312	0.91	0.86	17.2

Clustering dengan Berbagai Linkage

Berdasarkan Tabel 3, menunjukkan bahwa metode *ward linkage* memberikan performa terbaik di antara seluruh variasi linkage. Nilai Silhouette Coefficient dan Dunn Index tertinggi menunjukkan kluster yang kompak dan terpisah dengan baik. Selain itu, Cophenetic Correlation yang tinggi serta kombinasi nilai Average Connectivity yang tinggi dan Cluster Connectivity yang rendah mengindikasikan stabilitas dan kualitas struktur kluster yang optimal. Dengan demikian, *ward linkage* merupakan pendekatan yang paling efektif dalam pengelompokan data pada penelitian ini.

4.3 K-Medoids Clustering

Sebelum proses klusterisasi dilakukan, langkah awal yang penting adalah menentukan jumlah kluster yang optimal. Dalam penelitian ini digunakan metode *elbow*, yang menganalisis perubahan nilai distorsi terhadap berbagai jumlah kluster k . Hasil grafik *elbow* menunjukkan titik belok paling signifikan pada kluster, sehingga dipilih lima kluster sebagai konfigurasi terbaik. Setelah jumlah kluster ditetapkan, algoritma K-Medoids dijalankan dengan memetakan setiap data ke medoid terdekat, kemudian memperbarui medoid secara iteratif hingga proses konvergen dan kluster terbentuk secara stabil.



Gambar 3. Hasil Metode *Elbow* K-Medoids Clustering

Setelah jumlah kluster optimal ditentukan sebanyak lima melalui metode *elbow*, algoritma K-Medoids diterapkan pada data hasil reduksi PCA. Proses klusterisasi menghasilkan lima kelompok berdasarkan jarak ke medoid terdekat, yaitu titik data aktual yang paling representatif dalam setiap kluster.

Kluster	SC	DI	CHI	DBI	XBI	CC
5	0.39	0.28	387.4	0.84	0.63	18.5

Tabel 4. Hasil Evaluasi K-Medoids Clustering

Berdasarkan Tabel 4, Evaluasi menggunakan enam metrik internal menunjukkan bahwa metode ini menghasilkan nilai Silhouette Coefficient sebesar 0.3968 dan Dunn Index sebesar 0.2875, dengan CHI sebesar 387.4. Sementara itu, nilai DBI dan XBI masing-masing tercatat sebesar 0.842 dan 0.639, serta Cluster Connectivity sebesar 18.53. Hasil ini mencerminkan kluster yang cukup baik, meskipun tidak seoptimal Agglomerative Clustering pada beberapa indikator.

4.4 Evaluasi Model Terbaik

Berdasarkan evaluasi yang telah dilakukan, dapat disimpulkan bahwa metode *clustering* terbaik adalah Agglomerative Hierarchical Clustering dengan linkage *ward* dan jumlah lima *cluster*.

Tabel 5. Evaluasi Model Terbaik

Metode	N Clusters	CC	DI	SC	DBI
AHC Ward	5	0.18	0.21	0.54	0.72
K-Medoids	5	2.11	0.03	0.09	2.06

Metode	CHI	XBI
AHC Ward	37.8	0.17
K-Medoids	61.7	1.56

Hal tersebut dapat dilihat bahwa metode *Agglomerative Clustering Ward Linkage* memiliki lima kategori yang jauh lebih unggul dibandingkan metode K-Medoids. indeks tersebut terdiri dari *connectivity coefficient*, *dunn index*, *silhouette coefficient*, *davies bouldin index*, dan *xie and beni index*.

4.5 Profiling Analisis Cluster

Setelah menentukan metode *clustering* yang terbaik untuk penelitian ini, kita kemudian dapat melanjutkan pada tahap selanjutnya yaitu pengelompokan dan profiling hasil analisis *cluster*. Dengan metode *Agglomerative Clustering*, kita mendapatkan pengelompokan kabupaten/kota sebagai berikut untuk kategori pedesaan.

