

Klasterisasi Status Gizi Balita di Jawa Barat Menggunakan *Self-Organizing Maps* (SOM)

Nia Widianengsih *, Reny Rian Marliana

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Islam Bandung, Indonesia.

niawidia889@gmail.com, renyrianmarliana@unisba.ac.id

Abstract. Nutritional problems among toddlers remain a critical issue in Indonesia, particularly in West Java Province, which has the largest population and a stunting prevalence of 21.7% in 2023 exceeding the WHO threshold of 20%. This study aims to identify regional distribution patterns of toddler nutritional status using the Self Organizing Maps (SOM) clustering method. SOM, a neural network technique developed by Teuvo Kohonen, employs unsupervised learning to cluster high dimensional data into a two-dimensional map while preserving data proximity. This study classifies 27 districts/cities in West Java based on several nutritional factors: stunting prevalence, exclusive breastfeeding, low birth weight (LBW), access to proper sanitation, average years of schooling, complete basic immunization, and poverty rate. The data were obtained from the 2023 Indonesian Nutrition Status Survey (SSGI) and Statistics Indonesia (BPS). Clustering performance was evaluated using the Silhouette Coefficient and Dunn Index. The results revealed two main clusters: the first comprises 7 regions with better social and environmental indicators, lower stunting and LBW, and higher exclusive breastfeeding, sanitation, and education; the second cluster comprises 20 regions facing greater nutritional challenges. Internal validation indicates optimal clustering with a Silhouette Coefficient of 0.2538 and a Dunn Index of 0.3639.

Keywords: *Dunn index, Clustering, Nutritional Status.*

Abstrak. Permasalahan gizi balita masih menjadi isu penting di Indonesia, khususnya di Provinsi Jawa Barat yang memiliki prevalensi *stunting* sebesar 21,7% pada tahun 2023, melebihi ambang batas WHO. Penelitian ini bertujuan mengidentifikasi pola sebaran wilayah berdasarkan status gizi balita menggunakan metode klasterisasi *Self Organizing Maps* (SOM), yakni metode pembelajaran tanpa pengawasan pada jaringan syaraf tiruan yang memetakan data berdimensi tinggi ke dalam peta dua dimensi. Sebanyak 27 kabupaten/kota di Jawa Barat dikelompokkan berdasarkan tujuh indikator: prevalensi *stunting*, ASI eksklusif, persentase bayi lahir dengan berat rendah (BBLR), akses sanitasi layak, rata-rata lama sekolah, cakupan imunisasi dasar lengkap, dan persentase penduduk miskin. Data bersumber dari SSGI dan BPS tahun 2023. Evaluasi hasil klaster menggunakan *Silhouette Coefficient* dan *Dunn Index*. Hasil penelitian menunjukkan dua klaster utama: Klaster 1 terdiri dari 7 wilayah dengan kondisi sosial dan lingkungan yang lebih baik, prevalensi *stunting* dan BBLR rendah, serta cakupan ASI, sanitasi, dan pendidikan tinggi; Klaster 2 mencakup 20 wilayah dengan tantangan gizi lebih besar. Validasi internal menghasilkan *Silhouette Coefficient* sebesar 0,2538 dan *Dunn Index* sebesar 0,3639 yang menunjukkan kualitas klasterisasi cukup optimal.

Kata Kunci: *Dunn index, Pengelompokan, Status Gizi.*

A. Pendahuluan

Analisis kluster adalah metode statistik multivariat yang bertujuan untuk mengelompokkan sekumpulan objek berdasarkan kesamaan karakteristik yang dimilikinya. Prinsip utama dari analisis ini adalah untuk memaksimalkan homogenitas di dalam satu kluster (*intra-cluster homogeneity*) dan pada saat yang sama memaksimalkan heterogenitas antar kluster (*inter-cluster heterogeneity*). Dengan kata lain, objek-objek dalam satu kluster akan dibuat semirip mungkin satu sama lain, tetapi akan sangat berbeda dengan objek di kluster lainnya (Hair dkk., 2014).

Dalam analisis kluster, terdapat dua pendekatan utama, yaitu hierarki dan non-hierarki. Metode hierarki, seperti *agglomerative hierarchical clustering* dan *divisive hierarchical clustering*, membentuk kluster berdasarkan kedekatan antar objek dan divisualisasikan dalam dendrogram. Beberapa contoh metode dalam pendekatan ini adalah *ward's method*, *single linkage*, *average linkage*, dan *complete linkage*. Sedangkan dalam metode non-hierarki yang sering digunakan adalah *K-Means* yang lebih efisien dalam pemrosesan data dibandingkan metode hierarki, sehingga lebih sesuai untuk dataset berukuran besar (Kilo dkk., 2025).

