

PENGELOMPOKAN WILAYAH KABUPATEN/KOTA DI INDONESIA BERDASARKAN KETAHANAN PANGAN

Gabriel Nathanael Irawan¹⁾ Teny Handayani²⁾ Irvan Lewenusa³⁾

^{1) 2) 3)} Teknik Informatika Universitas Tarumanagara

Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

email : gabriel.535220142@stu.untar.ac.id ¹⁾, tenyh@fti.untar.ac.id ²⁾, irvanl@fti.untar.ac.id ³⁾

ABSTRAK

Penelitian ini bertujuan untuk melakukan pengelompokan terhadap kabupaten/kota di Indonesia berdasarkan empat indikator utama ketahanan pangan, yaitu Indeks Ketahanan Pangan (IKP), Skor Pola Pangan Harapan (PPH), Rasio Angka Kecukupan Energi (AKE), dan Rasio Angka Kecukupan Protein (AKP) pada periode 2021-2023. Keempat indikator tersebut digunakan untuk menggambarkan kondisi ketersediaan, keragaman, dan kecukupan pangan di setiap wilayah Indonesia. Hasil pengujian menunjukkan bahwa algoritma Spectral Bridges memberikan performa terbaik pada jumlah cluster dua, dengan nilai silhouette sebesar 0,62 dan Davies-bouldin Index (DBI) sebesar 0,53. Nilai tersebut mengindikasikan bahwa metode ini mampu mengelompokkan wilayah dengan karakteristik ketahanan pangan yang akurat, terpisah dengan baik antar clusternya, serta memberikan hasil yang stabil dan dapat menjadi representatif wilayah ketahanan pangan Indonesia.

Key words

Ketahanan pangan, Spectral Bridges, Clustering kabupaten/kota

1. Pendahuluan

Wilayah Indonesia memiliki keuntungan dalam segi geografis, karena sebagian besar wilayah Indonesia dijadikan lahan pertanian untuk beragam komoditas pangan. Namun, ketahanan pangan masih menjadi tantangan pada sejumlah daerah akibat pertumbuhan penduduk yang tinggi dan alih fungsi lahan pertanian menjadi area yang tidak produktif [1]. Beberapa wilayah di Indonesia seperti Indonesia bagian timur memiliki tantangan dari segi geografis, keterbatasan infrastruktur, dan fluktuasi harga pangan [2]. Selain faktor geografis, aspek sosial-ekonomi juga berperan dalam menentukan tingkat ketahanan pangan di suatu wilayah, seperti kondisi kemiskinan dan akses terhadap pangan bergizi [3]. Beberapa bahan pokok yang menjadi kebutuhan utama masyarakat Indonesia antara lain beras, daging

sapi, ayam, telur, bawang merah, bawang putih, cabai, minyak goreng, dan gula [4].

Ketahanan pangan sendiri merupakan isu strategis nasional yang juga menjadi bagian dari *Sustainable Development Goals (SDGs)* poin kedua, yaitu mengakhiri kelaparan dan meningkatkan ketahanan pangan berkelanjutan [5].

Terdapat beberapa indikator yang dapat menggambarkan ketahanan pangan suatu wilayah. Pertama terdapat Indeks Ketahanan Pangan (IKP) [6], yang digunakan sebagai alat ukur utama dalam mengevaluasi capaian ketahanan pangan pada setiap kabupaten/kota. Selain itu, terdapat Skor Pola Pangan Harapan (PPH) yang menilai keragaman konsumsi pangan masyarakat. Kemudian terdapat indikator Rasio Angka Kecukupan Energi (AKE) dan Rasio Angka Kecukupan Protein (AKP) yang menilai kuantitas konsumsi pangan [7]. Dengan memanfaatkan keempat indikator ini, kondisi ketahanan pangan setiap daerah dapat dianalisis secara lebih menyeluruh.

