

TINJAUAN SISTEMATIS PENERAPAN KLASTER K-MEANS DAN METODE ELBOW UNTUK ANALISIS DATA STUNTING

Suwanti, Siti Ganiyah Ramadhani, Dewi Yuliana, Nurfitriah Safrudin, Arif Budiman

Program Studi Informatika, Universitas Muhammadiyah Maumere

Jl. Jenderal Sudirman Waioti Maumere, Indonesia

Suwantis009@gmail.com

ABSTRAK

Stunting merupakan permasalahan gizi kronis yang berdampak jangka panjang terhadap kualitas sumber daya manusia dan masih menjadi tantangan utama di Indonesia. Seiring meningkatnya ketersediaan data kesehatan berskala besar, pendekatan berbasis *machine learning* menjadi alternatif potensial dalam analisis stunting. Penelitian ini bertujuan untuk mengkaji secara sistematis penerapan algoritma K-Means dan metode Elbow dalam analisis data stunting. Metode penelitian menggunakan pendekatan *Systematic Literature Review* (SLR) dengan menelaah artikel ilmiah nasional dan internasional yang diperoleh dari basis data jurnal bereputasi melalui tahapan identifikasi, seleksi, dan sintesis berdasarkan kriteria inklusi dan eksklusi. Hasil kajian terhadap 15 artikel terpilih menunjukkan bahwa algoritma K-Means merupakan metode klasterisasi yang paling dominan digunakan pada berbagai skala wilayah, mulai dari desa hingga provinsi. Metode Elbow dan *silhouette score* banyak dimanfaatkan dan umumnya dikombinasikan dengan teknik validasi internal lain seperti *Silhouette Coefficient* dan *Davies-Bouldin Index*, dengan jumlah klaster optimal. Meskipun demikian, masih ditemukan variasi metodologis, keterbatasan validasi klaster, sempitnya cakupan variabel, serta minimnya integrasi analisis spasial. Studi ini menegaskan perlunya pengembangan metodologi analisis stunting berbasis *machine learning* yang lebih terintegrasi dan relevan untuk mendukung perumusan kebijakan penanggulangan stunting yang berbasis bukti.

Kata kunci : *K-Means Clustering, Machine Learning, Metode Elbow, SLR, Stunting*

1. PENDAHULUAN

Stunting merupakan permasalahan gizi kronis yang ditandai dengan gangguan pertumbuhan pada anak akibat kekurangan asupan gizi dalam jangka panjang. Kondisi ini berdampak pada perkembangan fisik dan kognitif anak serta berpotensi menurunkan kualitas sumber daya manusia di masa depan. Tingginya prevalensi stunting di berbagai wilayah menunjukkan bahwa permasalahan ini masih memerlukan perhatian serius dan penanganan yang berbasis pada analisis data yang akurat dan komprehensif [1], [2].

Dalam upaya memahami dan menanggulangi stunting, tersedia berbagai data yang bersumber dari survei kesehatan, data demografi, serta indikator sosial ekonomi. Namun, kompleksitas dan besarnya volume data tersebut sering kali menyulitkan proses analisis menggunakan metode statistik konvensional. Akibatnya, pola keterkaitan antarvariabel dan karakteristik kelompok tertentu dalam data stunting belum dapat teridentifikasi secara optimal [3].

Seiring dengan perkembangan teknologi informasi, pendekatan data mining mulai banyak diterapkan dalam bidang kesehatan masyarakat, salah satunya melalui teknik klasterisasi. Algoritma K-Means merupakan metode klasterisasi yang paling umum digunakan karena kesederhanaan, efisiensi, serta kemampuannya dalam mengelompokkan data sisa berdasarkan tingkat kemiripan tertentu. Dalam konteks stunting, metode K-Means dimanfaatkan untuk mengelompokkan wilayah atau kelompok

populasi berdasarkan karakteristik dan tingkat prevalensi stunting[4], [5].

Meskipun demikian, salah satu kelemahan utama algoritma K-Means adalah perlunya menentukan jumlah klaster (K) sejak awal. Penentuan nilai K yang tidak tepat dapat menghasilkan klaster yang kurang representatif dan memengaruhi interpretasi hasil analisis. Untuk mengatasi permasalahan tersebut, metode Elbow sering digunakan sebagai pendekatan evaluatif dalam menentukan jumlah klaster optimal dengan menganalisis nilai *within-cluster sum of squares* (WCSS)[6], [7].

Penelitian menggunakan teknik studi literatur sistematis (*Systematic Literature Review*/SLR) sebagai pendekatan utama dalam pengumpulan dan analisis data [8], [9], [10]. SLR dilakukan secara terstruktur dan transparan untuk mengidentifikasi, mengevaluasi, serta mensintesis hasil-hasil penelitian terdahulu yang relevan dengan penerapan algoritma K-Means dan metode Elbow dalam analisis data stunting. Proses SLR meliputi penentuan kata kunci pencarian, pemilihan basis data ilmiah, penetapan kriteria inklusi dan eksklusi artikel, serta analisis tematik terhadap metode, variabel, dan hasil klasterisasi yang digunakan dalam setiap studi. Dengan pendekatan ini, hasil kajian diharapkan memiliki tingkat objektivitas dan replikasi yang lebih tinggi dibandingkan tinjauan literatur naratif, serta mampu memberikan gambaran komprehensif mengenai tren dan celah penelitian di bidang analisis data stunting berbasis klasterisasi.

Berbagai penelitian telah mengombinasikan algoritma K-Means dan metode Elbow dalam analisis data stunting. Namun, hingga saat ini belum banyak kajian yang secara sistematis meninjau penerapan kedua metode tersebut, baik dari sisi pendekatan metodologis, variabel yang digunakan, maupun hasil pengelompokan yang dihasilkan [3], [6], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21].

Penelitian ini bertujuan untuk melakukan tinjauan sistematis terhadap studi-studi yang menerapkan kluster K-Means dan metode Elbow dalam analisis data stunting, sehingga dapat memberikan gambaran menyeluruh serta menjadi rujukan bagi penelitian selanjutnya.

2. TINJAUAN PUSTAKA

2.1. Konsep Dan Defenisi Stunting

Stunting merupakan kondisi gagal tumbuh pada anak balita yang ditandai dengan panjang atau tinggi badan menurut umur (TB/U) berada di bawah minus dua standar deviasi (-2 SD) dari median standar pertumbuhan anak yang ditetapkan oleh World Health Organization (WHO) [22]. Kondisi ini mencerminkan akumulasi kekurangan gizi kronik, paparan infeksi berulang, serta lingkungan tumbuh kembang yang tidak mendukung dalam jangka waktu yang lama. Berbeda dengan wasting yang bersifat akut, stunting bersifat kronis dan sering kali tidak dapat diperbaiki sepenuhnya setelah anak melewati usia emas pertumbuhan (1.000 hari pertama kehidupan)[2], [23].

Dampak stunting tidak hanya terbatas pada hambatan pertumbuhan fisik, tetapi juga berimplikasi luas terhadap perkembangan kognitif, kemampuan belajar, produktivitas ekonomi di masa dewasa, serta peningkatan risiko penyakit tidak menular seperti diabetes dan penyakit kardiovaskular [23]. Oleh karena itu, stunting dipandang sebagai masalah kesehatan masyarakat yang multidimensi dan lintas sektor, yang memerlukan pendekatan intervensi berbasis bukti dan data.