Tabel 6. Hasil pengelompokan menggunakan metode ward linkage untuk kategori pedesaan

Clusters	Pedesaan
1	Ogan Komering Ilir, Musi Rawas, Ogan Komering Ulu Selatan, Ogan Komering Ulu Timur, Empat Lawang, Kepahiang, Lampung Barat, Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utarar, Way Kanan, Tulangbawang, Pesawaran, Pringsewu, Mesuji, Pesisir Barat, Kota Metro, Cilacap, Banyumas, Purworejo, Wonosobo, Boyolali, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Rembang, Temanggung, Kendal, Batang, Gunung Kidul, Pacitan, Ponorogo, Trenggalek, Tulungagung, Kediri, Lumajang, Banyuwangi, Mojokerto, Nganjuk, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, Kota Mojokerto, Jembrana, Tabanan, Badung, Gianyar, Klungkung, Bangli, Karangasem, Buleleng, Lombok Barat, Lombok Tengah, Lombok Timur, Lombok Utara, Kotawaringin Barat, Sukamara, Lamandau, Katingan, Pulang Pisau, Kota Baru, Tanah Bumbu, Kutai Kartanegara, Kutai Timur, Penajam Paser Utara, Malinau, Bolaang Mongondow, Minahasa, Bolaang Mongondow Utara, Minahasa Tenggara, Bolaang Mongondow Selatan, Bolaang Mongondow Timur, Kota Bitung, Kota Tomohon, Kota

	Kotamobagu, Morowali, Poso, Buol, Tojo Una-Una, Sigi, Banggai Laut, Morowali Utara, Luwu Timur, Kota Makasar, Boalemo, Bone Bolango, Tambrau, Kota Sorong, Biak Numfor, Keerom, Boven Digoel, Mimika, Deiyai, Nabire, Paniai, Intan Jaya, Puncak Jaya, Mamberamo Tengah
2	Indragiri Hilir, Kepulauan Meranti, Muara Enim, Musi Banyuasin, Ogan Ilir, Kota Prabumulih, Kota Pagar Alam, Bangka, Bangka Tengah, Bangka Selatan, Belitung Timur, Natuna, Lingga, Purbalingga, Kebumen, Klaten, Sukoharjo, Kudus, Jepara, Brebes, Kulon Progo, Bantul, Sleman, Jember, Bondowoso, Situbondo, Sidoarjo, Jombang, Bangkalan, Sampang, Pamekasan, Sumenep, Kota Batu, Sumbawa, Sumbawa Barat, Sabu Raijua, Sanggau, Sintang, Kotawaringin Timur, Kapuas, Barito Selatan, Barito Utara, Seruyan, Gunung Mas, Barito Timur, Tanah Laut, Barito Kuala, Tapin, Hulu Sungai Selatan, Hulu Sungai Utara, Tabalong, Balangan, Paser, Kutai Barat, Berau, Mahakam Ulu, Kota Bontang, Bulungan, Tana Tidung, Nunukan, Bulukumba, Gowa, Kolaka Timur, Maluku Tengah, Buru, Halmahera Utara, Halmahera Timur, Pulau Morotai, Fakfak, Kaimana, Teluk Wondama, Teluk Bintuni, Maybrat, Kepulauan Yapen, Nduga, Jayawijaya
3	Nias Barat, Cianjur, Garut, Ciamis, Kuningan, Majalengka, Purwakarta, Pandeglang, Bima, Kota Bima, Sumba Timur, Timor Tengah Selatan, Timor Tengah Utara, Belu, Alor, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Sumba Tengah, Nagekeo, Malaka, Sambas, Bengkayang, Landak, Mempawah, Kapuas Hulu, Melawi, Kayong Utara, Kota Singkawang, Murung Raya, Hulu Sungai Tengah, Banggai Kepulauan, Kepulauan Selayar, Bantaeng, Sinjai, Maros, Pangkajene dan Kepulauan, Barru, Bone, Soppeng, Wajo, Pinrang, Luwu, Tana Toraja, Toraja Utara, Kota Parepare, Kota Palopo, Buton, Muna, Konawe, Bombana, Wakatobi, Kolaka Utara, Buton Utara, Konawe Utara, Konawe Kepulauan, Buton Tengah, Buton Selatan, Kota Baubau, Majene, Mamasa, Maluku Tenggara Barat, Maluku Tenggara, Kepulauan Aru,