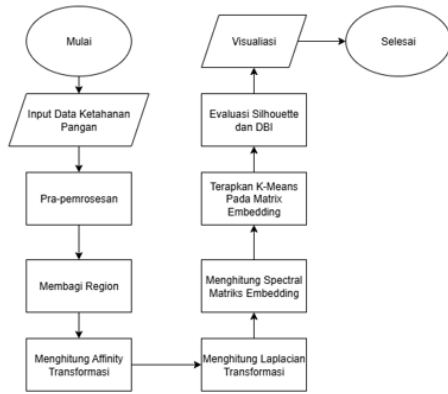
Pendekatan *clustering* pada penelitian ini menggunakan algoritma Spectral Bridges. Spectral Bridges merupakan algoritma clustering terkini yang menggabungkan konsep Spectral Clustering dan K-Means [8]. Algoritma ini bekerja dengan membangun struktur graf dan menentukan *affinity bridges* pada setiap data, sehingga mampu menghasilkan pembagian kelompok yang jelas.

2. Metodologi

2.1. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode *unsupervised learning*, yaitu clustering. Tujuan dari penelitian ini adalah untuk mengelompokkan wilayah di Indonesia berdasarkan indikator ketahanan pangan, sehingga diperoleh pola atau karakteristik dari masing-masing kelompok yang terbentuk. Spectral Bridges dipilih karena algoritma ini dapat menangkap struktur data yang kompleks dan non-linear, sehingga algoritma ini lebih akurat dibandingkan

metode *clustering* konvensional [8]. Alur penelitian dapat dilihat pada **Gambar 1**.



Gambar 1. Alur Penelitian

2.2. Pengumpulan Data

Dataset yang merupakan empat dataset berbeda, yaitu Indeks Ketahanan Pangan, Skor Pola Pangan Harapan, Rasio Angka Kecukupan Protein, dan Rasio Angka Kecukupan Lemak. Dataset ini diperoleh dari situs Badan Pangan Nasional pada periode 2021-2023. Dataset ini terdiri dari 514 baris pada masing-masing kabupaten/kota Indonesia. Dataset ini kemudian digabung menjadi satu dataset, yang kemudian dataset ini akan diproses untuk dilakukan *clustering* dengan *Spectral Bridges*.

2.3. Pre-Processing Data

Pre-processing data meliputi penanganan *missing values*, penanganan duplikat dan normalisasi rentang data. Metode normalisasi menggunakan MinMax, yaitu metode normalisasi dengan mengubah rentang data menjadi 0-1 [9]. Rumus

Pemrosesan data yang digunakan adalah mengisi missing value dengan menggunakan imputer dan melakukan normalisasi MinMax. Rumus untuk normalisasi dengan MinMax dapat dilihat pada persamaan berikut :

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

Keterangan :

x' = nilai hasil normalisasi

x_i = nilai tertentu yang akan dinormalisasi

(x) = nilai maksimal dari sebuah atribut

$\min(x)$ = nilai minimal dari sebuah atribut

2.4. Spectral Bridges

Spectral Bridges merupakan algoritma *clustering* yang menggunakan prinsip *K-Means* dan *Spectral Clustering* menjadi satu metode yang sama. Pada tahap awal, data dikelompokkan ke dalam beberapa *Voronoi*

region dengan masing-masing memiliki titik pusat. Hubungan antar *region* kemudian dihitung menggunakan konsep *bridge affinity*, yaitu pengukuran inertia antar *centroid* [8]. Langkah algoritma ini adalah sebagai berikut:

1. Menerapkan *K-Means* untuk membuat *region*.
2. Mencari nilai jarak antar dua *region* dengan rumus (1).
3. Menghitung bobot *affinity* dengan menggunakan nilai t_i dengan rumus (3)
4. Menghitung nilai *affinity* antar *region* dengan rumus (4)
5. Menghitung nilai γ dengan rumus (6)
6. Menghitung nilai *affinity* setelah transformasi dengan rumus (5)
7. Menghitung matriks Laplacian ternormalisasi dengan rumus (7)
8. Menghitung *eigenvalue* dengan rumus (8)
9. Melakukan *spectral embedding* dengan rumus (9)
10. Matriks embedding sudah terbentuk, dan dapat diselesaikan dengan algoritma *K-Means*.