2.2. Faktor Risiko dan Penentuan Stunting

Stunting dipengaruhi oleh berbagai faktor yang saling berinteraksi secara kompleks. Secara umum, determinan stunting dapat menjadi faktor langsung dan tidak langsung. Faktor langsungnya meliputi asupan gizi yang tidak memadai, praktik memberikan makanan bayi dan anak yang tidak tepat, serta penyakit infeksi berulang seperti diare dan infeksi saluran pernapasan [24], [25]. Sementara itu, faktor tidak langsung mencakup kondisi sosial ekonomi keluarga, tingkat pendidikan ibu, akses terhadap layanan kesehatan, sanitasi dan air bersih, serta kondisi lingkungan tempat tinggal [26], [27].

Sejumlah penelitian di Indonesia menunjukkan bahwa rendahnya tingkat pendidikan ibu, pendapatan keluarga yang terbatas, serta buruknya akses terhadap sanitasi layak memberikan kontribusi yang signifikan

terhadap tingginya prevalensi stunting [28], [29]. Di wilayah-wilayah dengan karakteristik geografis yang sulit dijangkau, seperti kawasan perdesaan dan kepulauan, masalah stunting diperparah oleh keterbatasan infrastruktur, distribusi pangan yang tidak merata, serta rendahnya jangkauan layanan kesehatan dasar. Kompleksitas faktor ini menuntut metode yang mampu menangkap analisis pola multidimensi dan heterogenitas antarwilayah.

2.3. Pendekatan Analisis Data dalam Studi Stunting

Pendekatan konvensional dalam analisis stunting umumnya menggunakan metode statistik deskriptif dan inferensial, seperti tabulasi silang, uji korelasi, dan regresi linier atau logistik[30], [31]. Metode ini bermanfaat untuk mengidentifikasi hubungan antara variabel independen dan kejadian stunting, namun memiliki keterbatasan dalam menangkap struktur pola yang kompleks, nonlinier, dan tersembunyi dalam data berukuran besar.

Seiring semakin tersedianya data kesehatan digital dan survei berskala besar, pendekatan berbasis *machine learning* (ML) mulai digunakan sebagai alternatif yang lebih fleksibel [32], [33], [34], [35], [36]. ML memungkinkan eksplorasi data multidimensi tanpa asumsi distribusi tertentu dan mampu mengidentifikasi kelompok atau pola laten yang sulit dijangkau oleh metode statistik tradisional. Dalam konteks ini, teknik unsupervised learning menjadi relevan untuk pemetaan wilayah dan karakteristik kelompok balita berdasarkan kesamaan atribut gizi, sosial ekonomi, dan lingkungan.

2.4. Algoritma Pengelompokan K-Means

K-Means merupakan salah satu algoritma clustering paling populer dalam unsupervised learning yang bertujuan untuk mengelompokkan data ke dalam sejumlah kluster berdasarkan kesamaan jarak antarobjek [31], [37]. Algoritma ini bekerja dengan cara menginisialisasi sejumlah pusat kluster (centroid), kemudian secara iteratif memperbarui posisi centroid berdasarkan rata-rata anggota kluster hingga konvergen [38].

Dalam konteks analisis stunting, K-Means banyak digunakan untuk mengelompokkan wilayah, rumah tangga, atau individu berdasarkan indikator gizi seperti TB/U, berat badan menurut umur (BB/U), berat badan menurut tinggi badan (BB/TB), serta variabel pendukung seperti tingkat pendidikan orang tua, pendapatan keluarga, dan akses sanitasi [39], [40]. Keunggulan K-Means terletak pada kemudahan implementasi, efisiensi komputasi, serta kemampuan menangani dataset berukuran besar. Namun demikian, algoritma ini sensitif terhadap inisialisasi centroid awal dan memerlukan penentuan jumlah kluster (K) secara tepat agar hasil pengelompokan tidak bias.

2.5. Metode Elbow untuk Penentuan Jumlah Optimal

Salah satu tantangan utama dalam penerapan K-Means adalah penentuan jumlah kluster (K) yang optimal. Jika nilai K terlalu kecil, maka data yang heterogen akan tergabung dalam kluster yang terlalu umum (*under-clustering*). Sebaliknya, jika K terlalu besar, maka kluster yang terbentuk menjadi terlalu spesifik dan sulit diinterpretasikan (*over-clustering*) [41].

Metode Elbow digunakan sebagai pendekatan heuristik untuk menentukan nilai K yang paling representatif. Metode ini bekerja dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk berbagai nilai K, kemudian memplotkannya dalam sebuah grafik. Titik “siku” (titik siku) pada kurva menunjukkan nilai K di mana penurunan WCSS mulai melambat secara signifikan. Titik ini dianggap sebagai kompromi terbaik antara kompleksitas model dan kualitas pengelompokan.

Dalam studi-studi analisis stunting, metode Elbow sering digunakan untuk memvalidasi jumlah kluster yang dihasilkan oleh K-Means agar lebih stabil dan mudah diinterpretasikan [42]. Meskipun bersifat subjektif dalam penentuan titik siku, metode ini tetap menjadi salah satu teknik paling umum dan praktis dalam eksplorasi kluster berbasis data kesehatan.

2.6. Studi Terdahulu tentang Penerapan K-means dan Elbow dalam Analisis Stunting

Beberapa penelitian di Indonesia telah menerapkan algoritma K-Means untuk mengelompokkan wilayah berdasarkan tingkat prevalensi stunting dan faktor determinannya [4]. Studi-studi tersebut menunjukkan bahwa K-Means mampu mengidentifikasi kluster wilayah dengan karakteristik risiko tinggi, sedang, dan rendah, sehingga dapat membantu pemerintah daerah dalam menetapkan prioritas intervensi.

Namun, sebagian besar penelitian tersebut belum secara sistematis mengombinasikan K-Means dengan metode Elbow dalam menentukan jumlah kluster yang optimal, atau masih terbatas pada wilayah tertentu seperti provinsi di Pulau Jawa [43]. Selain itu, masih sedikit penelitian yang mengeksplorasi penerapan metode ini pada wilayah dengan heterogenitas geografis dan sosial ekonomi yang tinggi seperti Nusa Tenggara Timur.

Keterbatasan ini menunjukkan adanya ruang penelitian yang relevan untuk dikaji lebih lanjut melalui studi literatur, guna mensintesis temuan-temuan yang ada, membandingkan pendekatan metodologis, serta memberikan kelebihan dan keterbatasan penerapan K-Means dan Elbow dalam analisis stunting berbasis data.

2.7. Kerangka Konseptual Studi Literatur

Berdasarkan uraian di atas, dapat disimpulkan bahwa stunting merupakan masalah multidimensi

yang memerlukan pendekatan analisis data yang mampu menangkap pola kompleks dan heterogenitas wilayah. Algoritma K-Means dan metode Elbow menawarkan kerangka metodologi yang potensial untuk mengelompokkan data stunting secara objektif dan menentukan struktur kluster yang optimal.

Studi literatur ini memposisikan diri sebagai upaya sintesis ilmiah terhadap penelitian-penelitian terdahulu yang menerapkan K-Means dan Elbow dalam konteks analisis stunting. Hasil pengamatan ini diharapkan dapat menjadi dasar konsep bagi penelitian empiris selanjutnya serta mendukung penelitian kebijakan dan intervensi gizi yang lebih tepat sasaran pada wilayah prioritas.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan studi literatur sistematis (*systematic literatur review*) yang bertujuan untuk mengkaji, membandingkan, dan mensintesis hasil-hasil penelitian terdahulu terkait penerapan algoritma K-Means dan metode Elbow dalam analisis stunting. Pendekatan ini dipilih untuk memperoleh pemahaman komprehensif mengenai tren metodologi, keunggulan, keterbatasan, serta ruang penelitian yang masih terbuka dalam analisis bidang stunting berbasis machine learning.

3.1. Penerapan Studi Literatur Sistematis (SLR)

Penelitian ini menerapkan pendekatan studi literatur sistematis (*systematic literatur review*) sebagai metode utama untuk mengkaji penerapan algoritma K-Means dan metode Elbow dalam analisis stunting. Pendekatan ini dilakukan melalui tahapan terstruktur yang meliputi penelitian pertanyaan penelitian, pencarian literatur berdasarkan data ilmiah bereputasi, seleksi artikel berdasarkan kriteria inklusi dan eksklusi, ekstraksi data, serta sintesis dan analisis hasil penelitian. Literatur dikumpulkan dari Google Scholar, Garuda, SINTA, dan Scopus dengan menggunakan kombinasi kata kunci terkait stunting, K-Means clustering, metode Elbow, dan *machine learning* kesehatan. Artikel yang dipilih dianalisis secara sistematis untuk mengidentifikasi tren metodologi, variasi pendekatan, serta temuan utama penelitian. Pendekatan ini dipilih untuk menghasilkan sintesis yang komprehensif, objektif, dan representatif terhadap perkembangan terkini analisis stunting berbasis *machine learning*.

3.2. Identifikasi Masalah dan Solusi

Berdasarkan hasil kajian literatur, dilakukan identifikasi terhadap permasalahan metodologi utama dalam penelitian analisis stunting, khususnya terkait keterbatasan pendekatan statistik konvensional, inkonsistensi penentuan jumlah kluster, serta minimnya validasi hasil clustering. Permasalahan tersebut dijelaskan melalui akurasi sistematis terhadap penelitian terdahulu yang disajikan dalam bentuk tabel komparasi, yang memuat informasi mengenai penulis & tahun, judul, objek, metode,

serta hasil utama dan keterbatasan masing-masing penelitian. Melalui pembahasan tabel tersebut, diidentifikasi pola umum, keunggulan pendekatan yang telah digunakan, serta celah penelitian yang masih terbuka. Berdasarkan temuan tersebut, solusi strategi yang dirumuskan berupa penguatan validasi klaster, perluasan analisis variabel, serta integrasi teknik tambahan untuk meningkatkan kualitas pengelompokan data stunting dan relevansi hasil bagi pengambilan keputusan.

3.3. Metodologi dan Implikasi Kebijakan Usulan

Sebagai pengayaan metodologis, penelitian ini mengusulkan integrasi metode Elbow dengan teknik validasi tambahan seperti *Silhouette Coefisien* untuk memperoleh struktur klaster yang lebih stabil dan objektif, serta pemanfaatan *Principal Component Analysis* (PCA) guna mereduksi dimensi data dan meningkatkan kualitas clustering. Selain itu, analisis integrasi spasial berbasis sistem informasi geografis dimaksudkan untuk memetakan distribusi klaster stunting secara visual dan mengidentifikasi prioritas intervensi wilayah. Metodologi usulan ini diharapkan dapat menjadi dasar bagi penelitian lanjutan yang lebih komprehensif dan kuat dalam analisis stunting berbasis data. Dari sisi kebijakan, hasil kajian ini memberikan pentingnya penyusunan strategi penanggulangan stunting yang lebih tepat sasaran, berbasis bukti, serta adaptif terhadap heterogenitas wilayah, sehingga dapat mendukung percepatan pencapaian target penurunan stunting nasional.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penerapan Studi Literatur Sistematis (SLR)

Tabel 1 menyajikan ringkasan hasil pencarian awal dari masing-masing database sebelum penghapusan duplikat.

Tabel 1. Hasil Pencarian Awal pada Tahap Identification

Database	Jumlah Rekord Ditemukan	Catatan
PubMed	10	Fokus pada studi kesehatan dan nutrisi balita
Scopus	10	Artikel peer-reviewed dengan kualitas tinggi
Google Scholar	15	Termasuk tesis, prosiding konferensi, dan artikel Indonesia
Portal Garuda	10	Literatur nasional terkait stunting di Indonesia
Total	50	Rekord sebelum penghapusan duplikat

Tabel 1 menyajikan ringkasan hasil pencarian awal pada tahap identifikasi dari masing-masing basis data sebelum dilakukan penghapusan duplikat. Dari total 50 rekaman awal yang diperoleh, artikel

selanjutnya diseleksi secara bertahap melalui penyaringan judul dan abstrak untuk mengidentifikasi kesesuaian topik dengan fokus penelitian. Sejumlah artikel dieliminasi karena tidak relevan dengan konteks stunting atau tidak menerapkan metode K-Means dan teknik penentuan jumlah klaster. Tahap telaah teks lengkap selanjutnya menghasilkan 15 artikel terpilih yang memenuhi seluruh kriteria inklusi, yaitu membahas analisis stunting, menerapkan algoritma K-Means, serta menjelaskan teknik penentuan jumlah klaster, khususnya metode Elbow. Kelima belas artikel inilah yang selanjutnya dianalisis secara komparatif pada tahap berikutnya. Artikel-artikel terpilih berasal dari berbagai wilayah studi dan rentang tahun publikasi yang beragam, sehingga merepresentasikan perkembangan metodologi analisis stunting berbasis machine learning dalam beberapa tahun terakhir.

Hasil sintesis menunjukkan bahwa penggunaan algoritma K-Means dalam analisis stunting cenderung meningkat seiring berkembangnya sistem informasi kesehatan digital dan kebutuhan pemetaan risiko berbasis wilayah. Meskipun demikian, variasi pendekatan metodologis masih cukup tinggi, baik dari sisi pemilihan variabel, jumlah klaster yang digunakan, maupun teknik validasi hasil clustering. Metode Elbow telah diterapkan pada sebagian penelitian sebagai dasar penentuan jumlah klaster, namun penggunaannya belum konsisten dan sering kali tidak disertai teknik validasi tambahan untuk menguji stabilitas klaster. Temuan ini mengindikasikan bahwa meskipun pendekatan clustering telah memberikan kontribusi penting dalam pemetaan risiko stunting, masih terdapat celah metodologis yang perlu diperkuat melalui standarisasi prosedur, integrasi teknik validasi tambahan, serta perluasan cakupan variabel analisis. Dengan demikian, hasil penerapan SLR ini menegaskan pentingnya pengembangan metodologi yang lebih robust dan terintegrasi guna meningkatkan akurasi serta kegunaan praktis analisis stunting berbasis data.

4.2. Hasil Identifikasi Masalah dan Solusi

Bagian ini menyajikan hasil perbandingan penelitian terdahulu yang relevan dengan penerapan algoritma K-Means dan metode Elbow dalam analisis stunting. Sebanyak lima belas artikel terpilih dianalisis secara komprehensif untuk mengidentifikasi pola penggunaan metode clustering, teknik penentuan jumlah klaster, cakupan wilayah penelitian, serta temuan utama dan keterbatasan masing-masing studi. Analisis perbandingan ini bertujuan untuk menggambarkan perkembangan metodologi terkini, mengidentifikasi konsistensi dan variasi pendekatan yang digunakan, serta menelaah celah penelitian yang masih terbuka. Hasil perbandingan disajikan pada Tabel 2 berikut.

Tabel 2. Perbandingan *State of the Art*

No	Penulis & Tahun	Judul (Ringkas)	Objek	Metode	Hasil Utama	Keterbatasan
1	Ibanez et al., 2024 (7)	PCA-K-Means Klasterisasi Stunting	81 Desa Dompu, NTB	PCA + K-Means; Elbow & Silhouette	2 klaster; Silhouette 0.53; variansi 72.9%	Skala kabupaten, tanpa MBG
2	Robbani et al., 2024 (8)	Clustering Stunting Tangerang (K-Means vs Others)	Anak Stunting Tangerang	K-Means; Silhouette; Hierarchical/DBSCAN	3 klaster; Silhouette 0.52	Skala kabupaten, tanpa integrasi program
3	Ramadhani et al., 2025 (14)	K-Means Spasial Stunting Sumut	Stunting Sumut (faktor lingkungan)	K-Means; Elbow & Silhouette	4 klaster; Silhouette 0.65	Fokus lingkungan saja
4	Putri et al., 2025 (15)	Hierarchical Clustering Determinan Stunting Jatim	Data SKI Jatim	PCA; Hierarchical; Silhouette	2 klaster; Silhouette 0.6155	Skala provinsi, tanpa K-Means
5	Ridha et al., 2025 (16)	Fuzzy C-Means Stunting Aceh Utara	32 Kecamatan Aceh Utara	Fuzzy C-Means; DBI	3 klaster; DBI 0.8281	Soft clustering, tanpa PCA
6	Irma Darmayanti et al., 2025	K-Means With Elbow for Stunting Risk in Toddlers	20 balita Purwekorto Selatan	K-Means; Elbow; Silhouette; PCA	2 klaster; Silhouette 0.252; Cluster 1 higher risk (50% potential stunting); PCA variansi 75.5%	Ukuran sampel kecil (N=20); bias recall; tanpa faktor lain (SES, infeksi)
7	Fanny Ramadhani et al., 2025	Spatial Clustering Stunting Sumut Berbasis Faktor Lingkungan	Kejadian stunting Sumatera Utara	K-Means Spasial; Elbow & Silhouette	4 klaster; Silhouette 0,65	Fokus faktor lingkungan, belum integrasi program nasional
8	Efmi Maiyana et al., 2025	Optimization Stunting Clustering with K-Means ++ & DBI	500 sampel resiko stunting bayi Indonesia (usia, gender, BB/TB lahir & saat ini)	K-Means++; DBI (uji K=2-10); preprocessing (imputasi, normalisasi, outlier)	5 klaster optimal; DBI 0,8379	Validasi hanya DBI; data sekunder
9	Adwin Falah Hasan Marsandy et al., 2024	Klasterisasi Prevalensi Stunting Menggunakan K-Prototype pada Data Campuran	34 Provinsi di Indonesia (2022)	K-Prototype; Elbow (WCSS)	4 klaster; WCSS optimal 65,83	Skala nasional, belum integrasi program kebijakan
10	Masita Handayani et al., 2024	Klasterisasi Prefalensi Stunting dengan K-Prototype	34 Provinsi di Indonesia 2022	Prototype; Siku (WCSS)	Terbentuk 4 klaster optimal; WCSS 65,83; kombinasi variabel lumerik dan kategorik Terbentuk 4 klaster optimal; WCSS 65,83; kombinasi variabel lumerik dan kategorik	Skala nasional; belum mempertimbangkan PCA dan validasi internal selain elbow
11	Rudi Rahmat Saputra et al., 2024	Klasterisasi Balita Stunting Kabupaten Ciamis	27 kecamatan di ciamis	K-Means; KDD; visualisasi	3 klaster rendah (<10%), sedang (10-20%), tinggi (>20%), hasil divisualisasikan	Variabel terbatas pada prevalensi; tanpa validasi Silhouette/PCA

No	Penulis & Tahun	Judul (Ringkas)	Objek	Metode	Hasil Utama	Keterbatasan
12	Indra et al., 2023	Perbandingan K-Means & Hierarchical Clustering Daerah Stunting	34 Provinsi Indonesia	K-Means & Hierarchical (Agglomerative, average linkage); Silhouette, Calinski-Harabasz, Davies-Bouldin	K-Means: 2 klaster, Silhouette 0.48, Calinski 10.49, DBI 0.54; Hierarchical: 3 klaster, Silhouette 0.47, Calinski 9.54, DBI 0.46	Skala provinsi nasional, data tahun 2021, tanpa data kabupaten lebih detail
13	Muhammad Ridha et al., 2025	FCM Clustering Stunting Vulnerability Aceh Utara	32 Kecamatan Aceh Utara (balita E-PPGBM, entry 2023, % stunting, wasting, underweight)	Fuzzy C-Means (fuzziness 2, max iter 100, tol 0.0001); Evaluasi DBI; Visualisasi scatter plot; Perbandingan K-Means	3 klaster: rendah (13 kecamatan), sedang (6), tinggi (13); DBI FCM 0.828 (lebih baik dari K-Means 0.979)	Data agregat 2023 terbatas pada indikator gizi dasar, tanpa faktor sosial-ekonomi/lain; Belum GIS; Potensi outdated
14	Wargijino Utomo 2021	Comparison K-Means & K-Medoids Clustering COVID-19 Spread Indonesia	4 Provinsi Indonesia (confirmed & death cases COVID-19, data Jan 2021)	K-Means & K-Medoids; Davies-Bouldin Index (K2-K9); RapidMiner	K-Means: optimal K=5, DBI 0.064; K-Medoids: optimal K=2, DBI 0.411; K-Means lebih baik	Hanya confirmed & death (bukan treatment/recovery); Skala nasional, data 2021 terbatas; Belum variabel lain seperti faktor sosial-ekonomi
15	Taufik Hidayat et al., 2023	K-Means Clustering Stunting-Prone Areas Jawa Barat	27 Kabupaten/Kota Jawa Barat (kejadian stunting & faktor risiko 2019-2021)	K-Means + CRISP-DM; Elbow Method; Silhouette Coefficient	3 klaster; Silhouette 0.61 (layak)	Data historis 2019-2021 terbatas pada faktor dasar, skala provinsi tanpa detail lebih mikro; Belum faktor eksternal seperti sosial-ekonomi

Tabel 2 perbandingan penelitian terdahulu menunjukkan bahwa pendekatan clustering telah banyak digunakan dalam analisis stunting di berbagai wilayah Indonesia, baik pada skala desa, kecamatan, kabupaten, maupun provinsi. Algoritma K-Means merupakan metode yang paling dominan digunakan, baik sebagai metode utama maupun sebagai pembanding dengan teknik clustering lain seperti hierarchical clustering, fuzzy c-means, k-prototype, dan k-medoids. Dominasi penggunaan K-Means menunjukkan bahwa algoritma ini dipandang efektif, sederhana, serta efisien untuk mengelompokkan wilayah atau individu berdasarkan karakteristik risiko stunting yang bersifat multidimensi.

Dari sisi penentuan jumlah klaster, sebagian besar penelitian menggunakan metode Elbow, Silhouette Coefficient, Davies-Bouldin Index, atau kombinasi di antaranya sebagai teknik validasi internal. Beberapa penelitian, seperti Ibanez et al. (2024), Ramadhani et al. (2025), dan Taufik Hidayat et al. (2023), mengombinasikan metode Elbow dan Silhouette untuk memperoleh struktur klaster yang lebih stabil dan representatif. Hasil-hasil tersebut menunjukkan bahwa jumlah klaster optimal umumnya berada pada rentang dua hingga lima klaster, dengan nilai Silhouette berkisar antara 0,48

hingga 0,65, yang mengindikasikan kualitas pengelompokan berada pada kategori sedang hingga baik. Namun demikian, masih ditemukan variasi signifikan dalam pemilihan nilai K dan teknik validasi, yang menunjukkan belum adanya standar baku dalam penentuan struktur klaster pada analisis stunting.

Dari perspektif cakupan wilayah dan skala data, sebagian besar penelitian masih berfokus pada skala kabupaten dan provinsi, seperti Jawa Barat, Jawa Tengah, Sumatera Utara, dan Aceh Utara. Penelitian pada skala mikro, seperti tingkat desa atau individu balita, masih relatif terbatas dan umumnya menggunakan ukuran sampel kecil. Kondisi ini berimplikasi pada keterbatasan generalisasi hasil serta potensi bias dalam interpretasi klaster risiko. Selain itu, sebagian penelitian hanya menggunakan indikator gizi dasar atau prevalensi stunting sebagai variabel utama, tanpa mengintegrasikan faktor sosial ekonomi, sanitasi, pendidikan ibu, maupun variabel lingkungan secara komprehensif.

Analisis terhadap metode tambahan menunjukkan bahwa beberapa penelitian telah mulai mengintegrasikan teknik reduksi dimensi seperti Principal Component Analysis (PCA) serta pendekatan clustering alternatif seperti fuzzy c-means

dan k-prototype untuk menangani data campuran. Penggunaan PCA terbukti mampu meningkatkan proporsi variansi yang menjelaskan hingga lebih dari 70% dan memperbaiki kualitas kluster. Namun, penerapan teknik ini masih terbatas pada sebagian kecil penelitian dan belum menjadi praktik umum dalam analisis stunting berbasis clustering.

Dari sisi keterbatasan metodologis, sebagian besar penelitian melaporkan keterbatasan pada ukuran sampel, cakupan variabel, ketiadaan validasi eksternal, serta belum adanya integrasi dengan program atau kebijakan penanggulangan stunting. Beberapa penelitian juga belum memanfaatkan analisis spasial berbasis sistem informasi geografis, sehingga pola distribusi risiko stunting secara geografis belum sepenuhnya tergambarkan. Temuan ini menunjukkan bahwa meskipun pendekatan clustering telah memberikan kontribusi signifikan dalam pemetaan risiko stunting, masih terdapat celah metodologis yang perlu diperkuat melalui integrasi teknik validasi yang lebih komprehensif, perluasan variabel multidimensi, serta pemanfaatan analisis spasial dan kerangka kebijakan.

Secara keseluruhan, hasil perbandingan penelitian terdahulu menegaskan bahwa kombinasi algoritma K-Means dan metode Elbow memiliki potensi besar dalam mendukung analisis stunting berbasis data. Namun, diperlukan pengembangan metodologi yang lebih robust melalui integrasi teknik validasi tambahan, reduksi dimensi, serta analisis spasial agar hasil clustering lebih stabil, interpretatif, dan relevan sebagai dasar perumusan kebijakan. Temuan ini sekaligus memperkuat urgensi pengusulan pendekatan metodologis terintegrasi pada penelitian selanjutnya sebagaimana diuraikan pada bagian berikutnya.

4.3. Usulan Metodologi dan Implikasi Kebijakan

Berdasarkan hasil sintesis dan perbandingan penelitian terdahulu, dapat disimpulkan bahwa penerapan algoritma K-Means dan metode Elbow telah memberikan kontribusi penting dalam pemetaan risiko stunting berbasis data. Namun, variasi pendekatan metodologis, keterbatasan validasi kluster, serta sempitnya cakupan variabel dan skala analisis menunjukkan perlunya pengembangan metodologi yang lebih terintegrasi dan robust. Oleh karena itu, penelitian ini mengusulkan kerangka metodologi lanjutan yang mengombinasikan K-Means dengan teknik validasi multi-metrik, reduksi dimensi, serta analisis spasial untuk meningkatkan stabilitas, interpretabilitas, dan relevansi kebijakan dari hasil clustering.

Secara metodologis, integrasi metode Elbow dengan teknik validasi tambahan seperti *Silhouette Coefficient* dan *Davies-Bouldin Index* direkomendasikan sebagai prosedur standar dalam penentuan jumlah kluster optimal. Penggunaan multi-

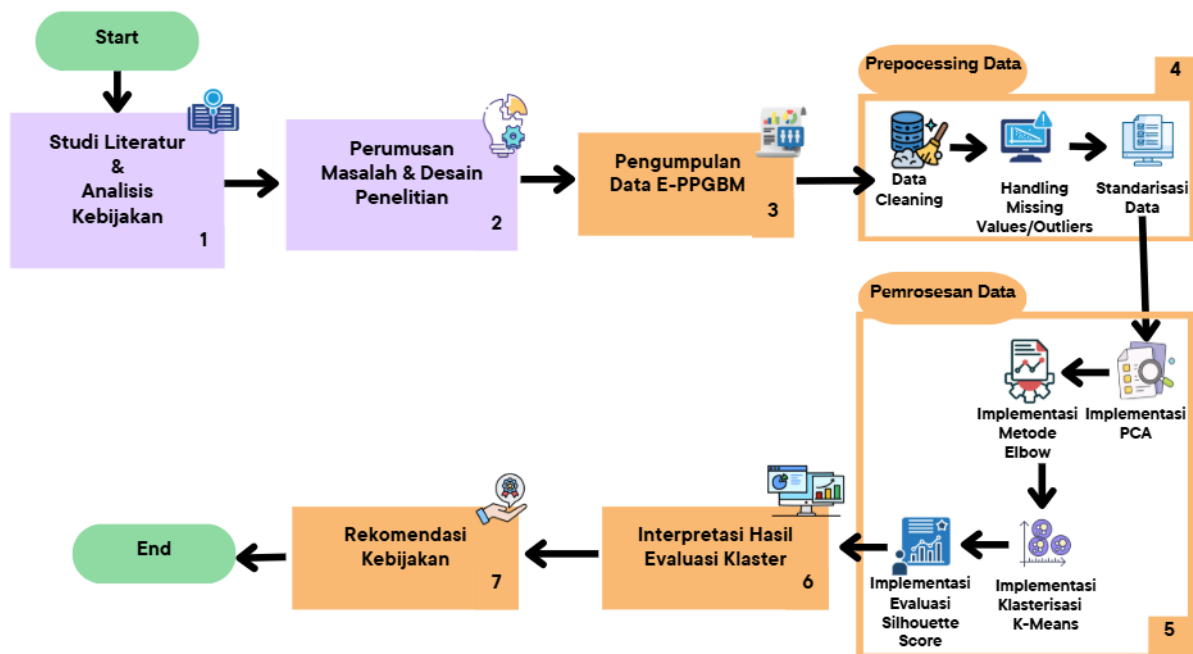
kriteria validasi ini memungkinkan evaluasi kualitas kluster yang lebih komprehensif, sehingga mengurangi risiko over-clustering maupun under-clustering yang sering terjadi pada penelitian sebelumnya. Selain itu, penerapan *Principal Component Analysis* (PCA) sebagai tahap pra-pemrosesan data diusulkan untuk menangani multikolinearitas antarvariabel dan meningkatkan proporsi variansi yang menjelaskan, sebagaimana ditunjukkan pada beberapa penelitian yang memperoleh kualitas kluster lebih baik setelah reduksi dimensi.

Pengembangan metodologi selanjutnya juga diarahkan pada integrasi analisis spasial berbasis sistem informasi geografis. Pendekatan clustering spasial memungkinkan identifikasi pola distribusi risiko stunting secara geografis yang lebih akurat, khususnya pada wilayah dengan heterogenitas tinggi seperti provinsi kepulauan dan daerah tertinggal. Dengan mengombinasikan atribut gizi, sosial ekonomi, dan faktor lingkungan dengan informasi spasial, hasil clustering diharapkan mampu merepresentasikan kondisi riil lapangan secara lebih komprehensif serta mendukung perencanaan intervensi berbasis wilayah prioritas.

Dari sisi variabel analisis, penelitian lanjutan disarankan untuk memperluas cakupan indikator yang tidak hanya terbatas pada prevalensi stunting dan antropometri, tetapi juga mencakup determinan sosial, sanitasi, pendidikan ibu, akses layanan kesehatan, serta program intervensi pemerintah. Integrasi variabel multidimensi ini penting untuk menghasilkan kluster risiko yang lebih interpretatif dan operasional, sehingga mampu mengidentifikasi kelompok wilayah atau individu yang memerlukan intervensi spesifik secara lebih tepat sasaran.

Implikasi kebijakan dari pengembangan metodologi ini bersifat strategis dalam mendukung percepatan penurunan stunting nasional. Hasil clustering yang lebih stabil dan terstandarisasi dapat dimanfaatkan sebagai dasar pemetaan wilayah prioritas, penentuan sasaran intervensi, serta evaluasi efektivitas program penanggulangan stunting. Pemerintah daerah dan pemangku kepentingan dapat menggunakan informasi kluster risiko sebagai alat bantu pengambilan keputusan dalam alokasi sumber daya, perancangan program lintas sektor, serta pemantauan capaian pembangunan gizi secara berkelanjutan.

Secara keseluruhan, usulan metodologi terintegrasi ini diharapkan tidak hanya memperkuat kualitas analisis stunting berbasis machine learning, tetapi juga meningkatkan keterhubungan antara hasil penelitian dan kebutuhan kebijakan publik. Dengan demikian, pendekatan ini berpotensi menjadi kerangka metodologis standar bagi penelitian dan implementasi sistem pemantauan stunting berbasis data di Indonesia pada masa mendatang.



Gambar 1. Diagram Alir Usulan Metode Penelitian

Berdasarkan hasil sintesis dan perbandingan penelitian terdahulu, dapat disimpulkan bahwa penerapan algoritma K-Means dan metode Elbow memiliki potensi besar dalam pemetaan risiko stunting berbasis data. Namun demikian, masih ditemukan berbagai keterbatasan, seperti variasi tahapan pra-pemrosesan data, penggunaan satu metode validasi kluster, serta keterbatasan integrasi hasil analisis dengan kebutuhan kebijakan. Oleh karena itu, penelitian ini mengusulkan suatu kerangka metodologi terintegrasi yang disusun secara sistematis sebagaimana ditunjukkan pada Gambar 1, dengan tujuan meningkatkan kualitas analisis, stabilitas kluster, serta relevansi hasil bagi perumusan kebijakan penanggulangan stunting.

Tahap awal metodologi dimulai dengan studi literatur dan analisis kebijakan, yang bertujuan mengidentifikasi indikator stunting, variabel determinan utama, serta kerangka kebijakan nasional dan daerah yang relevan. Tahap ini menjadi dasar dalam perumusan masalah dan desain penelitian, termasuk penentuan tujuan analisis clustering serta pemilihan variabel yang sesuai dengan konteks kebijakan penanggulangan stunting. Selanjutnya, dilakukan pengumpulan data E-PPGBM sebagai sumber data utama, mengingat sistem ini menyediakan informasi antropometri balita serta indikator pendukung yang bersifat rutin dan terstandarisasi.

Pada tahap pra-pemrosesan data, dilakukan serangkaian proses yang mencakup pembersihan data (data cleaning), penanganan data hilang dan outlier, serta standarisasi data. Tahapan ini krusial untuk memastikan kualitas data sebelum analisis clustering dilakukan. Selain itu, *Principal Component Analysis* (PCA) diusulkan sebagai bagian dari pra-pemrosesan

untuk mereduksi dimensi data, mengatasi multikolinearitas antarvariabel, serta meningkatkan proporsi variansi yang dapat dijelaskan, sehingga hasil clustering menjadi lebih stabil dan mudah diinterpretasikan.

Tahap berikutnya adalah proses clustering, yang diawali dengan penentuan jumlah kluster optimal menggunakan metode Elbow. Untuk meningkatkan validitas hasil, metode Elbow direkomendasikan untuk dikombinasikan dengan metrik evaluasi tambahan, seperti *Silhouette Coefficient* dan *Davies-Bouldin Index*, sehingga penentuan jumlah kluster tidak hanya bergantung pada satu kriteria. Setelah jumlah kluster ditetapkan, algoritma K-Means diimplementasikan untuk mengelompokkan data berdasarkan kemiripan karakteristik gizi, sosial ekonomi, dan lingkungan.

Hasil clustering selanjutnya masuk pada tahap interpretasi dan evaluasi kluster, di mana setiap kluster dianalisis karakteristik dominannya serta tingkat risiko stunting yang diwakilinya. Tahap ini menjadi jembatan penting antara analisis teknis dan pemanfaatan hasil secara praktis. Kluster dengan risiko tinggi dapat diidentifikasi sebagai wilayah atau kelompok prioritas intervensi, sementara kluster dengan risiko rendah dapat dijadikan acuan praktik baik (best practice).

Tahap akhir dari kerangka metodologi ini adalah perumusan rekomendasi kebijakan. Hasil clustering yang telah dievaluasi dan diinterpretasikan digunakan sebagai dasar penyusunan rekomendasi kebijakan yang lebih terarah dan berbasis bukti. Pemerintah daerah dan pemangku kepentingan dapat memanfaatkan informasi kluster risiko untuk menentukan prioritas wilayah, mengalokasikan sumber daya secara lebih efisien, merancang program

intervensi lintas sektor, serta melakukan pemantauan dan evaluasi program penanggulangan stunting secara berkelanjutan.

Secara keseluruhan, usulan metodologi terintegrasi ini tidak hanya memperkuat aspek teknis analisis stunting berbasis machine learning, tetapi juga memastikan keterhubungan yang jelas antara hasil analisis data dan kebutuhan kebijakan publik. Dengan mengikuti alur yang sistematis dari pengolahan data hingga rekomendasi kebijakan, pendekatan ini berpotensi menjadi kerangka metodologis yang aplikatif dan berkelanjutan bagi pengembangan sistem pemantauan dan penanggulangan stunting berbasis data di Indonesia.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil studi literatur sistematis terhadap 15 artikel terpilih, dapat disimpulkan bahwa penerapan algoritma K-Means yang dikombinasikan dengan PCA dan evaluasi metode Elbow dan silhouette score memiliki potensi yang signifikan dalam analisis dan pemetaan risiko stunting berbasis data pada berbagai skala wilayah. Metode ini terbukti mampu mengelompokkan wilayah atau individu berdasarkan karakteristik gizi, sosial ekonomi, dan lingkungan dengan kualitas kluster yang berada pada kategori sedang hingga baik, ditunjukkan oleh jumlah kluster optimal yang umumnya berada pada rentang dua hingga lima kluster serta nilai validasi internal yang memadai. Namun demikian, hasil kajian juga mengungkap adanya variasi pendekatan metodologis, keterbatasan penggunaan teknik validasi kluster yang komprehensif, sempitnya cakupan variabel, serta minimnya integrasi analisis spasial dan keterkaitan dengan kebijakan penanggulangan stunting. Oleh karena itu, penelitian selanjutnya disarankan untuk mengembangkan metodologi yang lebih terintegrasi melalui penerapan validasi kluster multi-metrik, pemanfaatan teknik reduksi dimensi, perluasan variabel multidimensi, serta integrasi analisis spasial agar hasil clustering lebih stabil, interpretatif, dan relevan sebagai dasar perumusan kebijakan penanggulangan stunting berbasis bukti.

DAFTAR PUSTAKA

- [1] J. Fauziah, K. D. Trisnawati, K. P. S. Rini, and S. U. Putri, "Stunting: Penyebab, Gejala, dan Pencegahan," *Jurnal Parenting dan Anak*, vol. 1, no. 2, pp. 1–11, Dec. 2024, doi: 10.47134/jpa.v1i2.220.
- [2] Samsuddin *et al.*, *Stunting*. Purbalingga: Penerbit Cv.Eureka Media Aksara, 2023.
- [3] R. Afifa *et al.*, "Implementasi Principal Component Analysis (PCA) dan Gap Statistic untuk Clustering Kanker Payudara pada Algoritma K-Means," vol. 13, no. 5, pp. 1852–1864, 2024, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [4] C. Syalom, U. Khaira, and D. Arsa, "Clustering Daerah Rawan Stunting Di Indonesia Menggunakan Algoritma K-Means," *Jurnal Teknologi Informasi*, vol. 6, no. 1, 2025, doi: 10.46576/djtechno.
- [5] N. Nur, M. Iqam, and N. Inayah, "Perbandingan K-Means dan Hierarchical Clustering dalam Pengelompokan Daerah Beresiko Stunting," vol. 8, no. 2, p. 2023.
- [6] R. Saputra and I. Purnama, "Optimizing K-Means Algorithm With Elbow And Silhouette Methods For National Exam Score Data Clustering," 2024.
- [7] I. Darmayanti, D. I. S. Saputra, A. B. Wijaya, A. Wijanarko, D. Fortuna, and A. F. Putranto, "K-Means Clustering with Elbow Method for Stunting Risk Detection in Toddlers Using Anthropometric and Nutritional Data," *Journal of Information Systems and Informatics*, vol. 7, no. 4, pp. 3735–3748, Dec. 2025, doi: 10.63158/journalisi.v7i4.1337.
- [8] I. Larasati, A. N. Yusril, and P. Al Zukri, "Systematic Literature Review Analisis Metode Agile Dalam Pengembangan Aplikasi Mobile," *Sistemasi*, vol. 10, no. 2, p. 369, 2021, doi: 10.32520/stmsi.v10i2.1237.
- [9] R. Belen-Saglam, E. Altuncu, Y. Lu, and S. Li, "A systematic literature review of the tension between the GDPR and public blockchain systems," *Blockchain: Research and Applications*, vol. 4, p. 100129, 2023, doi: 10.1016/j.bcr.2023.100129.
- [10] E. Indrisari, H. Febiansyah, and B. Adiwino, "A Systematic Literature Review on the Application of Machine Learning for Predicting Stunting Prevalence in Indonesia (2020–2024)," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 14, no. 3, pp. 277–283, Jul. 2025, doi: 10.32736/sisfokom.v14i3.2366.
- [11] Taufik Hidayat, Mohamad Jajuli, and Susilawati, "Clustering daerah rawan stunting di Jawa Barat menggunakan algoritma K-Means," *INFOTECH: Jurnal Informatika & Teknologi*, vol. 4, no. 2, pp. 137–146, Dec. 2023, doi: 10.37373/infotech.v4i2.642.
- [12] H. Darwis *et al.*, "K-Means and K-Medoid in Clustering Analysis of Network Congestion Level," *ILKOM Jurnal Ilmiah*, vol. 17, no. 3, pp. 323–335, 2025, doi: 10.33096/ilkom.v17i3.2083.323-335.
- [13] S. F. Raheem and M. Alabbas, "Optimal k-means clustering using artificial bee colony algorithm with variable food sources length," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 5, pp. 5435–5443, Oct. 2022, doi: 10.11591/ijece.v12i5.pp5435-5443.
- [14] M. Harahap, Y. Lubis, and Z. Situmorang, "Analisis Pemasaran Bisnis dengan Data Science: Segmentasi Kepribadian Pelanggan berdasarkan Algoritma K-Means Clustering,"

- Data Sciences Indonesia (DSI)*, vol. 1, no. 2, pp. 76–88, 2022, doi: 10.47709/dsi.v1i2.1348.
- [15] Y. Chen, P. Tan, M. Li, H. Yin, and R. Tang, “K-means clustering method based on nearest-neighbor density matrix for customer electricity behavior analysis,” *International Journal of Electrical Power and Energy Systems*, vol. 161, no. January, p. 110165, 2024, doi: 10.1016/j.ijepes.2024.110165.
- [16] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, “K-Means Clustering Algorithms: A Comprehensive Review, Variants Analysis, And Advances In The Era Of Big Data,” *Inf. Sci. (N Y.)*, vol. 622, pp. 178–210, Apr. 2023, doi: 10.1016/j.ins.2022.11.139.
- [17] B. Rizki, N. G. Ginasta, M. A. Tamrin, and A. Rahman, “Customer Loyalty Segmentation on Point of Sale System Using Recency-Frequency-Monetary (RFM) and K-Means,” *Jurnal Online Informatika*, vol. 5, no. 2, p. 130, Dec. 2020, doi: 10.15575/join.v5i2.511.
- [18] A. T. Widiyanto and A. Witanti, “Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global),” *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 1, no. 1, pp. 204–215, 2021, doi: 10.24002/konstelasi.v1i1.4293.
- [19] M. H. M. Rohman *et al.*, “Clustering Analysis of Stunting Risk Factors Using K-Means and Principal Component Analysis: A Case Study in Indonesian Regency,” *sinkron*, vol. 9, no. 1, pp. 65–77, Jan. 2025, doi: 10.33395/sinkron.v9i1.14311.
- [20] M. Ridha, “JITE (Journal of Informatics and Telecommunication Engineering) Clustering Analysis to Identify Stunting Vulnerability Areas in North Aceh District Using the Fuzzy C-Means Algorithm,” *JITE*, vol. 9, no. 1, 2025, doi: 10.31289/jite.v9i1.14892.
- [21] I. F. Ashari, E. Dwi Nugroho, R. Baraku, I. Novri Yanda, and R. Liwardana, “Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta,” *Journal of Applied Informatics and Computing*, vol. 7, no. 1, pp. 89–97, 2023, doi: 10.30871/jaic.v7i1.4947.
- [22] A. Anagnostopoulou, N. Eleftherakis, and E. Karanasios, “Obesity and Underweight in Children Who Undergo Catheterization for Congenital Heart Disease: A Retrospective Study,” *Global Pediatrics*, vol. 7, no. December, p. 100095, 2024, doi: 10.1016/j.gped.2023.100095.
- [23] N. K. Pane, U. H. Almadany, and E. Sujoko, “Status Gizi Ibu Hamil sebagai Prediktor Kejadian Stunting pada Anak Usia 24–59 Bulan di Kecamatan Padangsidimpuan Selatan,” *PubHealth Jurnal Kesehatan Masyarakat*, vol. 4, no. 1, pp. 46–53, Jul. 2025, doi: 10.56211/pubhealth.v4i1.1026.
- [24] F. Rahmawan and T. D. Widajantie, “Faktor – Faktor Pemicu Terjadinya Kesalahan Sistem Informasi Manajemen Terhadap Realisasi Penerimaan Puskesmas ‘X,’” *JIPM: Jurnal Informasi Pengabdian Masyarakat*, vol. 1, no. 4, pp. 33–39, 2023.
- [25] N. Komang Tri Apriastini, N. Putu Tia Adnyani, P. Onik Selvyani, K. Hendra Setiawan, and P. Kedokteran, “Stunting: Faktor Risiko, Diagnosis, Tatalaksana, Dan Prognosis.”
- [26] I. Purnamasari *et al.*, “Analisis Faktor Risiko Yang Mempengaruhi Kejadian Stunting Pada Balita Info Artikel Abstrak,” *Jurnal Penelitian dan Pengabdian Kepada Masyarakat UNSIQ*, vol. 9, no. 1, pp. 48–56.
- [27] C. Putri, W. Kase, and Y. J. Prasetyo, “Analisis Faktor Risiko Stunting pada Balita di Desa Kesetnana Menggunakan Metode Random Forest,” 2025. [Online]. Available: <https://journal.stmiki.ac.id>
- [28] Y. Nurdiansyah, E. P. A. Wulandari, and D. A. R. Wulandari, “Analisis Faktor Kepuasan Pengguna Layanan Website SKCK Online Menggunakan Metode End User Computing Satisfaction (EUCS) (Studi Kasus : Banyuwangi),” *Informatics Journal*, vol. 5, no. 2, pp. 72–76, 2020, [Online]. Available: <https://jurnal.unej.ac.id/index.php/INFORMAL/article/view/18669>
- [29] S. Anwar *et al.*, “Systematic Review Faktor Risiko, Penyebab Dan Dampak Stunting Pada Anak (Systematic Review Risk Factors, Causes And Impact Of Stunting In Children),” *Jurnal Ilmu Kesehatan*, vol. 11, no. 1, 2022.
- [30] M. A. Robbani, G. Firmansyah, A. M. Widodo, and B. Tjahjono, “Clustering of Child Stunting Data in Tangerang Regency Using Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods,” 2024. [Online]. Available: <https://ajosh.org/>
- [31] K. P. Sinaga and M. S. Yang, “Unsupervised K-means Clustering Algorithm,” *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [32] D. Shah, S. Patel, and S. K. Bharti, “Heart Disease Prediction using Machine Learning Techniques,” *SN Comput. Sci.*, vol. 1, no. 6, Nov. 2020, doi: 10.1007/s42979-020-00365-y.
- [33] Singh Archana and Kumar Rakesh, “Heart Disease Prediction Using Machine Learning Algorithms,” *International Conference on Electrical and Electronics Engineering*, 2020, doi: 10.1109/ICE348803.2020.9122958.
- [34] R. Radhika and S. Thomas George, “Heart Disease Classification Using Machine Learning Techniques,” *J. Phys. Conf. Ser.*, vol. 1937, no.

- 1, 2021, doi: 10.1088/1742-6596/1937/1/012047.
- [35] N. Novalina, I. A. A. Tarigan, F. K. Kameela, and M. Rizkinia, "Benchmarking machine learning algorithm for stunting risk prediction in Indonesia," *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 3, pp. 2252–2263, Jun. 2025, doi: 10.11591/eei.v14i3.8997.
 - [36] N. F. Sahamony, T. Terttiaavini, and H. Rianto, "Analisis Perbandingan Kinerja Model Machine Learning untuk Memprediksi Risiko Stunting pada Pertumbuhan Anak," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 413–422, Feb. 2024, doi: 10.57152/malcom.v4i2.1210.
 - [37] M. Thangamani, R. Vijayalakshmi, M. Ganthimathi, M. Ranjitha, P. Malarkodi, and S. Nallusamy, "Efficient classification of heart disease using K-means clustering algorithm," *International Journal of Engineering Trends and Technology*, vol. 68, no. 12, pp. 48–53, 2020, doi: 10.14445/22315381/IJETT-V68I12P209.
 - [38] G. Fouriza Ibanez, W. Giri, and R. Wiriasto, "Kombinasi Principal Component Analysis dengan Algoritma K-Means untuk Klasterisasi Data Stunting," *Media Online*, vol. 5, no. 1, pp. 131–141, 2024, doi: 10.30865/klik.v5i1.1977.
 - [39] I. Anggriawan and W. Gunawan, "Implementation of Data Mining Using K-Means Algorithm for Bicycle Sales Prediction," *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 284–293, Dec. 2022, doi: 10.33096/ilkom.v14i3.1291.284-293.
 - [40] K. Tabianan, S. Velu, and V. Ravi, "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data," *Sustainability (Switzerland)*, vol. 14, no. 12, pp. 1–15, 2022, doi: 10.3390/su14127243.
 - [41] S. Juanita and R. D. Cahyono, "K-Means Clustering With Comparison of Elbow and Silhouette Methods for Medicines Clustering Based on User Reviews," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 1, pp. 283–289, 2024, [Online]. Available: <https://doi.org/10.52436/1.jutif.2024.5.1.1349>
 - [42] E. Umargono, J. E. Suseno, and V. G. S. K., "K-Means Clustering Optimization using the Elbow Method and Early Centroid Determination Based-on Mean and Median," vol. 474, no. Isstec 2019, pp. 234–240, 2020, doi: 10.5220/0009908402340240.
 - [43] Taufik Hidayat, Mohamad Jajuli, and Susilawati, "Clustering daerah rawan stunting di Jawa Barat menggunakan algoritma K-Means," *INFOTECH: Jurnal Informatika & Teknologi*, vol. 4, no. 2, pp. 137–146, Dec. 2023, doi: 10.37373/infotech.v4i2.642