

Pemetaan Persebaran Status Gizi Balita untuk Penentuan Pemberian Penyuluhan dengan Pendekatan K-Means Clustering

Yunisa Ahgnia Zahrah^{1*}, Leny Tritanto Ningrum²

^{1,2}Program Studi Sistem Informasi, Fakultas Informatika dan Komputer, Universitas Binaniaga Indonesia
email: yunisaahgniaz@gmail.com

*Corresponding Author

ABSTRACT

Mapping the distribution of toddler nutritional status is an activity of grouping integrated health post (Posyandu) areas based on the conditions of undernutrition and overnutrition, which are still quite high. The problem faced is not only the high number of cases of undernutrition due to insufficient nutritional intake, but also the increasing number of cases of overnutrition triggered by unbalanced consumption patterns. However, to date, the available data has not been optimally utilized to produce accurate mapping information. This makes it difficult for community health centers (Puskesmas) to determine nutrition counseling targets that truly match the needs in the field. Furthermore, the limitations of manual analysis make identifying areas with cases of undernutrition or overnutrition less effective. The inaccurate distribution of toddler nutritional status can result in failure to achieve the target of nutrition counseling that is more targeted. The purpose of this study is to obtain the effectiveness and accuracy in mapping the distribution of toddler nutritional status for determining counseling using the K-Means Clustering approach. This study produces a prototype system that can map toddler nutritional status using the K-Means Clustering Algorithm by analyzing the initial data group, transforming and normalizing the data, and performing clustering calculations. The variables used include age, weight, height, z score bb/u, z score tb/u, and z score bb/tb. The grouping results were re-analyzed to see the posyandu area in each group based on the nutritional status of toddlers. This was done to map nutrition counseling through a decision support system as a reference for the Tanah Sareal Community Health Center in Bogor City in determining effective counseling targets. The application of the K-Means Clustering Algorithm resulted in Cluster 1 of 276 and Cluster 2 of 220 with the categories of Undernutrition and Overnutrition. The feasibility test by the system expert obtained a result of 100% "Very Feasible", while the user test using the PSSUQ questionnaire obtained an average of 90.4% "Very Feasible". The cluster validity test using the Silhouette Coefficient of 0.40 included in the category of "Weak Structure".

Keywords: Counseling, Decision Support System, Distribution Mapping, K-Means Clustering Algorithm, Toddler Nutritional Status, and Z-Score.

ABSTRAK

Pemetaan persebaran status gizi balita merupakan kegiatan pengelompokan wilayah posyandu berdasarkan kondisi gizi kurang maupun gizi lebih yang masih cukup tinggi. Permasalahan yang dihadapi tidak hanya tingginya kasus gizi kurang akibat asupan gizi yang tidak mencukupi, tetapi juga meningkatnya kasus gizi lebih yang dipicu pola konsumsi tidak seimbang. Namun, hingga saat ini data yang tersedia belum dimanfaatkan secara optimal untuk menghasilkan informasi pemetaan yang akurat. Hal ini menyebabkan pihak puskesmas kesulitan menentukan sasaran penyuluhan gizi yang sesuai dengan kebutuhan di lapangan. Selain itu, keterbatasan analisis manual membuat identifikasi wilayah posyandu dengan kasus gizi kurang atau gizi lebih menjadi kurang efektif. Persebaran status gizi balita yang belum akurat dapat menyebabkan belum tercapainya target penyuluhan gizi agar lebih tepat sasaran. Tujuan penelitian ini adalah untuk pemetaan persebaran status gizi balita untuk penentuan penyuluhan dengan pendekatan K-Means Clustering. Penelitian ini menghasilkan prototype sistem yang dapat melakukan pemetaan status gizi balita menggunakan Algoritma K-Means Clustering dengan menganalisis data awal, mentransformasi dan menormalisasi data, serta melakukan perhitungan pengelompokan. Variabel yang digunakan meliputi usia, berat badan, tinggi badan, z score bb/u, z score tb/u, dan z score bb/tb. Hasil pengelompokan menginformasikan wilayah posyandu pada setiap kelompok berdasarkan status gizi balita. Hal ini dilakukan untuk memetakan penyuluhan gizi melalui sistem pendukung keputusan sebagai acuan Puskesmas Tanah Sareal Kota Bogor dalam menentukan sasaran penyuluhan yang efektif. Penerapan Algoritma K-Means Clustering menghasilkan Cluster 1 sebanyak 276 dan Cluster 2 sebanyak 220 dengan kategori Gizi Kurang dan Gizi Lebih. Uji kelayakan oleh ahli sistem memperoleh hasil 100% "Sangat Layak", sedangkan uji pengguna menggunakan kuesioner PSSUQ memperoleh rata-rata 90,4% "Sangat Layak". Uji validitas cluster menggunakan Silhouette Coefficient sebesar 0,40.

Keywords: Algoritma K-Means, Klasterisasi, Pemberian Penyuluhan, Penyakit Diabetes, Persebaran

A. PENDAHULUAN

1. Latar Belakang Masalah

Kesehatan balita merupakan hal yang sangat penting karena balita berada pada masa pertumbuhan dan perkembangan yang membutuhkan perhatian khusus melalui pemenuhan gizi seimbang, pola makan sehat, aktivitas fisik, serta lingkungan yang bersih (Istiqomah et al., 2024). Namun, masalah gizi pada balita di Indonesia masih menjadi perhatian serius karena tingginya kasus gizi kurang, gizi lebih, dan beban gizi ganda yang berdampak pada pertumbuhan serta perkembangan anak (Fitri et al., 2022). Namun, hingga saat ini belum dapat terlihat secara jelas persebaran status gizi balita (gizi kurang, normal, dan gizi lebih) di setiap wilayah. Akibatnya, pemberian penyuluhan dan edukasi gizi yang bertujuan untuk meningkatkan kesadaran masyarakat dalam pencegahan dan penanganan masalah gizi belum berjalan secara optimal, karena belum disesuaikan dengan kondisi persebaran gizi di tiap posyandu. Dengan memanfaatkan teknik komputasi seperti algoritma K-Means Clustering, proses pengelompokan dan pemetaan persebaran status gizi balita dapat dilakukan secara lebih efektif dan efisien (Wijaksana et al., 2025).

2. Permasalahan

Dalam melakukan kegiatan penyuluhan maupun sosialisasi kepada masyarakat terkait pencegahan dan penanggulangan masalah gizi pada balita, terkadang pemilihan lokasi belum sepenuhnya sesuai dengan tingkat kerawanan kasus gizi di wilayah tersebut. Menjadikan pada saat ini masih belum dapat diketahui tingkat persebaran mengenai pemetaan pemberian penyuluhan terhadap masalah gizi balita. Dalam hal ini dapat mempengaruhi langkah pihak terkait dalam melakukan upaya pencegahan dan penanganan pemberian penyuluhan terhadap masalah gizi balita, sehingga diperlukan pendekatan berbasis data seperti metode K-Means *Clustering* untuk membantu memetakan status gizi balita secara lebih efektif dan efisien agar penyuluhan dapat dilakukan secara tepat sasaran dan mampu meningkatkan efektivitas penanganan gizi balita di setiap posyandu.