	Seram Bagian Barat, Seram Bagian Timur, Maluku Barat Daya, Buru Selatan, Kota Tual, Halmahera Barat, Halmahera Tengah, Kepulauan Sula, Kota Tidore Kepulauan, Raja Ampat, Sorong Selatan, Sarmi, Supiori, Merauke, Mappi, Asmat, Tolikara
4	Aceh Singkil, Aceh Tenggara, Aceh Timur, Aceh Tengah, Bireuen, Aceh Utara, Gayo Lues, Aceh Tamiang, Nagan Raya, Bener Meriah, Pidie Jaya, Kota Sabang, Kota Langsa, Kota Lhokseumawe, Tapanuli Selatan, Tapanuli Utara, Toba Samosir, Labuhan Batu, Asahan, Simalungun, Dairi, Karo, Deli Serdang, Langkat, Humbang Hasundutan, Samosir, Serdang Bedagai, Batu Bara, Labuhan Batu Selatan, Labuhan Batu Utara, Kota Tanjung Balai, Kota Pematang Siantar, Kota Tebing Tinggi, Kota Binjai, Pesisir Selatan, Solok, Sijunjung, Tanah Datar, Agam, Lima Puluh Kota, Solok Selatan, Dharmasraya, Kota Solok, Kota Sawah Lunto, Kota Bukittinggi, Kota Payakumbuh, Kota Pariaman, Kuantan Singingi, Indragiri Hulu, Pelalawan, Siak, Kampar, Rokan Hulu, Bengkalis, Rokan Hilir, Kota Dumai, Kerinci, Merangin, Sarolangun, Batang Hari, Tanjung Jabung Timur, Tanjung Jabung Barat, Tebo, Bungo, Kota Sungai Penuh, Lahat, Banyu Asin, Rejang Lebong, Mukomuko, Lebong, Karimun, Bintan, Pati, Demak
5	Simeulue, Aceh Selatan, Aceh Barat, Aceh Besar, Pidie, Aceh Jaya, Kota Subulussalam, Nias, Mandailing Natal, Tapanuli Tengah, Nias Selatan, Pakpak Bharat, Nias Utara, Kota Sibolga, Kota Gunungsitoli, Kepulauan Mentawai, Pasaman, Pasaman Barat, Ogan Komering Ulu, Musi Rawas Utara, Kota Lubuklinggau, Kaur, Seluma, Belitung, Bangka Barat, Kepulauan Anambas, Kep. Seribu, Indramayu, Subang, Karawang, Pangandaran, Lebak, Dompu, Sumba Barat, Ketapang, Sekadau, Kubu Raya, Kota Tarakan, Siau Tagulandang Biaro, Banggai, Donggala, Parigi Moutong, Sidenreng Rappang, Enrekang, Polewali Mandar, Halmahera Selatan

Dan sebagai berikut untuk kategori perkotaan:

Tabel 6. Hasil pengelompokan menggunakan metode ward linkage untuk kategori perkotaan