$$t_i = \min \left(1, \max \left(0, \frac{(\mu_l - \mu_k)}{\|\mu_l - \mu_k\|^2} \right) \right) \quad (2)$$

Keterangan :

t_i = Jarak antara dua *region*

μ_l = Centroid titik l

μ_k = Centroid titik k

x_i = Titik i

$$\alpha_i = \{t_i, \quad t_i \leq 0,5, 1 - t_i, \quad t_i > 0,5\} \quad (3)$$

Keterangan:

α_i = Nilai bobot *affinity* titik i

$$a_{ik} = \frac{\sum_{x_i \in v_k \cup v_l} a_i^2}{n_k + n_l} \quad (4)$$

Keterangan:

a_{ik} = Nilai *affinity* dari region i dan region k

α_i = nilai *affinity* titik i

n_k = jumlah titik pada region k

n_l = Jumlah titik pada region l

$$\widetilde{\alpha}_{kl} = \exp(\sqrt{\alpha_{kl}} \times \gamma) \quad (5)$$

Keterangan:

$\widetilde{\alpha}_{kl}$ = Nilai *affinity* setelah transformasi

α_{kl} = Nilai *affinity* sebelum transformasi

γ = Faktor skala

$$\gamma = \frac{\log(M)}{q_{90} - q_{10}} \quad (6)$$

Keterangan:

γ = Faktor skala

M = Jumlah elemen

q_{90} = Persentil ke-90

q_{10} = Persentil ke-10

$$L_{sym} = I - D^{-1/2}AD^{-1/2} \quad (7)$$

Keterangan:

L_{sym} = Nilai normalisasi Laplacian

I = Matriks identitas

A = Matriks bobot *affinity*

D = Matriks derajat

$$\det(L - \lambda I) = 0 \quad (8)$$

Keterangan:

L = Matriks normalisasi Laplacian

λ = *eigenvalue*

I = Matriks identitas berukuran L

$$\|u_i\|_2 = \sqrt{\sum_{j=1}^k u_{ij}^2}, \quad Y_{ij} = \frac{u_{ij}}{\|u_i\|_2} \quad (9)$$

Keterangan:

Y_{ij} = Elemen pada matriks *embedding* ternormalisasi

k = Jumlah *eigen vektor* yang dipakai

u_{ij} = Elemen pada baris i dan kolom j dari matriks U

2.5. Metode Evaluasi

Metode evaluasi pada penelitian ini menggunakan dua metrik evaluasi yaitu *silhouette coefficient* dan *Davies-bouldin Index (DBI)*. *Silhouette coefficient* merupakan metrik yang mengukur sejauh mana data berada dalam *cluster* yang tepat dengan membandingkan jarak rata-rata data terhadap anggota dalam *cluster* yang sama [10]. Rumus *silhouette coefficient* dapat dilihat pada:

$$s_i = \frac{(b(i) - a(i))}{(a(i), b(i))} \quad (10)$$

Keterangan :

a_i = rata-rata jarak dari data ke- i ke semua data lain yang berada pada *cluster* yang sama

b_i = rata-rata jarak dari data ke- i ke semua data pada *cluster* terdekat

Nilai dari *silhouette* ini berkisar dari -1 sampai 1, di mana nilai 1 menandakan bahwa data dapat tergolong dengan baik ke dalam *cluster*-nya, sedangkan bila data mendekati -1 menandakan bahwa data tidak memiliki hubungan yang kuat ke *cluster*-nya sendiri [11].

