

## Analisis Perbandingan Algoritma *K-Means* dan *K-Medoids* dalam Penentuan Status Gizi Balita

Jefania Tilman Soares<sup>1</sup>, Krisantus J. Tey Seran<sup>2\*</sup>, Fetronele Rambu Bobu<sup>3</sup>, Debora Chisinta<sup>4</sup>

<sup>1234</sup>Program Studi Teknologi Informasi, Universitas Negeri Timor, Indonesia

email: <sup>1</sup>[jevasoares4@gmail.com](mailto:jevasoares4@gmail.com)

email: <sup>2\*</sup>[krisantusteyseran@unimor.ac.id](mailto:krisantusteyseran@unimor.ac.id)

email: <sup>3</sup>[fetronela@unimor.ac.id](mailto:fetronela@unimor.ac.id)

email: <sup>4</sup>[deborachrisinta@unimor.ac.id](mailto:deborachrisinta@unimor.ac.id)

### Keywords:

*Clustering,*  
*K-Means,*  
*K-Medoids,*  
*Nutritional Status,*  
*Davies Bouldin Index.*

### ABSTRACT

Nutritional status in toddlers is an important indicator in determining child growth and development quality. Inaccurate classification of nutritional status can affect early intervention efforts. This study aims to compare the performance of *K-Means* and *K-Medoids* algorithms in clustering toddler nutritional status data at Puskesmas Betun. The dataset consists of 1,036 toddler records with variables including age, weight, height, and mid-upper arm circumference (MUAC). Data preprocessing was conducted through normalization before clustering. The performance of both algorithms was evaluated using the Davies Bouldin Index (DBI). The results show that *K-Means* converged in 24 iterations with a DBI value of 1.0281, while *K-Medoids* converged in 6 iterations with a DBI value of 1.1236. Based on the DBI evaluation, *K-Means* produced better clustering performance compared to *K-Medoids*. Therefore, *K-Means* is more suitable for determining toddler nutritional status in this study.

### Kata Kunci

*Clustering,*  
*K-Means,*  
*K-Medoids,*  
*Status Gizi,*  
*Davies Bouldin Index.*

### ABSTRAK

Status gizi balita merupakan indikator penting dalam menentukan kualitas pertumbuhan dan perkembangan anak. Kesalahan dalam pengelompokan status gizi dapat mempengaruhi ketepatan intervensi dini. Penelitian ini bertujuan untuk membandingkan kinerja algoritma *K-Means* dan *K-Medoids* dalam pengelompokan data status gizi balita di Puskesmas Betun. Dataset yang digunakan berjumlah 1.036 data balita dengan variabel usia, berat badan, tinggi badan, dan lingkar lengan atas (LiLA). Tahapan praproses dilakukan melalui normalisasi data sebelum proses *clustering*. Evaluasi performa algoritma menggunakan *Davies Bouldin Index* (DBI). Hasil penelitian menunjukkan bahwa *K-Means* mencapai konvergensi pada 24 iterasi dengan nilai *DBI* sebesar 1.0281, sedangkan *K-Medoids* mencapai konvergensi pada 6 iterasi dengan nilai *DBI* sebesar 1.1236. Berdasarkan hasil evaluasi *DBI*, algoritma *K-Means* memiliki performa *clustering* yang lebih baik dibandingkan *K-Medoids*. Dengan demikian, *K-Means* lebih direkomendasikan dalam penentuan status gizi balita pada penelitian ini.

### Korespondensi Penulis \*):

Krisantus Jumarto Tey Seran

Program Studi Teknologi Informasi, Fakultas Pertanian, Sains dan Kesehatan, Universitas Negeri Timor Timur Tengah Utara, Indonesia.

Diajukan: 25-02-2026 | Direvisi: 03-03-2025 | Diterima: 13-03-2026 | Diterbitkan: 13-04-2026

## 1. PENDAHULUAN

Status gizi balita merupakan indikator penting dalam menentukan kualitas pertumbuhan dan perkembangan anak, terutama pada masa *golden age* yang berpengaruh terhadap perkembangan fisik dan kognitif jangka panjang. Kekurangan gizi seperti *stunting*, *wasting*, dan *underweight* dapat menimbulkan dampak permanen apabila tidak ditangani secara dini [1]. Berdasarkan laporan *United Nations Children's Fund* (UNICEF) tahun 2023, prevalensi *stunting* di Indonesia mencapai 21,5%, *wasting* 8,5%, dan obesitas 4,2% [2]. Data tersebut menunjukkan bahwa permasalahan gizi balita masih menjadi tantangan dalam pembangunan kesehatan nasional.

Kabupaten Malaka di Provinsi Nusa Tenggara Timur juga menghadapi permasalahan serupa. Data Badan Pusat Statistik NTT mencatat adanya peningkatan kasus gizi buruk dari 177 kasus pada tahun 2022 menjadi 264 kasus pada tahun 2023 [3]. Kondisi ini menuntut adanya pendekatan berbasis data untuk mendukung proses identifikasi dan pengelompokan status gizi secara lebih sistematis.