<i>Clusters</i>	Kabupaten / Kota
1	Banjarnegara, Magelang, Semarang, Blitar, Malang, Madiun, Kota Madiun, Kota Denpasar, Kota Manado, Kota Palu, Gorontalo, Kota Gorontalo, Kota Ambon, Manokwari, Manokwari Selatan, Jayapura, Kota Jayapura,
2	Kota Palembang, Pekalongan, Pemalang, Tegal, Kota Magelang, Probolinggo, Pasuruan, Kota Kediri, Kota Blitar, Kota Malang, Kota Pasuruan, Tangerang, Banjar, Kota Samarinda, Mamuju Tengah, Kota Ternate
3	Sukabumi, Bandung, Tasikmalaya, Sumedang, Kota Tasikmalaya, Kota Probolinggo, Kupang, Kota Kupang, Kota Banjarmasin, Kota Kendari,
4	Padang Lawas Utara, Kota Medan, Kota Padangsidimpuan, Padang Pariaman, Kota Padang, Kota Padang Panjang, Kota Pekanbaru, Muaro Jambi, Kota Jambi, Bengkulu Utara, Bengkulu Tengah, Kota Pangkal Pinang, Kota Tanjung Pinang, Kota Bekasi, Kota Semarang, Kota Surabaya, Kota Tangerang Selatan, Kota Mataram
5	Kota Banda Aceh, Bengkulu Selatan, Kota Bengkulu, Kota Bandar Lampung, Kota Batam, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Jakarta Pusat, Kota Jakarta Barat, Kota Jakarta Utara, Bogor, Cirebon, Bekasi, Bandung Barat, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Depok, Kota Cimahi, Kota Banjar, Kota Surakarta, Kota Salatiga, Kota Pekalongan, Kota Tegal, Kota Yogyakarta, Serang, Kota Tangerang, Kota Cilegon, Kota Serang, Kota Pontianak, Kota Palangka Raya, Kota Banjar Baru, Kota Balikpapan, Mamuju, Mamuju Utara

Hasil segmentasi wilayah menunjukkan lima klaster utama untuk kategori pedesaan dan perkotaan berdasarkan karakteristik geografis, aksesibilitas, dan tingkat pembangunan sosial-ekonomi.

4.6 Interpretasi Hasil Clustering

Untuk wilayah pedesaan, Klaster 1 mencakup daerah-daerah perifer dan terpencil di bagian timur Indonesia, termasuk Papua,

Nusa Tenggara, serta sebagian Kalimantan dan Sulawesi. Wilayah ini ditandai oleh hambatan geografis yang tinggi, infrastruktur terbatas, dan akses pelayanan publik yang masih minim, sehingga memerlukan prioritas intervensi pembangunan dasar. Klaster 2 terdiri dari wilayah penyangga dan pesisir seperti Bangka Belitung, Kalimantan bagian tengah, dan sebagian Nusa Tenggara Barat. Meskipun tidak se-terpencil Klaster 1, wilayah ini menghadapi tantangan dalam pemerataan layanan dan pengembangan ekonomi berbasis sumber daya alam lokal. Klaster 3 mencakup wilayah pegunungan dan kepulauan di Sulawesi Selatan, Nusa Tenggara Timur, serta Maluku. Daerah ini menunjukkan karakter budaya yang kuat, namun terhambat oleh konektivitas antarpulau dan akses transportasi yang masih rendah. Klaster 4 didominasi oleh daerah-daerah pedesaan di Sumatra bagian tengah dan utara yang relatif lebih berkembang secara infrastruktur namun tetap menghadapi disparitas dalam distribusi kesejahteraan. Sementara itu, Klaster 5 terdiri dari kawasan perbatasan dan kepulauan seperti Kepulauan Mentawai, Anambas, dan Nias. Lokasi-lokasi ini menghadapi persoalan isolasi geografis dan keterbatasan infrastruktur dasar, menjadikannya prioritas dalam konteks pemerataan pembangunan nasional.