Sementara itu, *DBI* menilai kualitas *clustering* dengan rasio antara kekompakan dalam *cluster* dan jarak antar *cluster* [12]. Nilai *DBI* yang kecil menunjukkan bahwa hasil *cluster* dapat memisahkan kelompok data dengan baik [13]. Rumus perhitungan *DBI* dapat dilihat pada:

$$DBI = \frac{1}{k} \sum_{a=1}^k R_a \quad (11)$$

Keterangan :

DBI = nilai *davies bouldin index*

k = jumlah *cluster*

R_a = ukuran kemiripan antar *cluster*

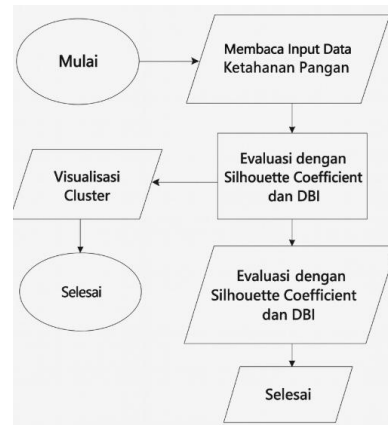
3. Hasil dan Pembahasan

3.1. Rancangan Sistem

Sistem yang dirancang bertujuan untuk mengelompokkan wilayah kabupaten/kota di Indonesia berdasarkan ketahanan pangannya. Alur kerja sistem dimulai dari proses pembacaan data hingga proses *clustering*. Setelah dibaca, data terlebih dahulu melalui tahap pra-pemrosesan yang akan menghapus nilai duplikat dan menangani *missing values*. Setelah *missing values* dan data duplikat ditangani, data dinormalisasi dengan MinMax agar rentang data sama (0-1). Setelah melalui tahap pra-pemrosesan, data siap diproses dengan algoritma *clustering*. Hasil *clustering* kemudian dievaluasi dengan metrik *silhouette* dan *DBI*. Hasil *cluster* yang terbentuk kemudian divisualisasikan dalam bentuk peta.

3.2. Flowchart Sistem

Flowchart sistem penelitian ini dapat dilihat pada Gambar 2.



Gambar 2 Flowchart Sistem

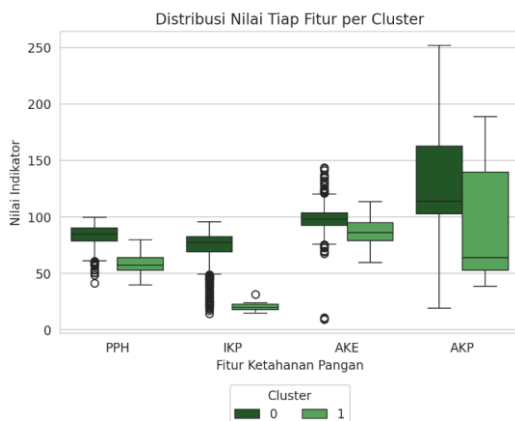
3.3. Hasil Clustering

Eksperimen dilakukan dengan menggunakan nilai *cluster* dari 2-7. Eksperimen ini menggunakan berbagai nilai *random state* dari 1-1000. Hasil pada Tabel 1 merupakan hasil *silhouette* dan *DBI* pada nilai *random state* terbaik.

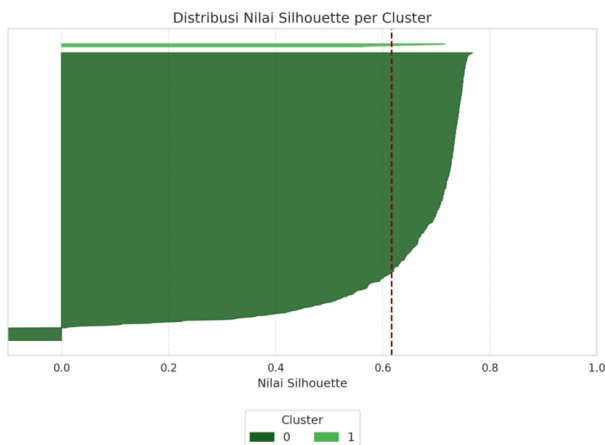
Tabel 1 Hasil eksperimen jumlah cluster metode CLARA