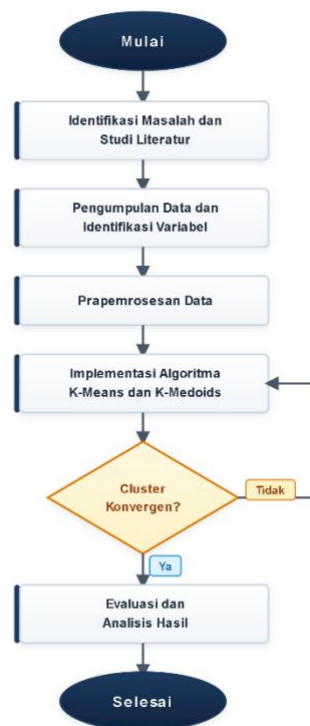
Pemantauan status gizi di Puskesmas Betun masih dilakukan secara manual sehingga kurang efisien dalam mengelola data dalam jumlah besar. Oleh karena itu, diperlukan pendekatan data *mining*, khususnya teknik *clustering*, untuk mengelompokkan data berdasarkan kemiripan karakteristik tanpa label awal. Dalam penelitian ini, indikator antropometri yang digunakan meliputi usia, berat badan, tinggi badan, dan lingkaran lengan atas (LiLA).

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma *K-Means* efektif dalam mengelompokkan status gizi balita dengan kualitas *cluster* yang baik berdasarkan nilai *Davies-Bouldin Index* (DBI) [4]. Penelitian lain menyatakan bahwa algoritma *K-Medoids* memiliki keunggulan dalam ketahanan terhadap *outlier* dan stabilitas pusat *cluster* [5]. Namun, penelitian yang secara langsung membandingkan performa *K-Means* dan *K-Medoids* pada data status gizi balita di wilayah Kabupaten Malaka yang masih terbatas.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membandingkan performa algoritma *K-Means* dan *K-Medoids* dalam pengelompokan status gizi balita di Puskesmas Betun menggunakan metode evaluasi *Davies-Bouldin Index* (DBI) dan pendekatan metode *Elbow* untuk menentukan jumlah *cluster* optimal. Hasil penelitian diharapkan dapat memberikan rekomendasi algoritma yang lebih efektif dalam mendukung pengambilan keputusan berbasis data pada layanan kesehatan tingkat dasar.

## 2. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan sistematis yang meliputi identifikasi masalah, pengumpulan data, prapemrosesan data, implementasi algoritma, serta evaluasi hasil clustering sebagaimana ditunjukkan pada gambar 1.



Gambar 1. Diagram Alir Penelitian

### 2.1 Tahapan Penelitian

Gambar 1 menunjukkan tahapan – tahapan penelitian yang dilakukan. Tahapan pertama adalah identifikasi masalah dan studi literatur, yaitu mengidentifikasi permasalahan yang berkaitan dengan penentuan status gizi balita di Puskesmas Betun serta mengkaji penelitian-penelitian terdahulu yang relevan sebagai dasar teoritis penggunaan algoritma *K-Means* dan *K-Medoids*. Selanjutnya dilakukan pengumpulan data dan identifikasi variabel, di mana data diperoleh dari pencatatan rutin unit program gizi balita di Puskesmas Betun, Kabupaten Malaka, periode 2023–2024, dengan variabel yang digunakan meliputi empat indikator antropometri yaitu usia (bulan), berat badan (kg), tinggi badan (cm), dan lingkaran lengan atas (LiLA) (cm).

Tahap prapemrosesan data meliputi pembersihan dan normalisasi. Pembersihan dilakukan dengan menghapus data duplikat, nilai kosong, dan data tidak valid sehingga diperoleh 1.025 data balita yang siap dianalisis. Selanjutnya,

normalisasi menggunakan metode *Min-Max Scaler* untuk mengubah nilai ke rentang 0–1 agar setiap variabel memiliki kontribusi yang seimbang dalam proses *clustering*.

Setelah data siap, dilakukan implementasi algoritma *K-Means* dan *K-Medoids* dengan jumlah *cluster*  $k=3$  yang ditentukan berdasarkan hasil analisis metode *Elbow*. Kedua algoritma menjalankan proses iterasi secara berulang hingga mencapai kondisi konvergen, yaitu ketika tidak terjadi perubahan signifikan pada posisi *centroid* maupun *medoid*. Apabila kondisi konvergen belum tercapai, proses iterasi diulang kembali hingga hasil pengelompokan menjadi stabil. Pada penelitian ini, algoritma *K-Means* mencapai konvergensi pada iterasi ke-24, sedangkan algoritma *K-Medoids* mencapai konvergensi lebih cepat pada iterasi ke-6.

Tahap akhir penelitian adalah evaluasi dan analisis hasil. Kualitas *clustering* dari kedua algoritma diukur menggunakan *Davies-Bouldin Index (DBI)*, di mana semakin kecil nilai *DBI* maka semakin baik kualitas *cluster* yang dihasilkan. Berdasarkan hasil evaluasi, algoritma *K-Means* memperoleh nilai *DBI* sebesar 1,0218 sedangkan algoritma *K-Medoids* memperoleh nilai *DBI* sebesar 1,1236. Hasil tersebut menunjukkan bahwa algoritma *K-Means* lebih efektif dibandingkan *K-Medoids* dalam pengelompokan status gizi balita di Puskesmas Betun.