Tabel 1 Data Sampel Permasalahan Status Gizi Balita pada Puskesmas Tanah Sareal

No	Nama	JK	Posyandu	Usia (Bulan)	Berat	Tinggi	ZS BB/U	ZS TB/U	ZS BB/TB
1.	MUHAMMAD FAQIH SURYA PRAMA	L	DEWI SARTIKA	29	10.5	85.3	-1.92	-1.77	-1.4
2.	NADIA	P	CUT NYA DIEN	22	10.8	84	-0.3	-0.39	-0.18
3.	FALISHA	P	CUT NYA DIEN	30	11.7	87	-0.71	-1.11	-0.19
4.	EL SHANUM	P	CUT NYA DIEN	3	5.3	60	-1.32	-0.65	-1.15
5.	ARABELLA Q. B	P	CUT NYA DIEN	5	6.2	63	-1.25	-1.05	-0.72
6.	RIZ Aidan ZAFRAN FATHARIANO	L	CUT NYA DIEN	10	9.6	74	0.41	0.28	0.38
7.	IKRIMA HUSNA A	P	CUT NYA DIEN	13	7.7	72	-1.55	-1.43	-1.21
8.	ABDUL ROHIM	L	CUT NYA DIEN	15	10.6	80	0.17	0.15	0.17
9.	SYAIRA CAHYA	L	CUT NYA DIEN	16	8.5	79	-2.1	-0.86	-2.35
10.	IMANUEL HERMAWAN	L	CUT NYA DIEN	17	8.5	75	-2.18	-2.54	-1.37
...
487.	M. UMAR MUZAHIR	L	HARAPAN MASA	25	10	85	-1.95	-1.16	-1.95
488.	M. SYAMIL	L	HARAPAN MASA	48	14.7	95.7	-0.83	-1.8	0.4
489.	ANINDYA	P	HARAPAN MASA	17	11.5	80.4	0.98	0	1.34
490.	RAYN	L	HARAPAN MASA	46	13.6	97	-1.33	-1.3	-0.85
491.	NAZRIL AHMAD	L	CUT NYA DIEN	0	2.9	49	-0.96	-0.47	-0.87
492.	M. ZAIN AKBAR	L	CUT NYA DIEN	0	3	50	-2.27	-1.54	-1.7
493.	KALISTA	P	DEWI SARTIKA	0	3	50	-1.56	-0.78	-1.74
494.	ALEXA	P	TERATAI	0	3	50	-2.27	-1.44	-1.74
495.	NAZEEB	L	DAHLIA	0	3	50	-1.95	-1.25	-1.7
496.	DAFA	L	ANGGREK	16	9.2	77.3	-1.33	-1.37	-0.96

Berdasarkan dari permasalahan yang diuraikan tersebut, maka dapat diidentifikasi masalah yaitu, belum akurat dalam menentukan pemberian penyuluhan berdasarkan status gizi balita dan belum efektif dalam proses menentukan pemberian penyuluhan berdasarkan status gizi balita.

3. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut, mendapatkan pemetaan penentuan pemberian penyuluhan berdasarkan status gizi balita, mendapatkan proses yang lebih efektif untuk penentuan pemberian penyuluhan, mengembangkan prototype untuk penentuan pemberian penyuluhan dan mengukur tingkat efektivitas penerapan algoritma K-Means dalam memetakan persebaran penyuluhan.

4. Tinjauan Pustaka

a. Pengertian Data Mining

Menurut (Rahayu et al., 2024, p. 2) data *mining* adalah proses memperoleh pengetahuan atau informasi berharga dari kumpulan data yang besar dan kompleks. (Jollyta et al., 2020, pp. 119–120) mengutarakan bahwa data *mining* adalah proses mengidentifikasi pola, tren, dan informasi berharga dari data yang besar proses logis untuk menemukan informasi dan pola yang berguna dari kumpulan data yang sangat besar melalui pengumpulan, ekstraksi, analisis, dan penerapan statistik, yang juga dikenal sebagai *Knowledge Discovery, Knowledge Extraction, data/pattern analysis, and information harvesting. Knowledge Discovery in Database* merupakan sebuah basis data, yang mengungkap pengetahuan dalam basis data, yang melibatkan ekstraksi atau identifikasi pola, pengetahuan, dan informasi dari kumpulan data besar (Sinurat, 2024, p. 4). Menurut (Habibi and Suryansah, 2020, p.7) KDD merupakan proses yang kompleks untuk mencari dan mengidentifikasi pola dalam data, di mana pola yang ditemukan harus valid, inovatif, bermanfaat, dan mudah dipahami.

b. Metode CRISP-DM

(Muttaqin et al., 2023, p. 18) mengutarakan bahwa Cross-Industry Standard Process for Data Mining atau CRISP-DM adalah salah satu model proses data mining (data mining framework) yang awalnya (1996) dibangun oleh 5 perusahaan yaitu Integral Solutions Ltd (ISL), Teradata, Daimler AG, NCR Corporation dan OHRA. Menurut (Muttaqin et al., 2023, p. 19) CRISP-DM merupakan model data mining model yang masih banyak diterapkan di kalangan industri secara luas, salah satunya karena kemampuannya dalam mengatasi berbagai masalah di proyek-proyek data mining. Berikut enam tahapan CRISP-DM dalam (Sumadireja et al., 2020, p. 68–70):

- (1) business understanding, pada tahapan ini mencakup penentuan tujuan bisnis, evaluasi situasi saat ini, penetapan tujuan data mining, dan pengembangan rencana penelitian;
- (2) data understanding, pada tahapan ini merupakan tahap pemahaman data yang diperoleh melalui observasi langsung dan analisis terhadap berkas data pelanggan pada objek penelitian;
- (3) data preparation, pada tahapan ini merupakan tahap pengolahan data atau disebut juga tahapan persiapan data;
- (4) modelling, pada tahapan ini dilakukan pengaplikasian teknik pemodelan, data yang dipilih pada fase persiapan digunakan sebagai parameter untuk melakukan peramalan dan optimasi;
- (5) evaluation, pada tahapan ini diperlukan analisis yang menginterpretasikan hasil pengolahan data sebelumnya untuk menghasilkan peramalan dan optimasi nilai;
- (6) deployment, pada tahapan ini mengatur dan mempresentasikan pengetahuan atau informasi yang diperoleh dalam format khusus agar dapat digunakan oleh pengguna.

c. Algoritma K-Means

Menurut (Arhami and Nasir, 2020, p.148) algoritma merupakan salah satu metode *clustering* yang sering digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik atau ciri-ciri yang serupa. Kelompok data yang terbentuk disebut *cluster*. Algoritma ini juga efisien dalam hal komputasi dan menghasilkan hasil yang baik jika *cluster - cluster* yang terbentuk padat, berbentuk hampir bulat, dan dapat memisahkan fitur-fitur ruang dengan baik. K-Means adalah algoritma yang sederhana, mudah dimengerti, dan mudah disesuaikan untuk mengolah data streaming. K-Means juga mudah diimplementasikan dan memiliki kompleksitas waktu serta ruang yang relatif rendah.

d. Silhouette Coefficient

Menurut (Prihandoko et al., 2024, p. 31) Silhouette Coefficient adalah metrik yang digunakan untuk menilai seberapa efektif objek dikelompokkan dalam cluster. (Ma'ady et al., 2024, p. 152) mengutarakan bahwa Silhouette Coefficient adalah teknik yang digunakan untuk menentukan jumlah cluster optimal. Nilai Silhouette Coefficient tertinggi diperoleh dari penjumlahan nilai s yang telah dikonversi, dan jika jumlah tersebut lebih besar dibandingkan cluster lainnya, maka cluster tersebut dianggap terbaik karena memiliki lebih sedikit nilai s di bawah 0 (Mulaab, 2021, p. 87).

B. METODE

1. Algoritma K-Means

Dalam penelitian ini, pemetaan persebaran status gizi balita untuk penentuan pemberian penyuluhan gizi diperlukan sebuah algoritma yang mampu mengelompokkan data berdasarkan kesamaan karakteristik. Dengan algoritma K-Means, data balita dapat dibagi ke dalam kelompok yang memiliki ciri atau karakteristik serupa, sedangkan data dengan karakteristik berbeda akan dikelompokkan ke dalam set yang berbeda pula. Proses ini bertujuan untuk meminimalkan variasi antar data dalam satu kelompok sehingga hasil pengelompokan menjadi lebih akurat dan dapat digunakan sebagai dasar dalam menentukan sasaran penyuluhan gizi yang tepat (Arhami and Nasir, 2020, p.148).

Menurut (Jollyta et al., 2020, p. 119–120) berikut ini langkah-langkah melakukan *clustering* dengan metode K-Means yaitu:

- (1) pilih jumlah *cluster* K ;
- (2) inialisasi K pusat *cluster* umumnya dilakukan secara acak dengan memberikan nilai awal berupa angka random;
- (3) alokasikan setiap data ke *cluster* terdekat berdasarkan jarak ke pusat *cluster*, menggunakan rumus *Euclidean* untuk menentukan data masuk ke *cluster* mana, dengan rumus:

$$D_{i,j} = \sqrt{((x^1_i - x^1_j)^2 + (x^2_i - x^2_j)^2 + \dots + (x^k_i - x^k_j)^2)}$$

dimana:

$D(i,j)$ = jarak data ke i ke pusat *cluster* j ;

$X(k,i)$ = data ke i pada atribut data ke k ;
 $X(k,j)$ = titik pusat ke j pada atribut ke k ;

- (4) hitung ulang pusat *cluster* berdasarkan anggota saat ini, dengan menggunakan rata-rata atau median data dalam *cluster* tersebut;
- (5) ulangi langkah ke 3 dan 4, apabila alokasi data *cluster* tidak berubah lagi maka proses *clustering* selesai.

2. Teknik Analisis Data

Pengujian model yang dilakukan untuk menekankan kedekatan relasi antar objek dan mengetahui seberapa jauh cluster terpisah dengan cluster lainnya dan kombinasi dari proses agregasi dan pemisahan disebut sebagai *Silhouette Coefficient*. Uji hasil pada penelitian ini menggunakan *Silhouette Coefficient* karena kemampuannya dalam memberikan evaluasi kualitas pengelompokan data serta membantu menentukan jumlah *cluster* optimal sehingga dapat mengambil keputusan yang terinformasi mengenai pengelompokan. Berikut tahapan perhitungan *Silhouette Coefficient* menurut (Mulaab, 2021) sebagai berikut;

- (1) untuk setiap objek i , dihitung rata-rata jarak dari objek i dengan seluruh objek lain yang berada dalam satu *cluster. Akan didapatkan nilai rata-rata yang disebut a_i ;*
- (2) untuk setiap objek i , dihitung rata-rata jarak dari objek i dengan objek yang berada di *cluster* lainnya. Dari semua jarak rata-rata tersebut ambil nilai yang paling kecil. Nilai ini disebut b_i ;
- (3) setelah itu, untuk objek i memiliki nilai *Silhouette Coefficient*;

$$S_i = \frac{(b_i - a_i)}{\max(a_i, b_i)}$$

Untuk menilai *Silhouette Coefficient* dapat dilihat pada tabel 2 sebagai berikut:

Tabel 2 Nilai Silhouette Coefficient

No	Rentang Nilai <i>Silhouette Coefficient</i>	Struktur
1.	$0,7 < SC \leq 1$	<i>Strong Structure</i>
2.	$0,5 < SC \leq 0,7$	<i>Medium Structure</i>
3.	$0,25 < SC \leq 0,5$	<i>Weak Structure</i>
4.	$SC \leq 0,25$	<i>No Structure</i>

Sumber : (Rangkuti et al., 2022, p. 30)

C. HASIL DAN PEMBAHASAN

1. Hasil

a. Penentuan Variabel Penelitian

Variabel yang digunakan dalam pemetaan persebaran status gizi balita untuk penentuan pemberian penyuluhan adalah usia, berat badan, tinggi badan, z score bb/u, z score tb/u dan z score bb/tb. Tabel 3 menjelaskan definisi tiap variabel yang digunakan dalam pemetaan persebaran status gizi balita untuk penentuan pemberian penyuluhan.

Tabel 3 Variabel Penelitian

No	Variabel	Definisi
1.	Usia	Menyatakan umur balita dalam bulan yang dihitung sejak tanggal lahir hingga hari pengukuran dilakukan.
2.	Berat Badan	Menyatakan massa tubuh balita yang diukur dalam satuan kilogram, digunakan sebagai indikator utama untuk menilai pertumbuhan dan status gizi.
3.	Tinggi Badan	Menyatakan ukuran tubuh balita dari tumit hingga puncak kepala dalam posisi berdiri (untuk balita ≥ 24 bulan) atau posisi berbaring (untuk balita < 24 bulan).
4.	Z Score BB/U	Menyatakan ukuran statistik yang menunjukkan sejauh mana berat badan anak dibanding rata-rata seusianya untuk menilai status gizi.
5.	Z Score TB/U	Menyatakan ukuran untuk menilai apakah tinggi badan anak sesuai

No	Variabel	Definisi
		dengan usianya untuk menilai status gizi.
6.	Z Score BB/TB	Menyatakan ukuran statistik yang menunjukkan sejauh mana tinggi badan anak dibanding berat badannya untuk menilai status gizi.

b. Data Transformasi dan Normalisasi

Diperlukan transformasi data pada beberapa atribut, yaitu usia, berat badan dan tinggi badan. Setelah dilakukan transformasi data maka langkah selanjutnya yaitu normalisasi dengan menggunakan *decimal scaling*. Adapun proses perhitungan *scaling* untuk normalisasi data menurut (Han, Kamber, & Pei, 2011) adalah sebagai berikut:

$$new\ data = \frac{data}{i}$$

Keterangan:

- New data = data hasil normalisasi
- Data = data yang digunakan
- i = nilai maksimal dari kriteria

Contoh perhitungan *decimal scaling* untuk normalisasi data pada atribut usia, berat badan dan tinggi badan yaitu:

- Nilai Maksimum (Xmaks) pada data Usia = 60
- Nilai Maksimum (Xmaks) pada data Berat = 29
- Nilai Maksimum (Xmaks) pada data Tinggi = 120