Metode klusterisasi seperti *hierarchical clustering* dan *K-Means* memiliki kelebihan dan kekurangan dalam penerapannya. *Hierarchical clustering* dapat membentuk struktur kluster yang jelas melalui dendrogram, namun memiliki keterbatasan dalam efisiensi komputasi, terutama untuk dataset berukuran besar karena membutuhkan perhitungan jarak antar setiap objek (Mulyana dkk., 2024). Sementara itu, *K-Means clustering* lebih cepat dan efisien dalam membagi data ke dalam sejumlah kluster yang telah ditentukan, tetapi metode ini memiliki kelemahan dalam menentukan jumlah kluster yang optimal serta rentan terhadap inisialisasi *centroid* dan *outlier* (Mubarok dkk., 2025).

Salah satu alternatif metode dalam klusterisasi adalah menggunakan jaringan saraf tiruan. Jaringan saraf tiruan (*neural network*) merupakan model komputasi yang meniru mekanisme kerja otak manusia dalam mengenali pola dan memproses informasi. Dengan arsitektur yang terdiri dari neuron buatan yang saling terhubung, *neural network* mampu belajar dari data dan menyesuaikan bobot sinaptiknya untuk meningkatkan kinerja dalam berbagai tugas komputasi, termasuk klusterisasi (Haykin, 2009). Salah satu pendekatan dalam klusterisasi yang berbasis jaringan saraf tiruan adalah *Self Organizing Maps* (SOM) yang diperkenalkan oleh Teuvo Kohonen pada tahun 1982. SOM bekerja dengan mengelompokkan data berdasarkan kemiripan pola dan memetakan data berdimensi tinggi ke dalam ruang dua dimensi. Salah satu keunggulan utama SOM adalah kemampuannya dalam mempertahankan struktur topologi, yaitu hubungan kedekatan antar data dalam ruang dimensi tinggi tetap terjaga ketika direpresentasikan dalam dimensi yang lebih rendah (Gunawan & Rais, 2024).

Arsitektur SOM terdiri dari lapisan input dan output yang saling terhubung melalui bobot, di mana setiap neuron dalam lapisan output mewakili sebuah kelas (*cluster*) yang terbentuk dari pola data yang ada. Berbeda dengan algoritma pembelajaran terawasi (*supervised learning*), SOM tidak memerlukan pengetahuan tentang output target dalam proses pengenalan pola (Hameed dkk., 2019).

Metode SOM dapat dianggap sesuai untuk digunakan dalam analisis kluster karena mampu menangani data yang bersifat multidimensi dan kompleks. Data status gizi balita umumnya mencakup berbagai variabel yang saling berkaitan, seperti faktor kesehatan, sosial, dan ekonomi yang sulit dianalisis secara komprehensif menggunakan metode konvensional. Dalam penelitian ini, penggunaan SOM menjadi relevan karena dapat mempermudah identifikasi pola-pola tersembunyi serta kesamaan kondisi antar wilayah. Kemampuan dalam memvisualisasikan hubungan antar data secara lebih sederhana memungkinkan pengambilan keputusan yang lebih tepat sasaran, khususnya dalam perumusan kebijakan intervensi gizi yang berbasis data dan kebutuhan spesifik tiap daerah.

Status gizi merupakan salah satu indikator utama dalam menilai kualitas kesehatan masyarakat, terutama dalam upaya peningkatan taraf hidup generasi mendatang. Gizi yang baik pada masa balita berperan dalam mendukung pertumbuhan fisik, perkembangan kognitif, serta ketahanan tubuh terhadap penyakit. Namun, masalah gizi seperti *stunting*, gizi buruk, dan kurang gizi masih menjadi tantangan utama di banyak negara berkembang, termasuk Indonesia (Kementerian Kesehatan RI, 2022).