## 2.2 Landasan Teori

### 2.2.1 Data Mining

*Data Mining* merupakan proses analitik yang memanfaatkan teknik-teknik statistika, matematika dan kecerdasan buatan untuk mengeksplorasi dan menggali informasi yang berharga atau pola yang tersembunyi dari kumpulan data yang besar dan kompleks [6].

*Data mining* sering disebut sebagai *Knowledge Discovery in Database (KDD)* karena berfokus pada penemuan pengetahuan dari basis data berukuran besar. Proses *KDD* meliputi beberapa tahap, yaitu data cleaning untuk menghapus nilai tidak valid, data integration untuk menggabungkan data yang relevan, selection untuk memilih data yang dibutuhkan, data transformation untuk menyesuaikan format data, data mining untuk menemukan pola, evaluation untuk menilai pola yang dihasilkan, serta *knowledge presentation* untuk menyajikan hasil analisis agar mudah dipahami dan mendukung pengambilan keputusan [7].

### 2.2.2 Clustering

*Clustering* adalah metode *data mining* yang digunakan untuk mengelompokkan sekumpulan data atau objek ke dalam kelompok (*cluster*) yang memiliki karakteristik khusus. *Clustering* yaitu kumpulan objek data yang memiliki karakteristik sama antara satu dengan yang lain dari kelompok yang sama dan berbeda objek dengan data kelompok lain[8]. Tujuan utama *clustering* adalah membagi data menjadi grup yang lebih seragam, di mana objek dalam satu grup lebih mirip satu sama lain dibandingkan dengan objek di grup lain.

Berbeda dengan metode pembelajaran berbasis pengawasan, proses *clustering* bersifat *unsupervised* sehingga pengelompokan data dilakukan tanpa menggunakan label atau kategori yang telah ditentukan sebelumnya. Algoritma bekerja secara otomatis dengan mencari kemiripan dan pola yang ada dalam data. Kemampuan ini menjadikan *clustering* sangat berguna bagi berbagai instansi dalam menganalisis data berukuran besar untuk menemukan informasi tersembunyi yang dapat mendukung proses pengambilan keputusan. Dalam penelitian ini, metode *clustering* digunakan untuk mengelompokkan data antropometri balita di Puskesmas Betun guna mengidentifikasi pola status gizi secara objektif berdasarkan karakteristik fisik yang dimiliki setiap balita.

### 2.2.3 Algoritma K-Means

*K-Means* merupakan algoritma *clustering* yang bekerja secara iteratif dengan mengelompokkan data ke dalam  $k$  *cluster* berdasarkan kemiripan karakteristik. Proses dimulai dengan menentukan jumlah *cluster*  $k$ , kemudian memilih titik *centroid* awal secara acak sebagai pusat masing-masing *cluster* [9]. Algoritma ini termasuk metode pembelajaran tanpa pengawasan (*unsupervised*) sehingga tidak memerlukan label atau kategori data sebelumnya dalam proses pengelompokan. Setelah *centroid* awal ditentukan, setiap data dihitung jaraknya terhadap seluruh *centroid* menggunakan rumus *Euclidean Distance* sebagaimana ditunjukkan pada Persamaan (1), kemudian data dikelompokkan ke *cluster* dengan jarak terdekat [10].

$$D(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \dots\dots\dots (1)$$

Keterangan:  $D$  menunjukkan jarak *Euclidean*,  $(x_1, y_1)$  menunjukkan koordinat titik awal, dan  $(x_2, y_2)$  menunjukkan koordinat titik kedua.

Selanjutnya posisi *centroid* diperbarui dengan menghitung nilai rata-rata seluruh data dalam setiap *cluster* menggunakan Persamaan (2) [11].

$$C_i = \frac{(x_1 + x_2 + x_3 + \dots + x_n)}{(\sum x)} \dots\dots\dots (2)$$

Keterangan:  $x_1$  adalah nilai *record* data ke-1,  $x_2$  adalah nilai *record* data ke-2, dan  $\sum x$  adalah jumlah *record* data. Proses perhitungan jarak dan pembaruan *centroid* diulang secara terus-menerus hingga posisi *centroid* tidak



### 2.2.9 Puskesmas Betun

Puskesmas Betun merupakan salah satu fasilitas pelayanan kesehatan masyarakat yang menyediakan berbagai layanan, termasuk pemantauan status gizi balita. Fasilitas ini berlokasi di Jl. Wehali, Kecamatan Malaka Tengah, Kabupaten Malaka, Nusa Tenggara Timur. Sebagai layanan kesehatan tingkat pertama, Puskesmas Betun memiliki peran strategis dalam mendeteksi permasalahan gizi di wilayah kerjanya serta melaksanakan intervensi yang diperlukan guna meningkatkan status gizi balita. Dalam pelaksanaan tugasnya, Puskesmas Betun secara rutin mencatat data antropometri balita, meliputi berat badan, tinggi badan, usia, dan lingkaran lengan atas (LiLA) sebagai bagian dari program pemantauan gizi. Data balita periode 2023–2024 yang diperoleh dari unit program gizi tersebut menjadi sumber utama dalam penelitian ini. Pemanfaatan data ini bertujuan untuk membantu pihak Puskesmas dalam mengelompokkan status gizi balita secara lebih objektif dan sistematis melalui metode *clustering*.