Untuk wilayah perkotaan, Klaster 1 memuat kota-kota kecil dan sedang di Jawa Tengah, Bali, serta bagian timur Indonesia yang menunjukkan fungsi regional meskipun belum menjadi pusat ekonomi utama. Kota-kota ini memiliki potensi pertumbuhan sebagai simpul wilayah tengah. Klaster 2 terdiri dari kota-kota yang berfungsi sebagai pusat transit atau industri penyangga, seperti di Jawa Timur, Kalimantan Timur, dan sebagian Sumatra. Kota-kota ini strategis dalam rantai pasok dan distribusi, namun memerlukan penguatan infrastruktur pelayanan. Klaster 3 menghimpun kota-kota menengah di Jawa Barat dan Nusa Tenggara, yang berkembang pesat namun belum optimal dalam tata kelola perkotaan. Klaster ini potensial dikembangkan melalui sektor UMKM dan ekonomi lokal. Klaster 4 berisi ibu kota provinsi di Sumatra dan Kalimantan yang telah mengalami urbanisasi tinggi dan menjadi pusat pemerintahan serta ekonomi, namun masih menghadapi tantangan tata ruang dan ketimpangan intra-kota.

Terakhir, Klaster 5 meliputi kota-kota metropolitan seperti Jakarta, Bandung, Surabaya, Medan, dan Batam. Sebagai pusat pertumbuhan nasional, klaster ini memiliki fasilitas dan konektivitas tertinggi, tetapi juga rentan terhadap masalah kepadatan, kemacetan, dan disparitas sosial.

Dan untuk kategori dari setiap *cluster* berdasarkan konsumsi rempah - rempahnya. Diperoleh bahwa *cluster* kelima adalah *cluster* dengan konsumsi rempah terendah di antara *cluster* lainnya. kemudian *cluster* keempat memiliki tingkat konsumsi bawang bombay yang paling tinggi di antara semu *cluster* lainnya. kemudian diikuti dengan *cluster* ketiga dengan konsumsi rempah - rempah yang ada di rata - rata di antara *cluster* lainnya. *cluster* kedua yang condong pada konsumsi cabai - cabai seperti cabai merah, cabai hijau, dan cabai rawit. sedangkan untuk *cluster* pertama yang condong pada konsumsi bawang - bawang non bawang bombay.

5. KESIMPULAN

Dari penelitian yang telah dilakukan dapat disimpulkan bahwa:

- Metode Agglomerative Hierarchical Clustering (AHC) dengan linkage *ward* terbukti paling efektif dalam mengelompokkan kabupaten/kota berdasarkan konsumsi rempah-rempah, dengan performa evaluasi terbaik pada metrik Silhouette Score, Dunn Index, Xie-Beni, dan Connectivity.
- Metode K-Medoid, meskipun lebih stabil terhadap outlier, menghasilkan klasterisasi yang kurang optimal dibandingkan AHC berdasarkan seluruh metrik evaluasi yang digunakan.
- Segmentasi wilayah menunjukkan adanya perbedaan konsumsi rempah yang signifikan antara wilayah pedesaan dan perkotaan, di mana klaster-klaster tertentu menggambarkan karakteristik geografis dan sosial ekonomi yang khas.
- Kelebihan penelitian ini terletak pada pendekatan komparatif dan penggunaan berbagai metrik evaluasi serta PCA untuk visualisasi, yang memperkuat validitas hasil klasterisasi.
- Keterbatasan penelitian meliputi cakupan jenis rempah yang terbatas serta tidak mempertimbangkan variabel pendukung

- lain seperti pendapatan, ketersediaan bahan pangan, atau budaya konsumsi.
- f. Pengembangan selanjutnya dapat diarahkan pada klasterisasi multivariabel dengan penambahan atribut sosial-ekonomi, serta pemanfaatan metode *clustering* berbasis deep learning atau spatio-temporal untuk hasil yang lebih komprehensif.