Berdasarkan data (World Health Organization, 2023), secara global terdapat sebanyak 148,1 juta anak di bawah usia lima tahun mengalami *stunting*, 45 juta anak mengalami *wasting*, dan 37 juta anak kelebihan berat badan (*overweight*). Di Indonesia, permasalahan gizi tidak merata terjadi di setiap wilayah, termasuk di Provinsi Jawa Barat. Berdasarkan hasil Survei Status Gizi Indonesia

(SSGI) tahun 2023, prevalensi *stunting* di Jawa Barat mencapai 21,7 persen yang menunjukkan bahwa permasalahan *stunting* masih menjadi isu serius di wilayah ini.

Pengelompokan wilayah berdasarkan faktor yang mempengaruhi status gizi balita dapat membantu memahami perbedaan kondisi secara lebih terstruktur. Faktor-faktor tersebut meliputi prevalensi *stunting*, persentase pemberian ASI secara eksklusif, persentase Berat Bayi Lahir Rendah (BBLR), persentase rumah tangga yang memiliki akses ke sanitasi layak, rata-rata lama sekolah (RRLS), persentase imunisasi dasar lengkap, dan persentase penduduk miskin (UNICEF, 1998). Dengan menggunakan SOM, wilayah dengan kondisi serupa dapat dikelompokkan, sehingga memudahkan analisis dan pengambilan keputusan untuk menanggulangi masalah gizi secara lebih optimal.

Berdasarkan uraian di atas, maka perumusan masalah dalam penelitian ini yaitu: bagaimana penerapan metode *Self Organizing Map* (SOM) untuk mengelompokkan kabupaten/kota di Provinsi Jawa Barat tahun 2023 berdasarkan faktor yang mempengaruhi status gizi balita?. Metode ini diharapkan mampu menghasilkan pengelompokan wilayah yang informatif dan memberikan rekomendasi berbasis data untuk perbaikan gizi balita.

B. Metode

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) dan Kementerian Kesehatan Republik Indonesia. Penelitian ini mencakup 27 kabupaten/kota di Provinsi Jawa Barat pada tahun 2023. Variabel yang digunakan dalam penelitian ini adalah prevalensi *stunting* (X_1), persentase pemberian ASI secara eksklusif (X_2), persentase Berat Bayi Lahir Rendah (BBLR) (X_3), persentase rumah tangga yang memiliki akses ke sanitasi layak (X_4), rata-rata lama sekolah (RRLS) (X_5), persentase imunisasi dasar lengkap (X_6), dan persentase penduduk miskin (X_7).

Proses analisis menggunakan SOM dilakukan melalui beberapa tahapan sebagai berikut:

1. Melakukan standarisasi data agar seluruh variabel memiliki skala yang seragam menggunakan persamaan berikut:

$$z = \frac{x - \mu}{\sigma} \quad \dots(1)$$

Keterangan :

z : nilai data yang telah distandarisasi.

x : nilai asli dari data.

μ : nilai *mean* (rata-rata) dari data

σ : nilai standar deviasi dari data.

2. Menentukan jumlah kluster yang akan dibentuk, dalam penelitian ini k yang digunakan yaitu $k=2$, $k=3$, $k=4$, dan $k=5$.
3. Menentukan bentuk dan ukuran topologi. Penelitian ini menggunakan topologi heksagonal karena bentuk ini mendukung pembaruan bobot yang lebih stabil dan representasi data yang lebih halus. Ukuran topologi yang ditetapkan adalah 9×3 yang menghasilkan 27 neuron. *Learning rate* (α) ditetapkan sebesar 0,1 dan secara bertahap menurun hingga 0,01 untuk menjaga stabilitas pembelajaran. Jumlah iterasi maksimum ditetapkan sebanyak 1000, dan radius ketetanggaan (R) ditetapkan sebesar 1, agar hanya neuron tetangga terdekat dari neuron pemenang (BMU) yang diperbarui.
4. Menginisialisasi bobot awal secara acak untuk setiap neuron dalam jaringan yang ditentukan dengan ukuran matriks sebesar jumlah variabel dikalikan jumlah neuron.
5. Menghitung jarak minimum antara data input terhadap neuron menggunakan jarak Euclidean pada persamaan berikut:

$$d_{ij} = \sqrt{\sum_{i=1}^n (w_{ik} - x_{ik})^2} \quad \dots(3)$$

6. Menghitung jarak antar neuron (d_{hex}) dalam bentuk topologi heksagonal pada persamaan:
- $$d_{hex} = \max(|q_1 - q_2|, |r_1 - r_2|, |s_1 - s_2|) \quad \dots (4)$$