## 3. HASIL DAN ANALISIS

Status gizi balita di Puskesmas Betun dikelompokkan berdasarkan usia, berat, tinggi, dan LiLA. Data – data yang dikumpulkan diolah menggunakan algoritma *K-Means* dan *K-Medoids*. Perbandingan kinerja keduanya dalam mengelompokkan status gizi balita didasarkan pada evaluasi *Davies-Bouldin Index*.

### 3.1 Deskripsi Dataset

Data yang digunakan dalam penelitian ini bersumber dari pencatatan rutin unit program gizi balita di Puskesmas Betun, Kabupaten Malaka, periode 2023–2024. Data awal yang diperoleh berjumlah 1.036 data balita, kemudian setelah melalui proses pembersihan data (*data cleaning*) untuk menghapus nilai kosong dan nilai tidak valid, jumlah data yang layak dianalisis menjadi 1.025 data balita. Variabel yang digunakan mencakup empat indikator antropometri, yaitu usia (bulan), berat badan (kg), tinggi badan (cm), dan lingkaran lengan atas (LiLA) (cm).

Sebelum proses *clustering* dilakukan, analisis statistik deskriptif dilakukan untuk memahami karakteristik awal data. Hasilnya disajikan pada Tabel 1.

**Tabel 1.** Statistik Deskriptif Data Balita

Variabel	Minimum	Maksimum	Rata-rata	Varians
Usia	38	61	49,023	2,18
Berat	6,1	19,0	11,787	4,22
Tinggi	67,9	110,7	90,223	72,81
LiLa	11,4	19,4	15,075	1,18

Berdasarkan tabel tersebut, variabel tinggi badan memiliki tingkat penyebaran tertinggi (varians = 72,81), sementara usia menunjukkan sebaran yang paling homogen (varians = 2,18). Hal ini mengindikasikan bahwa perbedaan karakteristik fisik balita dalam penelitian lebih dominan pada aspek tinggi badan dibandingkan variabel lainnya.

### 3.2 Praproses Data

Sebelum proses *clustering* dilakukan, data terlebih dahulu melewati tahapan prapemrosesan yang mencakup pembersihan data dan normalisasi data. Pada tahap pembersihan data (*data cleaning*), duplikasi data dihapus serta baris yang mengandung nilai kosong (*NaN*) dan nilai nol pada kolom numerik dieliminasi agar tidak mengganggu proses analisis selanjutnya. Setelah data bersih diperoleh, dilakukan normalisasi data menggunakan metode *Min-Max Scaler* karena setiap variabel memiliki satuan dan rentang nilai yang berbeda. Proses normalisasi ini mengubah seluruh nilai data ke dalam rentang 0 hingga 1 sehingga setiap variabel memberikan kontribusi yang seimbang dalam perhitungan jarak antar data pada proses *clustering*.

### 3.3 Penentuan Jumlah Cluster Optimal

Jumlah *cluster* optimal ditentukan menggunakan Metode *Elbow*, yaitu dengan memplot nilai *Sum of Squared Error* (SSE) terhadap berbagai jumlah *cluster*. Nilai *SSE* semakin menurun seiring bertambahnya jumlah *cluster*; titik "siku" pada kurva menunjukkan jumlah *cluster* yang paling efisien. Berdasarkan hasil analisis Metode *Elbow*, jumlah *cluster* optimal yang dipilih adalah  $k = 3$ , yang memberikan pemisahan data yang jelas berdasarkan karakteristik antropometri balita.

### 3.4 Hasil *Clustering K-Means*

Algoritma *K-Means* dijalankan dengan tiga *centroid* awal yang dipilih secara acak dari data yang telah dinormalisasi. Setiap iterasi, jarak setiap data terhadap ketiga *centroid* dihitung menggunakan rumus pada persamaan 1, kemudian data dikelompokkan ke *cluster* dengan jarak terpendek dan *centroid* diperbarui dengan menghitung rata-rata data dalam setiap *cluster*. Proses ini diulang secara terus-menerus hingga algoritma mencapai kondisi konvergen pada iterasi ke-24. Hasil akhir pengelompokan menunjukkan bahwa dari total 1.025 data balita, sebanyak 380 data masuk ke dalam *Cluster* 1, 249 data masuk ke dalam *Cluster* 2, dan 396 data masuk ke dalam *Cluster* 3.

Setelah konvergensi, nilai rata-rata variabel pada setiap *cluster* dihitung berdasarkan data asli (sebelum normalisasi) untuk keperluan interpretasi status gizi. Hasilnya disajikan pada Tabel 2.

**Tabel 2.** Karakteristik Rata-Rata *Cluster* Hasil *K-Means*

<i>Cluster</i>	Usia (bln)	Berat (kg)	Tinggi (cm)	LiLA (cm)	Kategori Gizi
C1	48,99	11,43	88,98	14,94	Gizi Kurang
C2	48,89	9,14	78,65	14,35	Gizi Buruk
C3	49,14	13,79	98,69	15,65	Gizi Normal

Berdasarkan tabel di atas, *Cluster 2* menunjukkan nilai berat badan dan tinggi badan paling rendah sehingga dikategorikan sebagai kondisi gizi paling buruk, sedangkan *Cluster 3* memiliki nilai antropometri tertinggi yang mencerminkan kondisi gizi normal, dan *Cluster 1* berada di posisi menengah dengan kategori gizi kurang. Ketiga *cluster* memiliki rentang usia yang relatif sama yaitu sekitar 49 bulan, sehingga perbedaan antar *cluster* lebih dominan dipengaruhi oleh variabel berat badan dan tinggi badan dibandingkan variabel usia. Hal ini selaras dengan penelitian Djaelani dkk. [18] yang membuktikan bahwa algoritma *K-Means* mampu menghasilkan pengelompokan status gizi balita yang berkualitas berdasarkan evaluasi *DBI*. Adapun perbedaan nilai *DBI* yang diperoleh dalam penelitian ini 1,0218 dibandingkan penelitian tersebut yang dipengaruhi oleh kompleksitas data yang lebih tinggi, mengingat penelitian ini melibatkan empat variabel antropometri dengan data yang lebih besar yaitu sebanyak 1.025 data balita. Hal serupa juga turut diperkuat oleh hasil penelitian Janah dkk. [19] yang memanfaatkan *DBI* sebagai metrik evaluasi kualitas *cluster* pada penerapan algoritma *K-Means* untuk data status gizi balita di Puskesmas, sehingga semakin mengonfirmasi bahwa *DBI* merupakan metrik yang tepat dan konsisten dalam mengukur kualitas pengelompokan data antropometri balita.

### 3.5 Hasil *Clustering K-Medoids*

Algoritma *K-Medoids* diinisialisasi dengan tiga *medoid* awal yang dipilih secara acak dari dataset, masing-masing mewakili baris ke-138 (C1), ke-70 (C2), dan ke-350 (C3). Berbeda dengan *K-Means*, pusat *cluster* pada *K-Medoids* adalah data aktual, bukan nilai rata-rata. Setiap iterasi dilakukan pemilihan kandidat *medoid* baru dan dihitung apakah penggantian *medoid* dapat menurunkan total *cost* (total jarak kedekatan).

Penurunan total *cost* terjadi secara konsisten pada setiap iterasi, di mana pada iterasi pertama total *cost* bernilai 280,014696 kemudian menurun menjadi 189,4497009 pada iterasi kedua, 162,3578967 pada iterasi ketiga, 161,3556136 pada iterasi keempat, dan 160,5477742 pada iterasi kelima. Pada iterasi keenam, total *cost* mencapai 160,471748 tanpa adanya perubahan *medoid* pada ketiga *cluster*, yang menandakan algoritma telah mencapai kondisi konvergen sehingga proses iterasi dihentikan. Dengan demikian, proses konvergensi *K-Medoids* tercapai lebih cepat dibandingkan *K-Means*, yaitu hanya pada iterasi ke-6. Hasil akhir pengelompokan menunjukkan bahwa dari total 1.025 data balita, sebanyak 369 data masuk ke dalam *Cluster 1*, 373 data masuk ke dalam *Cluster 2*, dan 283 data masuk ke dalam *Cluster 3*.

Karakteristik rata-rata setiap *cluster K-Medoids* disajikan pada Tabel 3.

**Tabel 3.** Karakteristik Rata-Rata Hasil *Cluster K-Medoids*

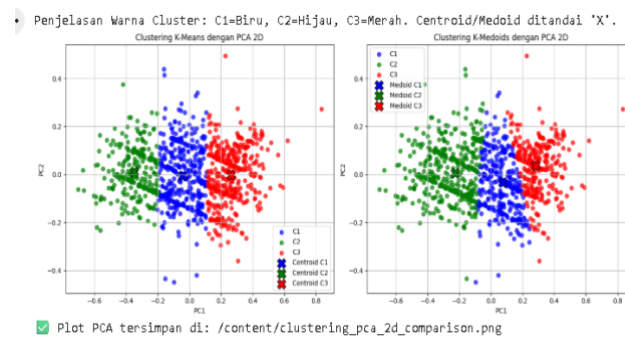
<i>Cluster</i>	Usia (bln)	Berat (kg)	Tinggi (cm)	LiLA (cm)	Kategori Gizi
C1	49,14	13,87	99,05	15,64	Gizi Normal
C2	48,97	11,58	89,68	15,08	Gizi Kurang
C3	48,91	9,32	79,42	14,31	Gizi Buruk

Hasil *clustering* menggunakan algoritma *K-Medoids* menunjukkan bahwa *Cluster 1* dengan rata-rata berat badan 13,87 kg dan tinggi badan 99,05 cm dikategorikan sebagai Gizi Normal, *Cluster 2* dengan rata-rata berat badan 11,58 kg dan tinggi badan 89,68 cm dikategorikan sebagai Gizi Kurang, serta *Cluster 3* dengan rata-rata berat badan 9,32 kg dan tinggi badan 79,42 cm dikategorikan sebagai Gizi Buruk. Perbedaan yang cukup signifikan terlihat antara *Cluster 1* dan *Cluster 3*, di mana selisih berat badan mencapai 4,55 kg dan selisih tinggi badan mencapai 19,63 cm, yang mengindikasikan adanya kesenjangan kondisi gizi yang nyata antar kelompok. Pembagian ke dalam tiga kategori status gizi ini sejalan dengan penelitian Hidayat dkk. [20] yang juga menyimpulkan bahwa jumlah *cluster* optimal adalah  $k=3$  dalam pengelompokan data gizi balita, baik menggunakan *K-Means* maupun *K-Medoids*. Hal ini mengonfirmasi bahwa pengelompokan menjadi tiga kategori yaitu Normal, Gizi Kurang, dan Gizi Buruk merupakan representasi yang tepat dan konsisten dalam menggambarkan variasi kondisi gizi pada populasi balita, terlepas dari perbedaan wilayah dan skala data yang digunakan.

Efektivitas algoritma *K-Medoids* dalam pengelompokan data balita juga ditunjukkan oleh hasil penelitian Pohan dkk. [21] yang berhasil menerapkan *K-Medoids* untuk mengelompokkan prevalensi *stunting* balita di Indonesia berdasarkan karakteristik wilayah dan menghasilkan pola pengelompokan yang informatif sebagai dasar perumusan strategi intervensi gizi. Hal ini memperkuat bahwa *K-Medoids* memiliki kemampuan yang baik dalam mengidentifikasi pola pengelompokan pada data kesehatan balita, meskipun dalam penelitian ini nilai *DBI* yang dihasilkan menunjukkan kualitas *cluster* yang sedikit lebih rendah dibandingkan *K-Means*.

### 3.6 Visualisasi *PCA*

Untuk mempermudah interpretasi visual, hasil *clustering* dari kedua metode divisualisasikan menggunakan reduksi dimensi *Principal Component Analysis* (*PCA*) ke dalam ruang dua dimensi sebagaimana ditunjukkan pada Gambar 2.

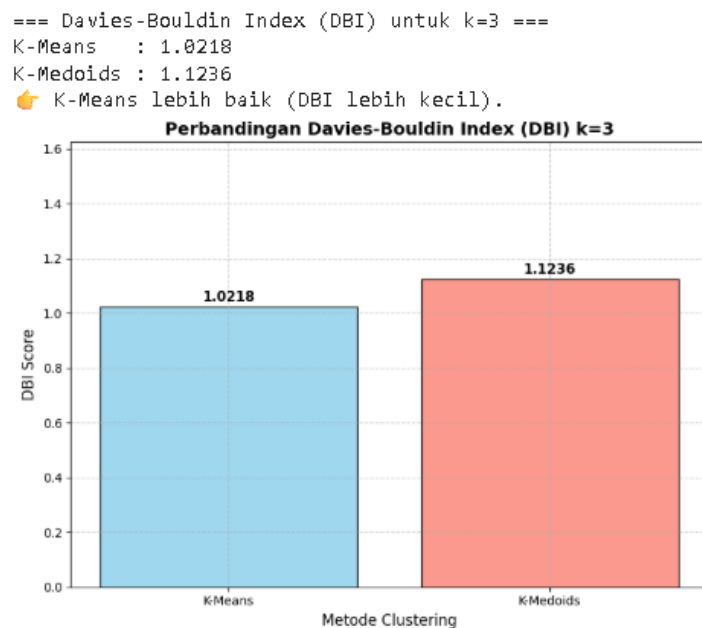


Gambar 2. Hasil Visualisasi PCA

Berdasarkan Gambar 2, dua komponen utama PCA mampu menjelaskan sekitar 92,42% dari total variansi data, di mana komponen pertama (PC1) menjelaskan 76,08% dan komponen kedua (PC2) sebesar 16,34%, sehingga representasi dua dimensi ini dinilai sudah sangat memadai dalam merepresentasikan keseluruhan informasi data. Secara visual, pola pengelompokan yang dihasilkan oleh K-Means dan K-Medoids tampak serupa, dengan ketiga cluster tersebar di sepanjang sumbu PC1. Terdapat sedikit tumpang tindih antar cluster pada beberapa bagian, yang menunjukkan adanya balita dengan karakteristik antropometri yang relatif mirip di antara kelompok yang berbeda.

### 3.7 Evaluasi Kualitas Clustering

Kualitas clustering dari kedua algoritma diukur menggunakan Davies-Bouldin Index (DBI), di mana semakin kecil nilai DBI maka semakin baik kualitas cluster yang dihasilkan. Hasil evaluasi disajikan pada gambar 3.



Gambar 3. Hasil Evaluasi Clustering

Berdasarkan pada Gambar 3, algoritma K-Means memperoleh nilai DBI sebesar 1,0218 sedangkan algoritma K-Medoids memperoleh nilai sebesar 1,1236 dengan selisih 0,1018. Meskipun K-Means secara teknis menghasilkan kualitas clustering yang lebih baik berdasarkan nilai DBI yang lebih kecil, selisih yang sangat kecil tersebut menunjukkan bahwa kedua metode menghasilkan kualitas pengelompokan yang relatif setara dalam mengelompokkan data status gizi balita. Dari sisi kecepatan konvergensi, K-Medoids lebih unggul karena hanya membutuhkan 6 iterasi dibandingkan K-Means yang membutuhkan 24 iterasi, namun keunggulan tersebut tidak cukup untuk mengimbangi kualitas cluster yang dihasilkan K-Means. Hal ini sejalan dengan temuan Muhammad dkk. [22] yang membandingkan kedua algoritma pada data status gizi anak dan menyimpulkan bahwa K-Means menunjukkan performa pengelompokan yang lebih baik dibandingkan K-Medoids. Dengan demikian, K-Means tetap direkomendasikan dalam penelitian ini karena menghasilkan cluster yang lebih kompak dan terpisah dengan baik, sekaligus memperkuat temuan sebelumnya bahwa K-Means lebih sesuai untuk pengelompokan data antropometri balita, khususnya pada dataset dengan distribusi yang relatif homogen seperti data di Puskesmas Betun. Perbandingan kualitas cluster antara K-Means dan K-Medoids ini juga sejalan dengan temuan Permatasari & Elfazza [23] yang membandingkan algoritma K-Means dan Partitioning Around Medoids (PAM) sebagai varian K-Medoids menggunakan evaluasi DBI, dan

menyimpulkan bahwa terdapat perbedaan kualitas *cluster* yang dihasilkan oleh kedua algoritma tersebut. Dengan demikian, hasil penelitian ini semakin memperkuat bahwa pemilihan algoritma *clustering* yang tepat sangat berpengaruh terhadap kualitas pengelompokan data status gizi balita, dan *K-Means* terbukti lebih unggul dalam menghasilkan *cluster* yang lebih kompak dan terpisah dengan baik pada data antropometri balita di Puskesmas Betun.

#### 4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *K-Means* dan *K-Medoids* untuk mengelompokkan 1.025 data balita di Puskesmas Betun ke dalam tiga *cluster* berdasarkan variabel antropometri, yaitu usia, berat badan, tinggi badan, dan lingkaran lengan atas (LiLA). Ketiga *cluster* yang terbentuk merepresentasikan kategori status gizi balita, yaitu Normal, Gizi Kurang, dan Gizi Buruk. Jumlah *cluster* optimal  $k=3$  ditentukan berdasarkan hasil analisis metode *Elbow*.

Berdasarkan hasil evaluasi menggunakan *Davies-Bouldin Index* (DBI), algoritma *K-Means* memperoleh nilai DBI sebesar 1,0218 sedangkan algoritma *K-Medoids* memperoleh nilai DBI sebesar 1,1236. Meskipun selisih keduanya sangat kecil yaitu 0,1018, algoritma *K-Means* dinyatakan lebih efektif dalam mengelompokkan status gizi balita di Puskesmas Betun dan dapat digunakan sebagai dasar pendukung pengambilan keputusan dalam perencanaan program intervensi gizi.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Puskesmas Betun, Kabupaten Malaka, Nusa Tenggara Timur yang telah memberikan izin dan menyediakan data penelitian sehingga penelitian ini dapat terlaksana dengan baik. Penulis juga mengucapkan terima kasih kepada Program Studi Teknologi Informasi, Fakultas Pertanian, Sains dan Kesehatan, Universitas Timor atas dukungan dan bimbingan yang diberikan selama proses penelitian hingga publikasi artikel ini.

#### REFERENSI

- [1] F. T. Wahyuni, N. S. Kirana, N. S. Ashifa, Dan P. Ella, "Kategorisasi Tingkat Gizi Balita Menggunakan K-Means : Studi Kasus Puskesmas Desa Karangsembung," Vol. 2024, No. Senada, Hal. 863–873, 2024.
- [2] Unicef, "Tren Malnutrisi Anak Di Indonesia." [Daring]. Tersedia Pada: <https://www.unicef.org/indonesia/id/gizi>
- [3] B. P. P. N. T. T. Statistik, "Jumlah Dan Persentase Balita Wasting Menurut Kabupaten/Kota, 2021-2023," Dinas Kesehatan Provinsi Ntt. [Daring]. Tersedia Pada: <https://ntt.bps.go.id/id/statistics-table/2/mtu3mimy/jumlah-dan-persentase-balita-wasting-menurut-kabupaten-kota.html>
- [4] Y. Eko, Y. Oktaviana Legu Rema, H. Herlina Ullu, B. Baso, Dan F. Pertanian Sains Dan Kesehatan, "Implementasi Metode K-Means Clustering Untuk Menentukan Kondisi Gizi Balita (Studi Kasus : Puskesmas Mamsena)," 2025.
- [5] Betrisandi, M. Hasan, Dan Bahrin, "Penerapan Algoritma Clustering K-Medoids Untuk Menentukan Status Gizi Balita Application Of The K-Medoids Clustering Algorithm To Determine The Nutritional Status Of Toddlers," Vol. 7, Hal. 94–99, 2025.
- [6] F. R. Janah, R. Kurniawan, Dan T. Suprapti, "Analisis Dataset Status Gizi Pada Balita Menggunakan Algoritma K-Means Clustering," 2023.
- [7] B. University, "Proses Data Mining Kdd." [Daring]. Tersedia Pada: <https://sis.binus.ac.id/2021/09/30/proses-data-mining-kdd/>
- [8] L. L. M. Betharia Sonatha Kaka, Andry Ananda P. T. Mara, "Penerapan Metode K-Means Clustering Bagi Balita Penderita Stunting Di Kecamatan Loli Kabupaten Sumba Barat Betharia," *J. Tek. Inform. Dan Sist. Inf.*, Vol. 10, No. 4, Hal. 256–269, 2023.
- [9] D. Dona Dan M. Rifqi, "Penerapan Metode K-Means Clustering Untuk Menentukan Status Gizi Baik Dan Gizi Buruk Pada Balita (Studi Kasus Kabupaten Rokan Hulu)," *Rabit J. Teknol. Dan Sist. Inf. Univrab*, Vol. 7, No. 2, Hal. 179–191, Jul 2022, Doi: 10.36341/Rabit.V7i2.2171.
- [10] L. Awaliyah, N. Rahaningsih, Dan R. Dinar Dana, "Implementasi Algoritma K-Means Dalam Analisis Cluster Korban Kekerasan Di Provinsi Jawa Barat," *Jati (Jurnal Mhs. Tek. Inform.)*, Vol. 8, No. 1, Hal. 188–195, 2024, Doi: 10.36040/Jati.V8i1.8332.
- [11] N. A. Maori Dan E. Evanita, "Metode Elbow Dalam Optimasi Jumlah Cluster Pada K-Means Clustering," *Simetris J. Tek. Mesin, Elektro Dan Ilmu Komput.*, Vol. 14, No. 2, Hal. 277–288, 2023, Doi: 10.24176/Simet.V14i2.9630.
- [12] D. Arbain, Sriyanto, Dan J. Triloka, "Perbandingan Kinerja Algoritma K-Medoids Dan K-Means Untuk Klasifikasi Penyakit Kanker Serviks," *Semin. Nas. Has. Penelit. Dan Pengabd. Masy.* 2023, No. 93, Hal. 118–131, 2023.
- [13] U. Linarti, A. H. Soleliza Jones, L. Zahrotun, Dan A. Rahmawati, "Penerapan Metode K-Medoids Guna Pengelompokan Data Usaha Mikro, Kecil Dan Menengah (Umkh) Bidang Kuliner Di Kota Yogyakarta," *J.*