Data ke-1

$$usia = \frac{29}{60} = 0,48$$

$$berat = \frac{10,5}{29} = 0,36$$

$$tinggi = \frac{85,3}{120} = 0,71$$

Data ke-2

$$usia = \frac{22}{60} = 0,37$$

$$berat = \frac{10,8}{29} = 0,37$$

$$tinggi = \frac{84}{120} = 0,70$$

Data ke-3

$$usia = \frac{30}{60} = 0,50$$

$$berat = \frac{11,7}{29} = 0,40$$

$$tinggi = \frac{87}{120} = 0,73$$

Data ke-4

$$\text{usia} = \frac{3}{60} = 0,05$$

$$\text{berat} = \frac{5,3}{29} = 0,18$$

$$\text{tinggi} = \frac{60}{120} = 0,50$$

Data ke-5

$$\text{usia} = \frac{5}{60} = 0,08$$

$$\text{berat} = \frac{6,2}{29} = 0,21$$

$$\text{tinggi} = \frac{63}{120} = 0,53$$

...

Data ke-496

$$\text{usia} = \frac{16}{60} = 0,050$$

$$\text{berat} = \frac{5,8}{29} = 0,200$$

$$\text{tinggi} = \frac{62}{120} = 0,517$$

Tabel 4 merupakan data hasil normalisasi variabel usia, berat badan dan tinggi badan.

Tabel 4 Normalisasi Variabel

No	Normalisasi Usia	Normalisasi Berat	Normalisasi Tinggi
1.	0,48	0,36	0,71
2.	0,37	0,37	0,70
3.	0,50	0,40	0,73
4.	0,05	0,18	0,50
5.	0,08	0,21	0,53
6.	0,17	0,33	0,62
7.	0,22	0,27	0,60
8.	0,25	0,37	0,67
9.	0,27	0,29	0,66
10.	0,28	0,29	0,63
...
487.	0,42	0,34	0,71
488.	0,80	0,51	0,80
489.	0,28	0,40	0,67
490.	0,77	0,47	0,81
491.	0,00	0,10	0,41
492.	0,00	0,10	0,42
493.	0,00	0,10	0,42
494.	0,00	0,10	0,42
495.	0,00	0,10	0,42
496.	0,27	0,32	0,64

c. Proses Algoritma K-Means

Dari data yang tertera pada tabel 4 akan dibuat kluster pemetaan persebaran penyakit diabetes untuk rekomendasi prioritas pemberian penyuluhan berdasarkan variabel yang telah ditentukan. Secara umum Algoritma K-Means bertujuan untuk membagi data menjadi beberapa kelompok dengan tahapan sebagai berikut:

(1) Menyiapkan Dataset

Data yang digunakan pada perhitungan ini menggunakan variabel usia, berat badan, tinggi badan, z score bb/u, z score tb/u dan z score bb/tb berdasarkan data status gizi balita tahun 2024 yang diperoleh dari Puskesmas Tanah Sareal Kota Bogor. Dataset pada Tabel 5 sebagai berikut;

Tabel 5 Dataset Perhitungan K-Means

No	Nama	Posyandu	Usia (Bulan)	Berat	Tinggi	ZS BB/U	ZS TB/U	ZS BB/TB
1.	MUHAMMAD FAQIH SURYA PRAMA	DEWI SARTIKA	0,48	0,36	0,71	-1,92	-1,77	-1,4
2.	NADIA	CUT NYA DIEN	0,37	0,37	0,70	-0,3	-0,39	-0,18
3.	FALISHA	CUT NYA DIEN	0,50	0,40	0,73	-0,71	-1,11	-0,19
4.	EL SHANUM	CUT NYA DIEN	0,05	0,18	0,50	-1,32	-0,65	-1,15
5.	ARABELLA Q. B	CUT NYA DIEN	0,08	0,21	0,53	-1,25	-1,05	-0,72
6.	RIZ Aidan ZAFRAN FATHARIANO	CUT NYA DIEN	0,17	0,33	0,62	0,41	0,28	0,38
7.	IKRIMA HUSNA A	CUT NYA DIEN	0,22	0,27	0,60	-1,55	-1,43	-1,21
8.	ABDUL ROHIM	CUT NYA DIEN	0,25	0,37	0,67	0,17	0,15	0,17
9.	SYAIRA CAHYA	CUT NYA DIEN	0,27	0,29	0,66	-2,1	-0,86	-2,35
10.	IMANUEL HERMAWAN	CUT NYA DIEN	0,28	0,29	0,63	-2,18	-2,54	-1,37
...
487.	M. UMAR MUZAHIR	HARAPAN MASA	0,42	0,34	0,71	-1,95	-1,16	-1,95
488.	M. SYAMIL	HARAPAN MASA	0,80	0,51	0,80	-0,83	-1,8	0,4
489.	ANINDYA	HARAPAN MASA	0,28	0,40	0,67	0,98	0	1,34
490.	RAYN	HARAPAN MASA	0,77	0,47	0,81	-1,33	-1,3	-0,85
491.	NAZRIL AHMAD	CUT NYA DIEN	0,00	0,10	0,41	-0,96	-0,47	-0,87
492.	M. ZAIN AKBAR	CUT NYA DIEN	0,00	0,10	0,42	-2,27	-1,54	-1,7
493.	KALISTA	DEWI SARTIKA	0,00	0,10	0,42	-1,56	-0,78	-1,74
494.	ALEXA	TERATAI	0,00	0,10	0,42	-2,27	-1,44	-1,74
495.	NAZEEB	DAHLIA	0,00	0,10	0,42	-1,95	-1,25	-1,7
496.	DAFA	ANGGREK	0,27	0,32	0,64	-1,33	-1,37	-0,96

(2) Menentukan Jumlah Kluster

Dari dataset pada Tabel 4.4 akan dibentuk 2 kluster sehingga nilai K pada perhitungan ini yaitu 2, dengan keterangan cluster 1 gizi kurang dan cluster 2 gizi lebih.

(3) Menentukan Titik Centroid

Dataset penimbangan bayi tahun 2024 Puskesmas Tanah Sareal Kota Bogor akan dibagi ke dalam 2 kluster sehingga diperlukan 2 titik awal centroid yang digunakan yaitu sebagai berikut. Menentukan pusat awal cluster dengan rumus sebagai berikut:

$$\frac{n \text{ Data}}{n \text{ Cluster} + 1}$$

Diperoleh titik centroid cluster 1 dan cluster 2 sebagai berikut:

$$C1 = \text{Jumlah data} / C1+1 = 496 / 1+1 = 248$$

$$C2 = \text{Jumlah data} / C2+1 = 496 / 2+1 = 165$$

yang berarti titik centroid C1 diambil dari data ke-248 dan titik C2 diambil dari data ke-165, Titik centroid awal yang digunakan yaitu seperti pada Tabel 4.6.

Tabel 6 Nilai Centroid Awal

Centroid	A	B	C	D	E	F
C1	0,65	0,46	0,79	-0,77	-0,69	-0,57
C2	0,77	0,68	0,86	1,58	0,13	2,23

(4) Hitung Jarak Dengan Centroid

Perhitungan jarak data dengan centroid dilakukan untuk menentukan jarak terpendek yang akan digunakan untuk menentukan pengelompokan klaster. Euclidean Distance digunakan untuk menghitung jarak data dengan centroid. Euclidean Distance merupakan perhitungan jarak dari 2 titik dan untuk mempelajari hubungan antara sudut dan jarak. Rumus Euclidean Distance adalah sebagai berikut :

$$D_{i,j} = \sqrt{((x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2)}$$

dimana:

$D(i,j)$ = jarak data ke i ke pusat *cluster* j ;

$X(k,i)$ = data ke i pada atribut data ke k ;

$X(k,j)$ = titik pusat ke j pada atribut ke k ;

Iterasi 1

Data ke 1:

$$D(i_1, j_1)$$

$$C1 = \sqrt{\left(\begin{aligned} &(0,48 - 0,65)^2 + (0,36 - 0,46)^2 + (0,71 - 0,79)^2 \\ &+ (-1,92 - (-0,77))^2 + (-1,77 - (-0,69))^2 + (-1,4 - (-0,57))^2 \end{aligned} \right)}$$

$$= 1,79$$

$$C2 = \sqrt{\left(\begin{aligned} &(0,48 - 0,77)^2 + (0,36 - 0,68)^2 + (0,71 - 0,86)^2 \\ &+ (-1,92 - 1,58)^2 + (-1,77 - 0,13)^2 + (-1,4 - 2,23)^2 \end{aligned} \right)}$$

$$= 5,41$$

Data ke 2:

$$D(i_2, j_2)$$

$$C1 = \sqrt{\left(\begin{aligned} &(0,37 - 0,65)^2 + (0,37 - 0,46)^2 + (0,70 - 0,79)^2 \\ &+ (-0,3 - (-0,77))^2 + (-0,39 - (-0,69))^2 + (-0,18 - (-0,57))^2 \end{aligned} \right)}$$

$$= 0,75$$

$$C2 = \sqrt{\left(\begin{array}{l} (0,37 - 0,77)^2 + (0,37 - 0,68)^2 + (0,70 - 0,86)^2 \\ +(-0,3 - 1,58)^2 + (-0,39 - 0,13)^2 + (-0,18 - 2,23)^2 \end{array} \right)}$$

$$= 3,15$$

Data ke 3:

$D(i_3, j_3)$

$$C1 = \sqrt{\left(\begin{array}{l} (0,50 - 0,65)^2 + (0,40 - 0,46)^2 + (0,73 - 0,79)^2 \\ +(-0,71 - (-0,77))^2 + (-1,11 - (-0,69))^2 + (-0,19 - (-0,57))^2 \end{array} \right)}$$

$$= 0,59$$

$$C2 = \sqrt{\left(\begin{array}{l} (0,50 - 0,77)^2 + (0,40 - 0,68)^2 + (0,73 - 0,86)^2 \\ +(-0,71 - 1,58)^2 + (-1,11 - 0,13)^2 + (-0,19 - 2,23)^2 \end{array} \right)}$$

$$= 3,58$$

Data ke 4:

$D(i_4, j_4)$

$$C1 = \sqrt{\left(\begin{array}{l} (0,05 - 0,65)^2 + (0,18 - 0,46)^2 + (0,50 - 0,79)^2 \\ +(-1,32 - (-0,77))^2 + (-0,65 - (-0,69))^2 + (-1,15 - (-0,57))^2 \end{array} \right)}$$

$$= 1,08$$

$$C2 = \sqrt{\left(\begin{array}{l} (0,05 - 0,77)^2 + (0,18 - 0,68)^2 + (0,50 - 0,86)^2 \\ +(-1,32 - 1,58)^2 + (-0,65 - 0,13)^2 + (-1,15 - 2,23)^2 \end{array} \right)}$$

$$= 4,62$$

Data ke 5:

$D(i_5, j_5)$

$$C1 = \sqrt{\left(\begin{array}{l} (0,08 - 0,65)^2 + (0,21 - 0,46)^2 + (0,53 - 0,79)^2 \\ +(-1,25 - (-0,77))^2 + (-1,05 - (-0,69))^2 + (-0,72 - (-0,57))^2 \end{array} \right)}$$

$$= 0,91$$

$$C2 = \sqrt{\left(\begin{array}{l} (0,08 - 0,77)^2 + (0,21 - 0,68)^2 + (0,53 - 0,86)^2 \\ +(-1,25 - 1,58)^2 + (-1,05 - 0,13)^2 + (-0,72 - 2,23)^2 \end{array} \right)}$$

$$= 4,35$$

...

Data ke 496:

$D(i_{20}, j_{20})$

$$C1 = \sqrt{\left(\begin{array}{l} (0,27 - 0,65)^2 + (0,32 - 0,46)^2 + (0,64 - 0,79)^2 \\ + (-1,33 - (-0,77))^2 + (-1,37 - (-0,69))^2 + (-0,96 - (-0,57))^2 \end{array} \right)}$$

$$= 1,06$$

$$C2 = \sqrt{\left(\begin{array}{l} (0,27 - 0,77)^2 + (0,32 - 0,68)^2 + (0,64 - 0,86)^2 \\ + (-1,33 - 1,58)^2 + (-1,37 - 0,13)^2 + (-0,96 - 2,23)^2 \end{array} \right)}$$

$$= 4,62$$

Tabel 7 Hasil Pengelompokan Iterasi 1

No	C1	C2	Jarak Terdekat	Cluster
1.	1,79	5,41	1,79	C1
2.	0,75	3,15	0,75	C1
3.	0,59	3,58	0,59	C1
4.	1,08	4,62	1,08	C1
5.	0,91	4,35	0,91	C1
6.	1,87	2,31	1,87	C1
7.	1,35	4,96	1,35	C1
8.	1,52	2,58	1,52	C1
9.	2,27	5,99	2,27	C1
10.	2,51	5,89	2,51	C1
...
487.	1,90	5,64	1,90	C1
488.	1,48	3,59	1,48	C1
489.	2,71	1,23	1,23	C2
490.	0,88	4,48	0,88	C1
491.	0,93	4,19	0,93	C1
492.	2,22	5,85	2,22	C1
493.	1,64	5,25	1,64	C1
494.	2,21	5,84	2,21	C1
495.	1,92	5,56	1,92	C1
496.	1,06	4,62	1,06	C1

Tabel 7 merupakan hasil perhitungan dengan menggunakan rumus *Euclidean Distance* antara tiap variabel dengan titik centroid 1 dan centroid 2 sehingga diperoleh jarak terdekat dari masing-masing atribut data ke-n, kemudian digunakan untuk ditentukan masuk ke *cluster* masing-masing. Setelah proses pengelompokan data lalu dilakukan penentuan titik centroid baru dengan menggunakan rumus:

$$\mu_k = \frac{1}{N_k} \sum_{t=1}^{N_k} x_i$$

Keterangan :

μ_k : Titik centroid dari *cluster* ke-K

N_k : Banyaknya data pada *cluster* ke-K

x_i : Data ke-i *cluster* ke-K

Nilai centroid baru pada centroid 1 untuk iterasi 2 :

$$A = \frac{209,20}{384} = 0,54 \quad B = \frac{159,63}{384} = 0,42 \quad C = \frac{282,91}{384} = 0,74$$

$$D = \frac{-320,00}{384} = -0,83 \quad E = \frac{-333,16}{384} = -0,87 \quad F = \frac{-194,20}{384} = -0,51$$

Nilai centroid baru pada centroid 2 untuk iterasi 2 :

$$A = \frac{63,57}{112} = 0,57 \quad B = \frac{61,26}{112} = 0,55 \quad C = \frac{86,71}{112} = 0,77$$

$$D = \frac{121,17}{112} = 1,08 \quad E = \frac{-17,03}{112} = -0,15 \quad F = \frac{187,17}{112} = 1,67$$

Tabel 8 Nilai Centroid Baru

Centroid	A	B	C	D	E	F
C1	0,65	0,46	0,79	-0,77	-0,69	-0,57
C2	0,77	0,68	0,86	1,58	0,13	2,23

Ulangi perhitungan jarak data dengan centroid hingga nilai titik centroid tidak berubah. Iterasi kedua dapat dilihat pada perhitungan berikut. Dari perhitungan yang dilakukan pada penelitian ini terjadi sebanyak 10 kali, titik pusat cluster tidak ada lagi perubahan dan data pertama hingga data yang terakhir tidak terdapat lagi yang bergeser dari satu cluster ke cluster yang lainnya. Dari perhitungan yang sudah dilakukan maka dapat disimpulkan bahwa hasil *clustering* dari dataset perhitungan K-Means menghasilkan pembagian objek ke masing-masing kluster, yaitu:

(1) Cluster 1 Gizi Kurang

Cluster 1 terdiri dari 276 data balita dan 10 data posyandu seperti pada Tabel 9 berikut:

Tabel 9 Hasil Cluster 1

No	Nama	Posyandu	Usia (Bulan)	Berat	Tinggi	BB/U	TB/U	BB/TB	Cluster
1.	MUHAMMAD FAQIH SURYA PRAMA	DEWI SARTIKA	29	10.5	85.3	-1.92	-1.77	-1.4	C1
2.	NADIA	CUT NYA DIEN	22	10.8	84	-0.3	-0.39	-0.18	C1
3.	FALISHA	CUT NYA DIEN	30	11.7	87	-0.71	-1.11	-0.19	C1
4.	EL SHANUM	CUT NYA DIEN	3	5.3	60	-1.32	-0.65	-1.15	C1
5.	ARABELLA Q. B	CUT NYA DIEN	5	6.2	63	-1.25	-1.05	-0.72	C1
6.	IKRIMA HUSNA A	CUT NYA DIEN	13	7.7	72	-1.55	-1.43	-1.21	C1
7.	SYAIRA CAHYA	CUT NYA DIEN	16	8.5	79	-2.1	-0.86	-2.35	C1
8.	IMANUEL HERMAWAN	CUT NYA DIEN	17	8.5	75	-2.18	-2.54	-1.37	C1
9.	M. ZAKI ABDULLAH	CUT NYA DIEN	20	10.3	80	-0.98	-1.71	-0.17	C1
10.	MUHAMMAD ALVIANO HERMAWAN	CUT NYA DIEN	20	10.3	84	-1	-0.36	-1.11	C1
...
487.	AR RAYAN TERANG PERDANA	HARAPAN MASA	57	17	107.8	-0.38	-0.13	-0.51	C1
488.	M. UMAR MUZAHIR	HARAPAN MASA	25	10	85	-1.95	-1.16	-1.95	C1
489.	M. SYAMIL	HARAPAN MASA	48	14.7	95.7	-0.83	-1.8	0.4	C1
490.	RAYN	HARAPAN MASA	46	13.6	97	-1.33	-1.3	-0.85	C1
491.	NAZRIL AHMAD	CUT NYA DIEN	0	2.9	49	-0.96	-0.47	-0.87	C1
492.	M. ZAIN AKBAR	CUT NYA DIEN	0	3	50	-2.27	-1.54	-1.7	C1
493.	KALISTA	DEWI	0	3	50	-1.56	-0.78	-1.74	C1

No	Nama	Posyandu	Usia (Bulan)	Berat	Tinggi	BB/U	TB/U	BB/TB	Cluster
		SARTIKA							
494.	ALEXA	TERATAI	0	3	50	-2.27	-1.44	-1.74	C1
495.	NAZEEB	DAHLIA	0	3	50	-1.95	-1.25	-1.7	C1
496.	DAFA	ANGGREK	16	9.2	77.3	-1.33	-1.37	-0.96	C1

(2) Cluster 2 Gizi Lebih

Cluster 2 terdiri dari 220 data balita dan 10 data posyandu seperti pada Tabel 10 berikut:

Tabel 10 Hasil Cluster 1

No	Nama	Posyandu	Usia (Bulan)	Berat	Tinggi	BB/U	TB/U	BB/TB	Cluster
1.	RIZ Aidan ZAFRAN FATHARIANO	CUT NYA DIEN	10	9,6	74	0,41	0,28	0,38	C2
2.	ABDUL ROHIM	CUT NYA DIEN	15	10,6	80	0,17	0,15	0,17	C2
3.	NALDA RASHIDA QURESHA	CUT NYA DIEN	21	12,4	83,5	0,99	0,02	1,29	C2
4.	SENADA ZEEYA	CUT NYA DIEN	24	12,8	84	0,76	-0,73	1,51	C2
5.	JENNAIRA	CUT NYA DIEN	30	13	93	0,07	0,46	-0,33	C2
6.	ANISA	CUT NYA DIEN	34	14,5	95	0,54	0,31	0,47	C2
7.	ALFA REZKY	CUT NYA DIEN	35	14	89	-0,18	-1,88	1,24	C2
8.	SHAHLA AULIA	TERATAI	11	10.1	77.6	1.18	1.89	0.53	C2
9.	ANINDYA JINAN	TERATAI	12	8.6	70	-0.34	-1.6	0.57	C2
10.	KHALISA MAHARANI	TERATAI	13	10	78	0.65	0.95	0.34	C2
...
487.	ADIBA SYAKIN	HARAPAN MASA	40	17.5	107	1.26	2.2	0.01	C2
488.	SYAKILA ARKAN DIMA	HARAPAN MASA	40	14.5	92.4	-0.05	-1.35	1	C2
489.	KALLEA	HARAPAN MASA	42	14.4	96	-0.28	-0.72	0.2	C2
490.	ABDULLAH FAUZAN AZHIMA	HARAPAN MASA	53	16.8	100.9	-0.2	-1.23	0.87	C2
491.	MICHIKO ARSA	HARAPAN MASA	57	18.6	110	0.29	0.38	0	C2
492.	KEANU KHALIF	HARAPAN MASA	53	18.2	105	0.4	-0.31	0.9	C2
493.	PUTU DANIEL SATYA WIRAGUNA	HARAPAN MASA	47	18.5	103	1.02	0.02	1.53	C2
494.	SAHANAYA SENJA NURHAFIDZAH	HARAPAN MASA	51	15.9	98.7	-0.32	-1.32	0.74	C2
495.	MUHAMAD REGAN ARSYA	HARAPAN MASA	43	15.2	98	-0.18	-0.64	0.31	C2
496.	ANINDYA	HARAPAN MASA	17	11.5	80.4	0.98	0	1.34	C2

2. Pembahasan

Berdasarkan hasil penerapan metode pada sistem Informasi dilakukan uji hasil dengan menggunakan metode Silhouette Coefficient. Metode pengujian hasil ini dilakukan bertujuan untuk menguji kualitas setiap cluster yang diperoleh dengan menggabungkan metode cohesion dan separation. Berikut adalah tahapan uji hasil

dengan menggunakan Silhouette Coefficient :

- (1) Menghitung jarak rata-rata dari suatu data terhadap semua data yang lain yang berada di dalam cluster yang sama, misalnya data ke-1 dengan semua data pada cluster tersebut dengan menghitung a(i) (average dissimilarity) menggunakan data yang ada pada tabel.

- (a) Data ke 1 terhadap semua data yang termasuk ke dalam *cluster 1*:

$$a(i)_1 = \sqrt{\frac{(0,48 - 0,48)^2 + (0,36 - 0,36)^2 + (0,71 - 0,71)^2 + (-1,92 - (-1,92))^2 + (-1,77 - (-1,77))^2 + (-1,4 - (-1,4))^2 + (0,48 - 0,37)^2 + (0,36 - 0,37)^2 + (0,71 - 0,70)^2 + (-1,92 - (-0,3))^2 + (-1,77 - (-0,39))^2 + (-1,4 - (-0,18))^2 \dots + (0,48 - 0,27)^2 + (0,36 - 0,32)^2 + (0,71 - 0,64)^2 + (-1,92 - (-1,33))^2 + (-1,77 - (-1,37))^2 + (-1,4 - (-0,96))^2}{276}}$$

$$a(i)_1 = \sqrt{\frac{444,22}{276}} = 1,61$$

- (b) Data ke 1 terhadap semua data yang termasuk ke dalam *cluster 2*:

$$a(i)_2 = \sqrt{\frac{(0,17 - 0,17)^2 + (0,33 - 0,33)^2 + (0,62 - 0,62)^2 + (0,41 - 0,41)^2 + (0,28 - 0,28)^2 + (0,38 - 0,38)^2 + (0,17 - 0,25)^2 + (0,33 - 0,37)^2 + (0,62 - 0,67)^2 + (0,41 - 0,17)^2 + (0,28 - 0,15)^2 + (0,38 - 0,17)^2 + \dots + (0,17 - 0,28)^2 + (0,33 - 0,40)^2 + (0,62 - 0,67)^2 + (0,41 - 0,98)^2 + (0,28 - 0)^2 + (0,38 - 1,34)^2}{220}}$$

$$a(i)_2 = \sqrt{\frac{365,52}{220}} = 1,66$$

Hasil perhitungan a(i) pada satu *cluster* dilihat pada Tabel 11

Tabel 11 Hasil Perhitungan a(i)

Data ke	a(i)
1.	1,61
2.	1,67
3.	1,34
4.	1,42
5.	1,30
6.	1,66
7.	1,37
8.	1,69
9.	2,10
10.	2,17
...	...
487.	1,75
488.	1,82
489.	
490.	1,22
491.	1,50
492.	1,96
493.	1,69
494.	1,96

Data ke	a(i)
495.	1,77
496.	1,56

- (2) Jarak rata-rata dihitung dari suatu data terhadap seluruh data yang lain yang berada pada *cluster* lain, misalnya data ke-1 dengan seluruh data pada *cluster* lain dengan menghitung $d(i,C)$ menggunakan data yang ada pada tabel *cluster* 1 ke *cluster* 2:

$$d(1,2) = \sqrt{\begin{aligned} &(0,17 - 0,48)^2 + (0,33 - 0,36)^2 + (0,62 - 0,71)^2 + \\ &(0,41 - (-1,92))^2 + (0,28 - (-1,77))^2 + \\ &(0,38 - (-1,4))^2 + (0,17 - 0,37)^2 + \\ &(0,33 - 0,37)^2 + (0,62 - 0,70)^2 + \\ &(0,41 - (-0,3))^2 + (0,28 - 0,39)^2 + \\ &(0,38 - 0,18)^2 + \dots (0,17 - 0,27)^2 + \\ &(0,33 - 0,32)^2 + (0,62 - 0,64)^2 + \\ &(0,41 - (-1,33))^2 + (0,28 - (-1,37))^2 + \\ &(0,38 - (-0,96))^2 \end{aligned}}$$

$$d(1,2) = \sqrt{\frac{871,28}{220}} = 3,96$$

- (a) *cluster* 2 ke *cluster* 1:

$$d(2,1) = \sqrt{\begin{aligned} &(0,48 - 0,17)^2 + (0,36 - 0,33)^2 + (0,71 - 0,62)^2 + \\ &(-1,92 - 0,41)^2 + (-1,77 - 0,28)^2 + \\ &(-1,4 - 0,38)^2 + (0,25 - 0,17)^2 + \\ &(0,37 - 0,33)^2 + (0,67 - 0,62)^2 + \\ &(0,17 - (-0,3))^2 + (0,15 - 0,39)^2 + \\ &(0,17 - 0,18)^2 + \dots (0,28 - 0,17)^2 + \\ &(0,40 - 0,33)^2 + (0,67 - 0,62)^2 + \\ &(0,98 - 0,41)^2 + (0 - 0,28)^2 + \\ &(1,34 - (-0,96))^2 \end{aligned}}$$

$$d(2,1) = \sqrt{\frac{728,13}{276}} = 2,64$$

Hasil perhitungan jarak rata-rata dari data i ke semua data pada *cluster* lainnya dapat dilihat pada Tabel 12

Tabel 12 Hasil Perhitungan $d(i,C)$

Data ke	$d(i,C)$	
	$d(i,1)$	$d(i,2)$
1.	-	3,96
2.	-	1,92
3.	-	2,26
4.	-	3,21
5.	-	2,95
6.	2,64	-
7.	-	3,52
8.	2,32	-
9.	-	4,53
10.	-	4,46

Data ke	d(i,C)	
	d(i,1)	d(i,2)
...
487.	-	4,18
488.	-	2,43
489.	3,41	-
490.	-	3,09
491.	-	2,84
492.	-	4,37
493.	-	3,81
494.	-	4,37
495.	-	4,09
496.	-	3,20

- (3) Menghitung nilai b(i) dengan mencari nilai terkecil dari semua d(i,C) pada Tabel 13.

Tabel 13 Hasil Perhitungan b(i)

Data ke	b(i)
1.	3,96
2.	1,92
3.	2,26
4.	3,21
5.	2,95
6.	2,64
7.	3,52
8.	2,32
9.	4,53
10.	4,46
...	...
487.	4,18
488.	2,43
489.	3,41
490.	3,09
491.	2,84
492.	4,37
493.	3,81
494.	4,37
495.	4,09
496.	3,20

- (4) Setelah mendapatkan nilai a(i) dan b(i), maka selanjutnya menghitung nilai *Silhouette Coefficient* pada setiap data di *cluster* nya.

Data ke-1

$$s(i) = \frac{(3,96-1,61)}{\max(1,61,3,96)} = 0,59$$

Data ke-2

$$s(i) = \frac{(1,92-1,67)}{\max(1,67,1,92)} = 0,13$$

Data ke-3

$$s(i) = \frac{(2,26-1,34)}{\max(1,34,2,26)} = 0,41$$

Data ke-4

$$s(i) = \frac{(3,21-1,42)}{\max(1,42,3,21)} = 0,56$$

Data ke-5

$$s(i) = \frac{(2,95-1,30)}{\max(1,30,2,95)} = 0,56$$

...

Data ke-496

$$s(i) = \frac{(3,20-1,25)}{\max(1,25,3,20)} = 0,61$$

Dari hasil perhitungan *silhouette* diatas, maka *Silhouette Coefficient* semua data berdasarkan urutan data pada Tabel 13 dapat dilihat pada Tabel 14.

Tabel 14 Hasil Perhitungan Silhouette Coefficient

Data ke	s(i)
1.	0,59
2.	0,13
3.	0,41
4.	0,56
5.	0,56
6.	0,37
7.	0,61
8.	0,27
9.	0,54
10.	0,51
...	...
487.	0,58
488.	0,25
489.	0,54
490.	0,60
491.	0,47
492.	0,55
493.	0,56
494.	0,55
495.	0,57
496.	0,61

Pada Tabel 14 menunjukkan hasil *Silhouette Coefficient* dari tiap data yang berjumlah . Berdasarkan nilai *Silhouette Coefficient* pada Tabel 3.15 dapat disimpulkan bahwa jika nilai *silhouette* semakin mendekati nilai 1, maka semakin baik pengelompokkan data ke dalam suatu *cluster*. Sebaliknya jika nilai *silhouette* semakin mendekati nilai 0, maka semakin kurang baik pengelompokkan data ke dalam suatu *cluster*. Hasil perhitungan uji hasil dengan menerapkan *Silhouette Coefficient* dengan dua *cluster* diperoleh nilai rata-rata 0,40 yang berarti masuk ke dalam kategori Struktur Lemah.

D. KESIMPULAN

Berdasarkan hasil penelitian, kesimpulan yang dapat diperoleh adalah bahwa proses pemetaan persebaran status gizi balita di Puskesmas Tanah Sareal Kota Bogor dengan pendekatan K-Means menggunakan sistem pendukung keputusan dengan pendekatan K-Means mendapatkan hasil lebih efektif dari yang sebelumnya. Penelitian ini menghasilkan dua cluster, 2 Cluster yaitu Cluster 1 sebanyak 276 dan Cluster 2 sebanyak 220 dengan kategori data pada Cluster 1 yaitu Gizi Kurang dan kategori data pada Cluster 2 artinya Gizi Lebih, yang keduanya menjadi acuan untuk dilakukan penyuluhan. Pengembangan prototype sistem dikembangkan dengan bahasa pemrograman python dan data CSV terdapat 5 menu yaitu menu beranda, menu unggah file, menu halaman Perhitungan Clustering yang menampilkan hasil cluster, menu Diagram Hasil Clustering menampilkan diagram/grafik hasil cluster, persentase cluster, dan menu Evaluasi Cluster hasil menampilkan Silhouette Coefficient. Hasil pengujian sistem memperoleh tingkat keakuratan presentase kelayakan oleh ahli sistem sebesar 100% "Sangat Layak" dan oleh pengguna memperoleh hasil sebesar 90,4% "Sangat Layak", serta tingkat keefektifitasan dengan metode uji Silhouette Coefficient memperoleh hasil sebesar 0,40.

E. DAFTAR PUSTAKA

- [1] Arhami, M., & Nasir, M. (2020). *Data Mining Algoritma dan Implementasi* (R. I. Utami (ed.)). Penerbit Andi;Politeknik Negeri Lhokseumawe. <https://books.google.co.id/books?id=AtcCEAAQBAJ&printsec=frontcover&hl=id#v=onepage&q&f=false>
- [2] Fitri, N., Astuti, W., & Huriyati, E. (2022). *Usia ibu dan jumlah anak berhubungan dengan kejadian beban gizi ganda pada tingkat rumah tangga di desa dan kota di Indonesia*. 18(3), 104–114. <https://doi.org/10.22146/ijcn.69959>
- [3] Habibi, R., & Suryansah, A. (2020). *Aplikasi prediksi jumlah kebutuhan perusahaan*. Kreatif. <https://books.google.co.id/books?id=sjDwDwAAQBAJ>
- [4] Istiqomah, A., S, K. M., Amali, R. A., & Tiawati, S. (2024). *Peran Gizi Terhadap Pertumbuhan Dan Perkembangan Balita*. 2(2).
- [5] Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining dan Penerapan*. https://www.google.co.id/books/edition/Konsep_Data_Mining_Dan_Penerapan/vjhWEQAAQBAJ?hl=ban&gbpv=0
- [6] Ma'ady, M. N. P., Lidiawaty, B. R., Dzulkarnain, A., & Ramadan, A. (2024). *Data Mining Algoritma dan Contoh Perhitungan Matematis*. DEEPUBLISH. https://www.google.co.id/books/edition/Data_Mining_Algoritma_Dan_Contoh_Perhitu/xsQ0EQAAQBAJ?hl=ban&gbpv=1&dq=Data+Mining+Algoritma+dan+Contoh+Perhitungan+Matematis&printsec=frontcover
- [7] Mulaab. (2021). *Data Mining: Konsep dan Aplikasi*. Media Nusa Creative (MNC Publishing). <https://books.google.co.id/books?id=X1FKEAAQBAJ>
- [8] Muttaqin, Widiyanto, W. W., Munsarif, M., Mandias, G. F., Pungus, S. R., Widarman, A., Hapsari, W. K., Hardiyanti, S. A., Fatkhudin, A., Pasnur, Bisono, E. F., Anshori, M., & Suryani, N. S. (2023). *Pengenalan Data Mining* (Issue July).
- [9] Prihandoko, P., Alam, R. G. G., Gunawan, G., Abdullah, D., Efitra, E., & Uzma, I. (2024). *Memahami Konsep dan Implementasi Machine Learning*. PT. Sonpedia Publishing Indonesia. https://books.google.co.id/books?id=_831EAAQBAJ
- [10] Rahayu, P. W., Sudipa, I. G. I., Suryani, S., Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, M. N., Slamet, I., Harlina, S., & Maysanjaya, I. M. D. (2024). *Buku Ajar Data Mining*. PT. Sonpedia Publishing Indonesia. <https://books.google.co.id/books?id=vCruEAAQBAJ>
- [11] Sinurat, Y. F. (2024). *Data Mining Pengelompokan Siswa Berprestasi Menggunakan Metode Clustering*. Penerbit NEM. <https://books.google.co.id/books?id=0TMZEQAAQBAJ>
- [12] Wijaksana, V. P., Hairah, U., Wati, M., & Puspitasari, N. (2025). *Penerapan Metode K-Means Clustering Status Gizi Balita Di UPT Puskesmas Barong Tongkok*. 5(1), 156–165.