7. Menghitung fungsi kedekatan neuron (h_{ci}) menggunakan persamaan :
- $$h_{ci} = \exp\left(\frac{-d_{hex}^2}{2R^2}\right) \quad \dots (5)$$

Keterangan :

h_{ci} : Nilai fungsi kedekatan antara neuron pemenang c dan neuron tetangga i dalam grid

d_{hex}^2 : Jarak heksagonal antara neuron pemenang c dan neuron tetangga i dalam grid

R : Parameter yang mengontrol luas area pembelajaran di sekitar neuron pemenang

8. Memperbarui nilai bobot neuron pemenang dan neuron tetangganya menggunakan persamaan:
- $$w_{ij,new} = w_{ij,current} + \alpha * h_{ci} (x_{ij} - w_{ij,current}) \quad \dots (6)$$

Keterangan :

$w_{ij,new}$: bobot baru dari neuron baris ke- i dan kolom ke- j

$w_{ij,current}$: bobot lama dari neuron baris ke- i dan kolom ke- j

α : nilai laju pembelajaran (*learning rate*)

x_{ij} : data input pada baris ke- i dan kolom ke- j

9. Memeriksa kondisi konvergensi. Jika iterasi telah mencapai batas maksimum atau perubahan bobot sudah sangat kecil, maka proses pelatihan dihentikan. Jika belum, maka kembali ke langkah 4.
10. Memetakan data ke neuron terbaik (BMU). Data yang dipetakan ke neuron yang sama atau berdekatan dianggap memiliki karakteristik yang mirip.
11. Mengelompokkan neuron ke dalam kluster menggunakan metode *hierarchical clustering* dengan pendekatan Ward. Secara matematis, peningkatan nilai SSE terjadi ketika dua kluster A dan B digabungkan yang dapat dihitung dengan persamaan berikut (Aggarwal & Chandan, 2014).

$$\Delta SSE = \frac{n_A n_B}{n_A + n_B} \|\bar{c}_A - \bar{c}_B\|^2 \quad \dots (7)$$

di mana n_A dan n_B adalah jumlah elemen dalam kluster A dan B, serta $\bar{c}_A$ dan $\bar{c}_B$ merupakan *centroid* masing-masing kluster.

12. Menghitung *centroid* masing-masing kluster, yaitu dengan mencari nilai rata-rata bobot dari neuron-neuron yang termasuk dalam satu kluster menggunakan persamaan:

$$\bar{c}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i \quad \dots (8)$$

di mana x_i menyatakan data anggota ke- i dalam kluster ke- k dan n_k jumlah total anggota kluster ke- k .

13. Menentukan kluster akhir untuk setiap data, dengan cara menghitung jarak Euclidean antara data dengan seluruh *centroid* kluster, lalu memasukkan data ke kluster dengan jarak terdekat.
14. Melakukan evaluasi kualitas hasil klusterisasi. Evaluasi kluster bertujuan untuk mengukur sejauh mana kualitas hasil klusterisasi yang diperoleh. Dalam penelitian ini, evaluasi kluster yang digunakan untuk menilai hasil klusterisasi ada dua, yaitu:
- a) Metode *Dunn Index* merupakan validasi kluster yang menilai seberapa baik pemisahan antar kluster dan kekompakan dalam kluster. *Dunn Index* dapat dihitung berdasarkan persamaan:

$$DI = \frac{\min_{1 \leq k \leq j \leq K} d(C_k, C_j)}{\max_{1 \leq k \leq K} \delta(C_k)} \quad \dots (9)$$

Keterangan:

$d(C_k, C_j)$ = jarak minimum antar titik dalam kluster k dan titik dalam kluster j

$\delta(C_k)$ = jarak maksimum antar dua titik dalam kluster k

- b) *Silhouette coefficient* adalah metrik yang digunakan untuk mengukur seberapa baik suatu titik data dikelompokkan dalam klasternya dibandingkan dengan klaster lain. Dalam memastikan klaster yang dihasilkan memiliki pemisahan yang baik dapat menggunakan persamaan berikut.

$$s(i) = \frac{b(i) - a(i)}{\max((a(i), b(i)))}, i = 1, 2, 3, \dots, N \quad \dots (10)$$

Keterangan

$a(i)$: rata-rata jarak objek ke- i terhadap objek lainnya pada klaster yang sama

$b(i)$: nilai jarak terkecil atau paling minimum

$s(i)$: indeks *silhouette*

Melakukan profilisasi hasil klaster dan membuat kesimpulan.