Jumlah Cluster	Rata-rata Silhouette	DBI	Waktu Komputasi (detik)
2	0,62	0,53	0,1536
3	0,57	1,94	0,1580
4	0,48	1,43	0,0581
5	0,34	2,07	0,0557
6	0,29	2,19	0,0608
7	0,24	1,53	0,0593

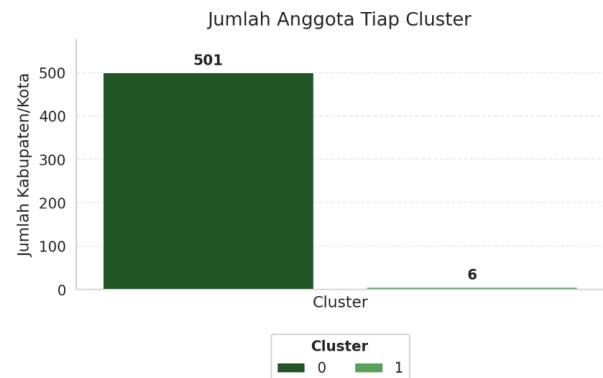
Pada tabel 1 dapat dilihat bahwa algoritma *spectral bridges* memiliki hasil *clustering* terbaik pada jumlah *cluster* 2 dengan nilai rata-rata *silhouette* 0,62 dan nilai *DBI* 0,53. Hal ini menunjukkan bahwa algoritma ini dapat menghasilkan pemisahan *cluster* yang baik. Visualisasi algoritma ini dapat dilihat pada Gambar 3, Gambar 4, Gambar 5, dan Gambar 6.



Gambar 3 Distribusi Nilai Fitur Tiap Cluster



Gambar 4 Visualisasi Silhouette



Gambar 5 Distribusi Anggota Tiap Cluster



Gambar 6 Visualisasi Peta

Pada visualisasi di atas, dapat dilihat pada wilayah yang tergolong dalam *cluster* dengan ketahanan pangan rendah umumnya berada di kawasan Indonesia bagian timur. Distribusi anggota *cluster* menunjukkan ketidakseimbangan, di mana *cluster* 0 memiliki 501 anggota, sedangkan *cluster* 1 hanya terdiri dari 6 anggota. Hal ini menandakan bahwa sebagian besar wilayah di Indonesia memiliki karakteristik ketahanan pangan yang relatif serupa, sementara hanya sebagian kecil wilayah yang memiliki karakteristik berbeda secara signifikan.

Pada visualisasi *boxplot* dapat dilihat bahwa terdapat perbedaan karakteristik antar *cluster*. Di mana *cluster* 0 merepresentasikan wilayah dengan tingkat ketahanan pangan yang lebih baik dibandingkan *cluster* 1, berdasarkan keempat indikator ketahanan pangan, yaitu IKP, PPH, AKE, dan AKP. Sementara itu, *cluster* 1 menunjukkan nilai yang lebih rendah pada seluruh indikator tersebut, menandakan bahwa wilayah yang termasuk dalam *cluster* 1 memiliki kondisi ketahanan pangan yang kurang baik dibandingkan dengan *cluster* 0.

4. Kesimpulan

Metode *spectral bridges* mampu mengelompokkan data ketahanan pangan menggunakan indikator IKP, PPH, AKE, dan AKP dengan baik, menghasilkan pembagian *cluster* yang jelas dan terpisah. Hasil ini dapat dilihat pada nilai *silhouette* (0,62) dan *DBI* (0,53) yang baik.