DAFTAR PUSTAKA

- [1] S. Mayasari, N. Isnawati, F. Aprillia, D. Wigati, K. Agung, and Z. Permana, "Edukasi Tanaman Rempah Untuk Meningkatkan Pengetahuan Pada Karang Taruna Garuda Cempaka di Kelurahan Gebang-Jember," *BERNAS J. Pengabd. Kpd. Masy.*, vol. 4, no. 3, pp. 2196–2200, 2023, [Online]. Available: <https://doi.org/10.31949/jb.v4i3.5978>
- [2] G. . Muliana, R. Magfirah, and U. Hasanah, "Potensi Herba dan Rempah Sebagai Tanaman Obat Keluarga," *Bioma*, vol. 5, no. 2, pp. 1–11, 2023.
- [3] A. Fitriani, R. Wiwin Dari, I. Siregar, and B. Purnomo, "Jejak Historis dan Peran Indonesia Dalam Jalur Rempah Sebagai Warisan Budaya Bahari," *Krinok*, vol. 2, no. 2, pp. 9–18, 2023, doi: 10.22437/krinok.v2i2.18495.
- [4] Y. Robi, S. M. Kartikawati, and . Muflihati, "Etnobotani Rempah Tradisional Di Desa Empoto Kabupaten Sanggau Kalimantan Barat," *J. Hutan Lestari*, vol. 7, no. 1, pp. 130–142, 2019, doi: 10.26418/jhl.v7i1.31179.
- [5] Y. M. Roza, G. Razali, E. Fatmawati, S. Syamsuddin, and G. A. Wibowo, "Identitas Budaya Dan Sosial Pada Makanan Khas Daerah: Tinjauan Terhadap Perilaku Konsumsi Masyarakat Muslim Pada Bulan Ramadan Di Indonesia," *Komitmen J. Ilm. Manaj.*, vol. 4, no. 1, pp. 305–315, 2023, doi: 10.15575/jim.v4i1.25031.
- [6] V. Number, I. Rizqi, Z. Nisa, E. Zulfiani, F. I. Kesehatan, and U. I. A. Kuningan, "Lentera : Multidisciplinary Studies LITERATURE REVIEW PARADIGMA GIZI : EKSPLORASI TABU POLA KONSUMSI MAKANAN DAN PENGETAHUAN GIZI DI MASYARAKAT PEDESAAN," vol. 3, no. 1, pp. 1–10, 2024.
- [7] E. Widodo, P. Ermayani, L. N. Laila, and A. T. Madani, "Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative Clustering (Indonesian Province Grouping Based on Poverty Level Using Hierarchical Agglomerative Clustering Analysis)," *Semin. Nas. Off. Stat.*, vol. 2, pp. 557–566, 2021.
- [8] F. Zahra, A. Khalif, and B. N. Sari, "Pengelompokan Tingkat Kemiskinan di Setiap Provinsi di Indonesia Menggunakan Algoritma K-Medoids," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, pp. 2830–7062, 2024, [Online]. Available: <http://dx.doi.org/10.23960/jitet.v12i2.4199>
- [9] L. Hakim, J. Batoro, and K. Sukenti, "Etnobotani Rempah-Rempah di Dusun Kopen Dukuh, Kabupaten Banyuwangi," *J. Pembang. dan alam lestari*, vol. 6, no. 2, pp. 133–142, 2015.
- [10] R. Swastika, S. Mukodimah, F. Susanto, M. Muslihudin, and S. Ipnuwati, *IMPLEMENTASI DATA MINING : (Clustering, Association, Prediction, Estimation, Classification)*. Penerbit Adab, 2023.
- [11] R. Wahyusar, "Penerapan Algoritma K-Medoids Untuk Mengelompokkan Status Obesitas," *Simetris2*, vol. 18, no. 1, pp. 329–334, 2024.
- [12] R. Kusumastuti, E. Bayunanda, A. M. Rifa'i, M. R. G. Asgar, F. I. Ilmawati, and K. Kusriani, "Clustering Titik Panas Menggunakan Algoritma Agglomerative Hierarchical Clustering (AHC)," *CogITO Smart J.*, vol. 8, no. 2, pp. 501–513, 2022, doi: 10.31154/cogito.v8i2.438.501-513.
- [13] R. O. Pratikto and N. Damastuti, "Klasterisasi Menggunakan Agglomerative Hierarchical Clustering Untuk Memodelkan Wilayah Banjir," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 6, no. 1, p. 13, 2021.
- [14] T. B. Nugroho and E. Sugiharti, "The Improvement of C4.5 Algorithm Accuracy in Predicting Forest Fires Using Discretization and AdaBoost," *J. Adv. Inf. Syst. Technol.*, vol. 3, no. 1, pp. 43–52, 2021, doi: 10.15294/jaist.v3i1.49094.
- [15] A. Riski, D. Saputra, and B. Prasetyo, "Sentiment Analysis on Twitter Social Media Regarding Covid-19 Vaccination with Naive Bayes Classifier (NBC) and Bidirectional Encoder Representations from Transformers (BERT)," vol. 2, no. 2, pp. 106–113, 2024, doi: 10.15294/rji.v8i1.25356.
- [16] A. Rofifah, R. Goenjantoro, and D. Yuniarti, "Perbandingan Pengelompokan K-Means dan K-Medoids Pada Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Studi Kasus : Data Titik Panas Di Indonesia Pada 28 April 2018)," *J. EKSPONENSIAL*, vol. 10, no. 2, pp. 143–152, 2019.
- [17] F. Ardiansyah, F. Hamdan, S. Sugiyanto, and I. Wahyu Siadi, "Klasifikasi Customer

- Relationship Management Menggunakan Dataset KDD Cup 2009 dengan Teknik Reduksi Dimensi,” *Komputika J. Sist. Komput.*, vol. 11, no. 2, pp. 193–202, 2022, doi: 10.34010/komputika.v11i2.6498.
- [18] L. Ramadhani, I. Purnamasari, and F. D. T. Amijaya, “Penerapan Metode Complete Linkage dan Metode Hierarchical Clustering Multiscale Bootstrap (Studi Kasus: Kemiskinan Di Kalimantan Timur Tahun 2016),” *Ekspansional*, vol. 9, no. 2016, pp. 1–10, 2018.
- [19] N. Mirantika, T. S. Syamfithriani, and R. Trisudarmo, “Implementasi Algoritma K-Medoids Clustering Untuk Menentukan Segmentasi Pelanggan,” *J. Nuansa Inform.*, vol. 17, no. 1, pp. 2614–5405, 2023, [Online]. Available: <https://journal.uniku.ac.id/index.php/ilkom>
- [20] M. Rahmah, A. Candra, and R. W. Sembiring, “Identifikasi Predikat Hasil Pengelompokan Data Kualitas Udara dengan Menggunakan Affinity Propagation dan Silhouette Coefficient,” *InfoTekJar J. Nas. Inform. dan Teknol. Jar.*, vol. 6, no. 2, pp. 177–180, 2022, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/4670>
- [21] B. D. Satoto, B. K. Khotimah, and A. Muhammad, “Pengelompokan Tingkat Kesehatan Masyarakat Menggunakan Self Organizing Map Dengan Cluster Validation Idb dan I-Dunn,” *Semin. Nas. Apl. Teknol. Inf.* 2015, vol., pp. 1–6, 2015.
- [22] C. Tjipta *et al.*, “COMPARISON OF K-MEANS ++ AND AGGLOMERATIVE HIERARCHICAL METHODS IN CLUSTERING HEALTHCARE WORKERS KOMPARASI METODE K-MEANS ++ DAN AGGLOMERATIVE,” vol. 10, no. 2, pp. 717–728, 2025.
- [23] E. Klaster, F. C-means, and S. Pelanggan, “EVALUASI KLASSTER ALGORITMA FUZZY C-MEANS UNTUK PENENTUAN JUMLAH KELOMPOK OPTIMAL PADA SEGMENTASI PELANGGAN Annisa Permatasari Ayuningtyas , Deden Istiawan,” 2017.
- [24] G. R. Suraya and A. W. Wijayanto, “Comparison of Hierarchical Clustering, K-Means, K-Medoids, and Fuzzy C-Means Methods in Grouping Provinces in Indonesia according to the Special Index for Handling Stunting,” *Indones. J. Stat. Its Appl.*, vol. 6, no. 2, pp. 180–201, 2022, doi: 10.29244/ijsa.v6i2p180-201.