C. Hasil Penelitian dan Pembahasan

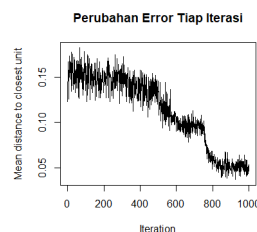
Langkah pertama dalam tahapan analisis adalah melakukan standarisasi data. Proses ini bertujuan agar seluruh variabel memiliki skala yang seragam satu sama lain.

Tabel 1. Hasil Standarisasi Data

Kab/Kota	Prevalensi <i>Stunting</i>	ASI Eklus if	BBL R	Sanitasi layak	RRL S	Pendudu k Miskin	Imunisa si lengkap
Bogor	0.98	-0.65	-0.27	-0.23	-0.35	-0.35	1.43
Sukabumi	1.49	-0.66	1.17	-1.39	-1.00	-0.44	-0.11
Cianjur	-1.26	0.45	-1.23	-0.72	-1.22	0.78	2.01
Bandung	1.00	-0.66	0.81	-0.53	0.26	-0.68	3.33
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Kota Cimahi	-0.70	0.34	-0.22	0.08	1.65	-1.34	0.56
Kota Tasikmala ya	0.48	0.03	0.74	-1.20	0.52	1.28	-1.04
Kota Banjar	-0.13	0.41	1.01	0.89	-0.02	-0.78	-0.81

Pengelompokan Berdasarkan Jumlah Klaster (k)

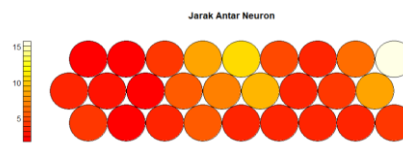
Proses pelatihan SOM selama 1000 iterasi menghasilkan sebuah peta neuron yang terorganisir. Untuk menganalisis kualitas peta yang terbentuk, disajikan dua grafik berikut yang mengilustrasikan proses konvergensi dan struktur klaster yang dihasilkan.



Gambar 1. Perubahan *Error* Iterasi

Plot perubahan *error* pada gambar di atas menunjukkan jarak rata-rata dari vektor data ke neuron pemenangnya (BMU) pada setiap iterasi. Grafik tersebut menunjukkan konvergensi model, di mana nilai *error* menurun tajam pada awal pelatihan dan kemudian stabil setelah iterasi ke-800. Hal ini menandakan bahwa bobot neuron telah berhasil beradaptasi dengan struktur data. Selanjutnya

implementasi SOM menggunakan beberapa kluster sebagai berikut.

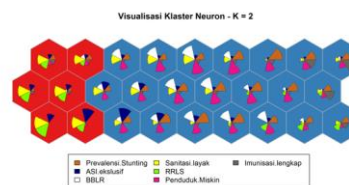


Gambar 2. Visualisasi Jarak antar Neuron

Gambar 2 memvisualisasikan jarak antar neuron tetangga (U-Matrix) untuk mengidentifikasi pola kluster. Pada peta ini, warna gelap (merah) menandakan neuron-neuron yang sangat mirip dan cenderung membentuk inti dari sebuah kluster. Sebaliknya, warna terang (kuning atau putih) mengindikasikan jarak yang jauh antar neuron dan seringkali berfungsi sebagai batas pemisah antar kluster.

Implementasi dengan 2 Kluster ($k = 2$)

Setelah pelatihan SOM selama 1000 iterasi menghasilkan 27 vektor bobot akhir, dilakukan tahap pengelompokan terhadap bobot-bobot tersebut. Visualisasi dari hasil pengelompokan neuron ini disajikan pada gambar berikut.



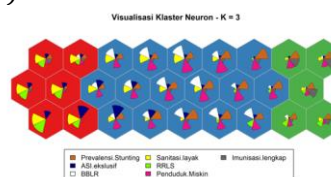
Gambar 3. Hasil Pengelompokan 2 Kluster

Hasil analisis membentuk dua kluster utama. Kluster 1 (merah) merepresentasikan wilayah dengan kondisi sosial-lingkungan yang lebih baik, ditandai oleh tingginya cakupan ASI eksklusif, sanitasi layak, dan Rata-rata Lama Sekolah (RRLS). Sebaliknya, Kluster 2 (biru) adalah kelompok wilayah dengan tantangan gizi dan kesejahteraan yang lebih serius, terutama pada tingginya angka *stunting*, BBLR, dan kemiskinan, sehingga memerlukan intervensi khusus. Berikut merupakan anggota wilayah untuk dua kluster.

Tabel 3. Nilai Jarak Antara Bobot dengan Data Pada 2 Kluster

Kluster	Jumlah Anggota	Anggota Kluster
1	7	Subang, Bekasi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi
2	20	Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Purwakarta, Karawang, Bandung barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Tasikmalaya, Kota Banjar

Implementasi dengan 3 Kluster ($k = 3$)



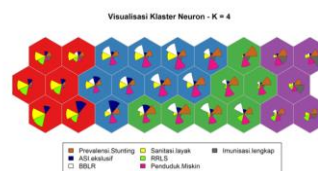
Gambar 4. Hasil Pengelompokan 3 Kluster

Hasil pengelompokan membagi wilayah menjadi tiga klaster dengan profil yang lebih spesifik. Klaster 1 (merah) dapat diidentifikasi sebagai kelompok dengan kondisi sosial lingkungan pendukung terbaik yang unggul dalam indikator ASI eksklusif, sanitasi layak, dan Rata-rata Lama Sekolah (RRLS). Klaster 2 (biru) merupakan kelompok dengan tantangan gizi dan kesejahteraan paling kompleks, ditandai oleh tingginya angka *stunting*, BBLR, dan kemiskinan. Sementara itu, Klaster 3 (hijau) menunjukkan profil unik, di mana tantangan *stunting* dan kemiskinan terjadi bersamaan dengan tingginya cakupan imunisasi dasar, mengindikasikan bahwa program kesehatan tertentu tetap berjalan baik meskipun kondisi sosial-ekonomi masih menjadi tantangan.

Tabel 4. Anggota dan Jumlah Klaster ($k = 3$)

Klaster	Jumlah Anggota	Anggota Klaster
1	6	Subang, Bekasi, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi.
2	14	Sukabumi, Garut, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Purwakarta, Karawang, Bandung Barat, Pangandaran, Kota Sukabumi, Kota Tasikmalaya, Kota Banjar
3	7	Bogor, Cianjur, Bandung, Tasikmalaya, Ciamis, Kota Bogor, Kota Bandung

Implementasi dengan 4 Klaster ($k = 4$)



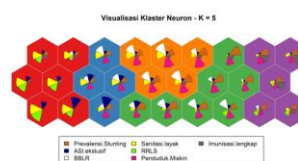
Gambar 5. Hasil Pengelompokan 4 Klaster

Klaster 1 (merah) merepresentasikan wilayah dengan kondisi pendukung terbaik, unggul dalam indikator ASI eksklusif, sanitasi, dan RRLS. Klaster 2 (biru) dan 3 (hijau) merupakan kelompok wilayah dengan tantangan gizi dan kesejahteraan, di mana keduanya memiliki prevalensi *stunting*, BBLR, dan kemiskinan yang tinggi. Sementara itu, Klaster 4 (ungu) menunjukkan tantangan ganda pada *stunting* dan kemiskinan, namun memiliki keunggulan pada cakupan imunisasi dasar yang tinggi.

Tabel 5. Anggota dan Jumlah Klaster ($k = 4$)

Klaster	Jumlah Anggota	Anggota Klaster
1	5	Bekasi, Kota Cirebon, Kota Bekasi, Kota Depok, dan Kota Cimahi.
2	9	Kuningan, Cirebon, Majalengka, Indramayu, Subang, Purwakarta, Karawang, Pangandaran, Kota Banjar.
3	7	Sukabumi, Garut, Tasikmalaya, Sumedang, Bandung Barat, Kota Sukabumi, Kota Tasikmalaya.
4	6	Bogor, Cianjur, Bandung, Ciamis, Kota Bogor, Kota Bandung.

Implementasi dengan 5 Klaster ($k = 5$)



Gambar 6. Hasil Pengelompokan 5 Klaster

Klaster 1 (merah) merepresentasikan wilayah dengan kondisi pendukung terbaik, unggul dalam indikator ASI eksklusif, sanitasi, dan RRLS. Klaster 2 (biru), klaster 3 (oranye), dan klaster 4 (hijau) merupakan kelompok wilayah dengan tantangan gizi dan kesejahteraan, yang ditandai oleh tingginya angka prevalensi *stunting*, BBLR, dan kemiskinan yang tinggi. Sementara itu, Klaster 5 (ungu) menunjukkan tantangan profil unik yang di mana tantangan *stunting* dan kemiskinan yang tinggi terjadi bersamaan dengan cakupan imunisasi dasar yang baik.

Tabel 6. Anggota dan Jumlah Klaster ($k = 5$)

Klaster	Jumlah Anggota	Anggota Klaster
1	5	Bekasi, Kota Cirebon, Kota Bekasi, Kota Depok, dan Kota Cimahi.
2	4	Subang, Purwakarta, Karawang, Kota Banjar
3	7	Sukabumi, Garut, Tasikmalaya, Sumedang, Bandung barat, Kota Sukabumi, Kota Tasikmalaya
4	6	Bogor, Cianjur, Bandung, Ciamis, Kota Bogor, Kota Bandung
5	5	Kuningan, Cirebon, Majalengka, Indramayu, Pangandaran

Evaluasi Hasil Klaster

Untuk menemukan jumlah klaster optimum akan digunakan validasi internal dengan metode Indeks Dunn dan *Silhouette Coefficient*. Jumlah klaster terbaik ditentukan berdasarkan nilai Indeks Dunn tertinggi pada persamaan (9), serta untuk nilai *Silhouette Coefficient* tertinggi pada persamaan (10). Berdasarkan persamaan tersebut diperoleh ukuran validasi internal untuk setiap k yang disajikan pada tabel berikut.

Tabel 7. Nilai Ukuran Validasi untuk Setiap k

Ukuran Validasi	Ukuran Klaster			
	2	3	4	5
Dunn	0.3639	0.1763	0.2770	0.3023
<i>Silhouette</i>	0.2538	0.1936	0.1924	0.1659

Hasil validasi internal pada Tabel 7 menunjukkan bahwa nilai Indeks Dunn yang mendekati 1 terletak pada klaster 2 yaitu sebesar 0.3639. Untuk nilai indeks *Silhouette* tertinggi terletak pada klaster 2 yaitu sebesar 0,2538. Berdasarkan kedua ukuran validasi tersebut, maka jumlah klaster optimal untuk pengelompokan wilayah di Jawa Barat berdasarkan faktor-faktor yang mempengaruhi status gizi balita adalah sebanyak 2 klaster.

Profilisasi Hasil Klaster



Gambar 7. Peta Pengelompokan Kab/Kota di Jawa Barat

Gambar 7 memperlihatkan hasil pengelompokan wilayah kabupaten/kota di Provinsi Jawa Barat ke dalam dua klaster berdasarkan karakteristik status gizinya. Wilayah yang tergabung dalam Klaster 1 (ditampilkan dalam warna hijau) cenderung memiliki kondisi yang lebih baik. Sementara

itu, Klaster 2 (berwarna oranye) didominasi oleh wilayah dengan tantangan yang lebih besar, seperti tingginya prevalensi *stunting* dan keterbatasan akses terhadap layanan kesehatan dasar. Karakteristik masing-masing klaster yang terbentuk dapat dianalisis melalui Tabel 8 yang menyajikan perbandingan dari rata-rata antara Klaster 1 dan Klaster 2.

Tabel 8. Nilai rata-rata klaster

Variabel	Klaster 1	Klaster 2
Angka Prevalensi <i>Stunting</i> (X_1)	14.99	21.71
Persentase Pemberian Asi Secara Eksklusif (X_2)	86.05	70.23
Persentase Berat Bayi Lahir Rendah (BBLR) (X_3)	0.59	3.37
Persentase Rumah Tangga yang Memiliki Akses ke Sanitasi Layak (X_4)	86.08	71.22
Rata-Rata Lama Sekolah (RRLS) (X_5)	10.15	7.99
Persentase Imunisasi Dasar Lengkap (X_6)	5.53	9.10
Persentase Penduduk Miskin (X_7)	2.90	3.99

D. Kesimpulan

Berdasarkan hasil pengelompokan 27 kabupaten/kota di Provinsi Jawa Barat menggunakan metode *Self Organizing Map* (SOM), terbentuk dua klaster utama yang mencerminkan perbedaan kondisi sosial dan kesehatan terkait status gizi balita. Klaster 1 terdiri dari 7 wilayah dengan kondisi yang lebih baik, ditandai oleh rendahnya prevalensi *stunting* dan BBLR, serta tingginya cakupan ASI eksklusif, akses sanitasi layak, dan rata-rata lama sekolah. Sebaliknya, Klaster 2 mencakup 20 wilayah dengan tantangan gizi yang lebih besar, ditandai oleh prevalensi *stunting* dan BBLR yang tinggi, serta rendahnya indikator pendidikan, sanitasi, dan cakupan ASI. Kelompok ini juga memiliki persentase penduduk miskin yang lebih tinggi. Secara umum, pengelompokan ini membedakan wilayah dengan faktor penunjang status gizi yang baik dan wilayah dengan faktor risiko yang lebih tinggi, dengan hasil validasi internal menunjukkan nilai *Dunn Index* sebesar 0,3639 dan *Silhouette Coefficient* sebesar 0,2538.

Ucapan Terimakasih

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada keluarga, sahabat, dosen pembimbing, dan semua pihak atas doa, dukungan, dan bimbingannya selama proses penulisan artikel ini. Semoga segala kebaikan yang diberikan mendapat balasan terbaik dari Allah SWT.

Daftar Pustaka

- Dewi, E. F., & Hajarisman, N. (2023). Penanganan Data Hilang pada Pemodelan Persamaan Terstruktur melalui Metode Full Information Maximum Likelihood (FIML). *DataMath: Journal of Statistics and Mathematics*, 1(1), 11–18. <https://doi.org/10.29313/datamath.v1i1.10>
- Firdaus, A. (2022). Aplikasi Algoritma K-Nearest Neighbor pada Analisis Sentimen Omicron Covid-19. *Jurnal Riset Statistika*, 85–92. <https://doi.org/10.29313/jrs.v2i2.1148>
- Aggarwal, C. C., & Chandan, K. R. (2014). *Data Clustering Algorithms and Applications* (1 ed.). Taylor & Francis Group. <https://doi.org/10.1201/b12990>

- Gunawan, A. R., & Rais, Z. (2024). Applied the Self-Organizing Maps (SOM) Method for Clustering Educational Equity in South Sulawesi. *ARRUS Journal of Mathematics and Applied Science*, 4. <https://doi.org/10.35877/mathscience2607>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7. Auflage, Pearson new internat. ed). Pearson.
- Hameed, A. A., Karlik, B., Salman, M. S., & Eleyan, G. (2019). Robust adaptive learning approach to self-organizing maps. *Knowledge-Based Systems*. <https://doi.org/10.1016/j.knosys.2019.01.011>
- Haykin, S. (2009). *Neural Networks and Learning Machines* (third). Pearson.
- Kilo, N. A., Katili, M. R., & Hasan, I. K. (2025). Perbandingan Metode K-Means dan K-Medoids Dengan Validitas Davies-Bouldin Indeks, Dunn Indeks dan Indeks Connectivity Pada Pengelompokan Masyarakat Penerima Bantuan Langsung Tunai. *Research in the Mathematical and Natural Sciences*, 4, 8–15. <https://doi.org/10.55657/rmns.v4i1.190>
- Kohonen, T. (2001). *Self-Organizing Maps* (Vol. 30). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-56927-2>
- Mubarak, R., Syahputra, A. A., Permana, A. T., Sholiah, L., & Tarwoto. (2025). Implementasi Data Mining untuk Clustering Lowongan Pekerjaan Menggunakan Metode Algoritma K-Means. *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, 9(2), 703–712. <https://doi.org/10.35870/jtik.v9i2.3438>
- Mulyana, A. A. R., Juwita, A. R., Siregar, A. M., & Fauzi, A. (2024). Penerapan Algoritma K-means Clustering dan Hierarchical Clustering dalam Mengelompokkan Data Pengangguran di Karawang. *Jurnal Algoritma*, 21(2). <https://doi.org/DOI:10.33364/algoritma/v.21-2.2155>
- UNICEF. (1998). *The State of The World Children*. Oxford University Press. World Health Organization. (2023). *Levels and trends in child malnutrition*. WHO. <https://iris.who.int/bitstream/handle/10665/368038/9789240073791-eng.pdf?sequence=1>
- Shafira Raudya Putri, & Tasya Aspiranti. (2022). Analisis Pengukuran Kualitas Pelayanan Hotel dengan Menggunakan Metode Kano. *Jurnal Riset Manajemen Dan Bisnis*, 19–24. <https://doi.org/10.29313/jrmb.v2i2.1321>