- Ilmu Komput. Dan Sist. Inf.*, Vol. 7, No. 1, Hal. 37–45, 2024, Doi: 10.55338/Jikoms.V7i1.2194.
- [14] M. Qusyairi, Zul Hidayatullah, Dan Arnita Sandi, “Penerapan K-Means Clustering Dalam Pengelompokan Prestasi Siswa Dengan Optimasi Metode Elbow,” *Infotek J. Inform. Dan Teknol.*, Vol. 7, No. 2, Hal. 500–510, Jul 2024, Doi: 10.29408/Jit.V7i2.26375.
- [15] Nanda Shalsadilla, Shantika Martha, Dan Hendra Perdana, “Penentuan Jumlah Cluster Optimum Menggunakan Davies Bouldin Index Dalam Pengelompokan Wilayah Kemiskinan Di Indonesia,” *Stat. J. Theor. Stat. Its Appl.*, Vol. 23, No. 1, Hal. 63–72, 2023, Doi: 10.29313/Statistika.V23i1.1743.
- [16] N. A. Hidayatullah Dan W. Prihartono, “Clustering Algoritma K-Means Untuk Meningkatkan Efektivitas Program Sosial Di Kota / Kabupaten Cirebon,” Vol. 13, No. 1, Hal. 629–636, 2025.
- [17] K. Kadarsih Dan S. Andrianto, “Penerapan Metode Clustering K-Means Dalam Pengelompokan Penerima Kwh Metergratis Dengan Bahasa Pemrograman Python,” *Jtim J. Tek. Inform. Mahakarya*, Vol. 03, No. 2, Hal. 37–44, 2022.
- [18] M. Djaelani, U. Hayati, Dan A. Kesehatan, “Jurnal Informatika Dan Rekayasa Perangkat Lunak Clustering Status Gizi Balita Menggunakan Metode K-Means Pada Posyandu Desa Mekar Wangi,” 2024.
- [19] I. Iskandar Dan D. Jollyta, “Penerapan Algoritma K-Means Untuk Clusterisasi Kasus Stunting Di Provinsi Riau,” *J. Mhs. Apl. Teknol. Komput. Dan Inf.*, Vol. 6, No. 1, Hal. 18–22, 2024.
- [20] F. M. Hidayat, T. Rohana, E. Nurlaelasari, Dan A. F. N. Masruriah, “Klasterisasi Kabupaten Dan Kota Di Jawa Barat Dalam Kasus Gizi Buruk Menggunakan Algoritma K-Means Dan K-Medoids,” *J. Tekinkom (Teknik Inf. Dan Komputer)*, Vol. 7, No. 1, Hal. 251–261, 2024, Doi: 10.37600/Tekinkom.V7i1.1387.
- [21] H. Pohan, M. Zarlis, E. Irawan, H. Okprana, Dan Y. Pranayama, “Penerapan Algoritma K-Medoids Dalam Pengelompokan Balita Stunting Di Indonesia,” *Juki J. Komput. Dan Inform.*, Vol. 3, No. 2, Hal. 97–104, 2021, Doi: 10.53842/Juki.V3i2.69.
- [22] M. Muhammad, A. Mahmudi, Dan K. Auliasari, “Perbandingan Metode K-Means Dan K-Medoids Untuk Klasifikasi Status Gizi Anak,” 2023.
- [23] R. Permatasari Dan F. Elfazza, “Perbandingan Optimasi Penyuluhan Penyakit Stunting Pada Balita Integrasi Algoritma K- Means Dan Partitioning Around Medoids ( Pam ),” Vol. 1, No. 2, Hal. 90–99, 2024.