REFERENSI

- [1] V. V. Rumawas, H. Nayoan, and N. Kumayas, "Peran Pemerintah Dalam Mewujudkan Ketahanan Pangan di

- Kabupaten Minahasa Selatan (Studi Dinas Ketahanan Pangan Minahasa Selatan),” *JURNAL GOVERNANCE*, vol. 1, no. 1, pp. 1–12, 2021.
- [2] J. P. Immanuel and T. Handhayani, “PREDIKSI HARGA PANGAN JAYAPURA MENGGUNAKAN ELM, LSTM, LIGHTGBM, DAN GB,” *JIKSI*, vol. 13, no. 2, pp. 1–12, 2025, doi: <https://doi.org/10.24912/jiksi.v13i2.35122>.
- [3] M. Oni, M. D. Lauro, A. Winata, and T. Handhayani, “Analysis And Forecasting of Foodstuffs Prices in Bandung Using Gated Recurrent Unit,” *Jurnal Esensi Infokom : Jurnal Esensi Sistem Informasi dan Sistem Komputer*, vol. 7, no. 2, pp. 15–21, Oct. 2023, doi: 10.55886/infokom.v7i2.651.
- [4] A. Winata, M. D. L. Sitorus, and T. Handhayani, “Analysis and Prediction of Foodstuffs Prices in Tasikmalaya Using ELM and LSTM,” *SISTEMASI*, vol. 12, no. 3, pp. 874–887, 2023.
- [5] K. PPN/Bappenas, “Tujuan 2: Mengakhiri Kelaparan.” [Online]. Available: <https://sdgs.bappenas.go.id/17-goals/goal-2/>
- [6] B. P. N. (BAPANAS), *INDEKS KETAHANAN PANGAN TAHUN 2023*. 2023. [Online]. Available: https://drive.google.com/file/d/1P5KIdhdmZkVLWlpC82TaCH_3rCxQaLG6/view
- [7] B. P. N. (BAPANAS), *PEDOMAN PENILAIAN SKOR POLA PANGAN HARAPAN*. 2023. [Online]. Available: https://drive.google.com/file/d/13tNvG06JhO7xx4oQtOY9VjdlkNAfp_J/view
- [8] F. Laplante and C. Ambroise, “Spectral Bridges,” *Computo*, 2024.
- [9] P. P. Allorerung, A. Erna, M. Bagussahrir, and S. Alam, “Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 9, no. 3, pp. 178–191, Sep. 2024, doi: 10.14421/jiska.2024.9.3.178-191.
- [10] R. Hidayati, A. Zubair, A. H. Pratama, and L. Indana, “Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering,” *Techno.Com*, vol. 20, no. 2, pp. 186–197, May 2021, doi: 10.33633/tc.v20i2.4556.
- [11] V. Nellie, V. C. Mawardi, and N. J. Perdana, “IMPLEMENTASI METODE AGGLOMERATIVE HIERARCHICAL CLUSTERING UNTUK SISTEM REKOMENDASI FILM,” *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 11, no. 1, Jun. 2023, doi: 10.24912/jiksi.v11i1.24070.
- [12] Gusrianty and M. Siddik, “ANALISIS PENGUKURAN NUMERIK PADA OPTIMALISASI CLUSTER OBAT DI APOTEK X MENGGUNAKAN DAVIES BOULDIN INDEX,” *JOISIE (Journal Of Information Systems And Informatics Engineering)*, vol. 5, no. 1, pp. 60–66, 2021.
- [13] A. Chandra, “PERBANDINGAN ALGORITMA CLUSTERING K-MEANS, GAUSSIAN MIXTURE MODEL, DAN DBSCAN PADA DATA INDEKS STANDAR PENCEMAR UDARA (ISPU) DI PROVINSI DKI JAKARTA,” *Jurnal Komputer dan Informatika*, vol. 19, no. 1, pp. 1–8, 2024.

Gabriel Nathanael Irawan, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara