

PERBANDINGAN METODE K-MEANS DAN K-MEDOIDS UNTUK KLASIFIKASI STATUS GIZI ANAK

Maulana Muhammad, Ali Mahmudi, Karina Auliasari
Teknik Informatika, Institut Teknologi Nasional Malang
Jalan Raya Karanglo km 2 Malang, Indonesia
1918008@scholar.itn.ac.id

ABSTRAK

Status gizi anak, yang diukur melalui berat dan tinggi badan, mencerminkan keberhasilan pemenuhan nutrisi. Gizi buruk atau kurang, terutama di kalangan balita, berdampak serius pada kesehatan, pertumbuhan, perkembangan, dan produktivitas masa dewasa. Di Kecamatan Pringgarata, Puskesmas Bagu, lebih dari 4600 balita mengalami gizi kurang. Masalah ini diakibatkan oleh konsumsi pangan yang tidak memadai dan faktor ekologis seperti kemiskinan serta sanitasi yang buruk. Penelitian ini bertujuan untuk memantau status gizi anak melalui sistem klasifikasi otomatis. Melalui pendekatan *k-means* dan *k-medoids*, sistem ini membantu mengatasi permasalahan gizi dengan memantau balita menggunakan indikator pendukung. Harapannya, metode ini dapat membantu petugas puskesmas dalam pemantauan gizi balita secara berkala.

Kata kunci : *K-means, k-medoids, klasifikasi, perbandingan, status gizi*

1. PENDAHULUAN

Status gizi adalah ukuran keberhasilan dalam pemenuhan nutrisi untuk anak yang diindikasikan oleh berat badan dan tinggi badan anak, hal ini dijadikan sebagai acuan untuk menentukan status Kesehatan gizi. Gizi buruk atau gizi kurang merupakan fenomena yang lekat pada permasalahan gizi yang dialami oleh kelompok balita yang berdampak pada masalah kesehatan, pertumbuhan hingga perkembangan balita, serta produktivitas di masa dewasa. Dapat disimpulkan bahwa stunting merupakan pertumbuhan yang melambat pada balita dan gizi kurang merupakan suatu keadaan dimana balita mudah mengalami infeksi karena kekebalan tubuhnya yang rendah.

Berdasarkan, data Pemerintah Kabupaten Lombok Tengah, dikutip pada radarmandalika.id menyatakan dari hasil pemantauan status gizi Kabupaten Lombok Tengah pada tahun 2022, kasus gizi kurang tembus diangka 20.509 yang menimpa balita. Sedangkan di Kecamatan Pringgarata, Puskesmas Bagu berada, angka gizi kurang mencapai angka 3588 balita. Masalah gizi kurang ini diantara lain disebabkan karena konsumsi yang tidak adekuat dipandang sebagai suatu permasalahan ekologis yang tidak tidak saja disebabkan oleh ketidakcukupan ketersediaan pangan dan zat-zat gizi tertentu juga dipengaruhi oleh kemiskinan, sanitasi lingkungan yang kurang baik dan ketidaktahuan tentang gizi.

Meskipun perkembangan teknologi sangat pesat, Puskesmas Desa Bagu melakukan pengambilan dan penginputan data secara manual menggunakan kertas dan menggunakan Microsoft excel. Dimana ukuran yang dihasilkan menjadi besar dan proses komputasi menjadi lebih berat dan masyarakat hanya bisa melakukan pengecekan dalam kurun waktu sebulan sekali pada kegiatan posyandu yang dilakukan.

Oleh karena itu perlu ada perancangan sistem klasifikasi untuk status gizi untuk memantau keadaan balita dengan beberapa indikator pendukung guna

mendukung inovasi dan pemahaman masyarakat akan pentingnya gizi terhadap balita. Dengan dibuatnya sistem ini diharapkan petugas puskesmas dapat melakukan pengukuran indikator gizi pada balita khususnya yang berkaitan dengan kondisi gizi kurang dan stunting secara berkala.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Menurut Nurul Rizki Octaviyani dan rekan-rekan dalam penelitiannya yang berjudul Implementasi Algoritma *K-means* Clustering Status Gizi Balita yang bertujuan melakukan pengelompokan status gizi balita dengan algoritma *k-means* memiliki hasil akhir yang menunjukkan *k-means* menggambarkan status gizi balita dinilai cukup baik karena DBI (Index Davies Bouldin) dinyatakan optimal secara esensial jika nilai yang didapatkan serendah mungkin (non-negatif ≥ 0) [1].

Menurut Eni Irfani dan Siti Sulistia Rani dalam penelitiannya yang berjudul Algoritma *K-means* Clustering untuk Menentukan Nilai Gizi Balita yang bertujuan untuk melakukan klasterisasi nilai gizi balita di Desa Sukamantri Bogor menggunakan metode *k-means* memiliki hasil bahwa nilai gizi balita Desa Sukamantri Bogor dapat diklasterisasikan dengan menggunakan algoritma *k-means* melalui dua parameter yaitu berat badan dan tinggi badan yang dibagi menjadi 5 cluster yaitu obesitas, gizi lebih, gizi baik, gizi kurang, dan gizi buruk, dari hasil tersebut diketahui masih terdapat 30% balita obesitas serta 11% balita kekurangan gizi [2].

Menurut Windha Mega Pradnya Dhuhiita dalam penelitiannya yang berjudul Clustering Menggunakan Metode *K-means* untuk Menentukan Status Gizi Balita yang bertujuan untuk melakukan pengelompokan nilai gizi balita di Desa Karang Songo, Jetis, Bantul menggunakan metode *k-means*, menghasilkan perbandingan menggunakan table growth chart dan

algoritma *k-means* didapat 17 data yang memiliki kelompok yang sama. Dapat disimpulkan bahwa algoritma *k-means* hanya memiliki *accuracy* 34% benar dan *accuracy*nya bisa berubah dengan seiring penambahan data latih[3].

2.2. Data Mining

Data mining merupakan gabungan sejumlah disiplin ilmu komputer yang didefinisikan sebagai proses penemuan pola-pola baru dari kumpulan-kumpulan data sangat besar, meliputi metode-metode yang beririsan dari kecerdasan buatan, mesin pembelajaran, statistic dan sistem basis data[4].

Data mining yaitu suatu proses penggalian data data dari sebuah informasi yang sangat penting. Data Mining juga merupakan suatu proses untuk menggali pola-pola dari data. Pola-pola itu didapatkan dari berbagai jenis basis data seperti basis data relasional, data warehouse, data transaksi, dan data berorientasi objek. Penggunaan data mining dapat membantu para pebisnis dalam pengambilan keputusan secara cepat dan tepat[5].

2.3. Status Gizi

Status gizi adalah keadaan tubuh sebagai akibat dari pemakaian, penyerapan, dan penggunaan makanan. Pengertian lain menyebutkan bahwa status gizi merupakan ekspresi dari keadaan keseimbangan dalam bentuk variabel tertentu, atau perwujudan dari status tubuh yang berhubungan dengan gizi dalam bentuk variabel tertentu[6].

Status gizi balita dipengaruhi oleh beberapa faktor diantaranya adalah faktor langsung dan faktor tidak langsung. Faktor langsung yang dapat mempengaruhi status gizi balita adalah adanya penyakit infeksi dan asupan makanan. Sedangkan faktor tidak langsung yang mempengaruhi status gizi balita adalah ketahanan pangan di dalam keluarga, pola asuh, sanitasi lingkungan, akses terhadap pelayanan kesehatan, umur anak, jenis kelamin anak, tempat tinggal, pendidikan, dan pekerjaan orang tua[7].

Berdasarkan hasil Riset Kesehatan Dasar (Kemenkes, 2018) didapatkan status gizi anak 5- 12 tahun menurut indeks massa tubuh/umur di Indonesia, yaitu prevalensi kurus adalah 9,3, terdiri dari 2,5% sangat kurus dan 6,8% kurus. Masalah gemuk pada anak di Indonesia juga masih tinggi dengan prevalensi 20,6% terdiri dari gemuk 11,1% dan sangat gemuk (obesitas) 9,5%. Sedangkan prevalensi pendek yaitu 23,6% terdiri dari 6,7 sangat pendek dan 16,9% pendek[8].

2.4. K-means

K-Means merupakan algoritma clustering yang berulang ulang dimana dimulai dengan penentuan jumlah *cluster* atau kelompok, kemudian menentukan nilai *centroid* secara random. Setelah mendapatkan nilai *centroid* dilanjutkan dengan menghitung jarak data terhadap setiap *centroid* hingga ditemukan

jarak terdekat data dengan *centroid*, kemudian mengelompokkan data berdasarkan kedekatannya dengan *centroid*. Setelah mengelompokkan data menghitung kembali nilai *centroid* baru. Kemudian langkahnya berulang sampai data di setiap kelompok tidak berubah posisinya[9].

Menurut Eko Prasetyo mengatakan bahwa metode *K-means* ini mempartisi data kedalam kelompok sehingga data berkarakteristik sama dimasukan kedalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain[10].

Berikut merupakan ilangkah-langkah metode *K-means*:

1. Inisialisasi: tentukan nilai k sebagai jumlah klaster yang diinginkan dan matriks jarak yang diinginkan.
2. Pilih k data dari set data X sebagai *centroid*.
3. Alokasikan semua data ke *centroid* terdekat dengan matriks jarak yang sudah ditetapkan (memperbarui klaster ID pada setiap data). Pada penelitian ini menggunakan jarak *euclidean* seperti berikut :

$$D(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^2} \quad (1)$$

x = *centroid*

y = data

4. Hitung kembali *centroid* berdasarkan data yang mengikuti klaster masing-masing.

$$C_1 = \frac{1}{M} \sum_{j=1}^M x_j \quad (2)$$

M = jumlah data suatu cluster

x = data dalam suatu cluster

5. Ulangi langkah 3 dan 4 hingga kondisi konvergen tercapai, yaitu tidak ada data yang berpindah kluster.

2.5. K-medoids

K-Medoids atau *Partitioning Around Medoid* (PAM) ditemukan oleh Kaufman dan Rousseeuw pada tahun 1990. *Clustering* melibatkan pengelompokkan, pemisahan, atau pembagian objek data menjadi k cluster, dimana jumlah k lebih kecil atau sama dengan data cluster. Medoid adalah objek yang diasumsikan mewakili cluster dan pusat cluster[11].

K-medoids Clustering, juga dikenal sebagai *Partitioning Around Medoids* (PAM), adalah varian dari metode *K-means*. Hal ini didasarkan pada penggunaan *medoids* bukan dari pengamatan mean yang dimiliki oleh setiap klaster, dengan tujuan mengurangi sensitivitas dari partisi yang dihasilkan sehubungan dengan nilai-nilai ekstrim yang ada dalam dataset[12].

Adapun tahapan *K-medoids Clustering* adalah :

1. Secara acak pilih k objek pada sekumpulan n objek sebagai *medoid*.
2. Ulangi.
3. Tempatkan objek non-*medoid* dengan menghitung jaraknya dengan *medoid* ke dalam klaster yang paling dekat dengan *medoid*. Pada penelitian ini menggunakan jarak *euclidean* seperti berikut :

$$D(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^2} \quad (3)$$

x = *centroid*
y = data

4. Secara acak pilih Orandom (sebuah objek non *medoid*).
5. Hitung total cost, S, dari pertukaran *medoid* Oj dengan Orandom.
6. Jika $S < 0$ maka tukar Oj dengan Orandom untuk membentuk sekumpulan k objek baru sebagai *medoid*.
7. Jika $S > 0$ maka kondisi konvergen tercapai dan perhitungan dihentikan.

3. METODE PENELITIAN

3.1. Metode Pengumpulan Data

Penelitian ini menggunakan data sekunder yang merupakan data tidak langsung didapatkan dari obyek penelitian, melainkan data yang telah dikumpulkan oleh perangkat desa melalui dokumen hasil survey dan data pada website pemerintah yang diberikan kepada perangkat desa.

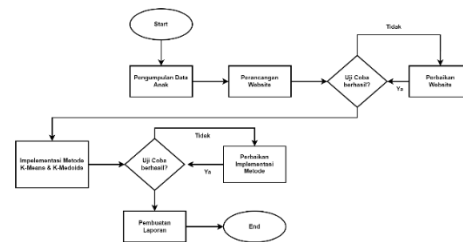
Berikut merupakan metode pengumpulan data yang dilakukan pada penelitian ini:

1. Wawancara
Peneliti melakukan wawancara dengan Kepala Desa dan admin yang terkait untuk mendapatkan informasi beserta data di tempat objek penelitian sesuai dengan kebutuhan.
2. Studi Literatur
Peneliti mengumpulkan berbagai informasi dan mempelajari semua materi serta sumber-sumber data yang berhubungan dengan penelitian yang dilakukan.

3.2. Analisa dan Perancangan

Analisis yang di lakukan guna mengetahui kebutuhan yang di harapkan sehingga dapat dilakukan perancangan website yang tepat, Maka di buat sistem ini agar meringankan proses pendataan dan meminimalisir adanya kecurangan dalam pemrosesan data yang sebenarnya.

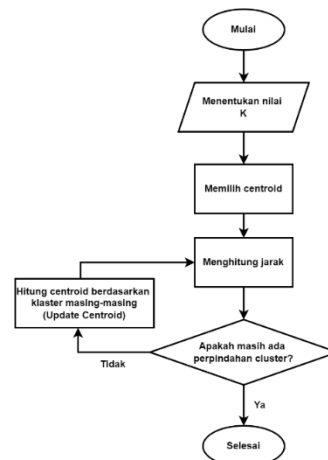
3.3. Diagram alir sistem



Gambar 1. Diagram alir sistem

Pada Gambar 1. merupakan proses dari pengerjaan website. Pertama akan dilakukan pengumpulan data anak yang diperoleh dari Puskesmas Bagu, kemudian akan dilakukan perancangan website, disini akan menghasilkan proses create, read, update, delete pada data yang diperlukan kemudian diuji coba, apabila tidak berhasil akan dilakukan perbaikan, apabila berhasil akan dilanjutkan pada proses selanjutnya yaitu implementasi metode *k-means* dan *k-medoids*, apabila implementasi belum berhasil akan dilakukan perbaikan, apabila berhasil akan dilanjutkan keproses pembuatan laporan.

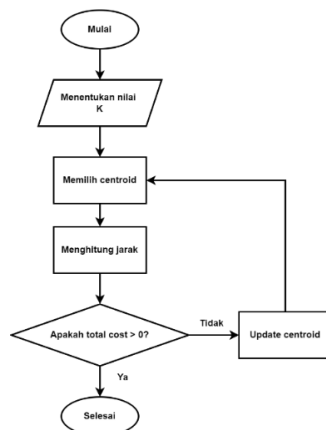
3.4. Flowchart Metode K-means



Gambar 2. Flowchart metode K-means

Pada Gambar 2, proses pertama adalah menentukan nilai K atau cluster kemudian memilih titik data yang akan dijadiakann *centroid*, setelah itu menghitung jarak, apabila masih ada perpindahan cluster maka akan dilakukan perhitungan untuk memperbarui titik *centroid*, apabila tidak maka nilai bisa disebut konvergen dan proses selesai.

3.5. Flowchart Metode K-medoids



Gambar 3. Flowchart metode K-medoids

Pada Gambar 3, proses pertama adalah menentukan nilai K atau cluster kemudian memilih titik data yang akan dijadikan *centroid*, setelah itu menghitung jarak, kemudian akan dihitung total cost dari jarak, apabila > 0 maka akan melakukan update *centroid* dengan memilih *centroid* baru, apabila tidak maka nilai bisa disebut konvergen dan proses selesai.

3.6. Perancangan Metode K-means

Dalam penelitian ini terdapat beberapa nilai k adalah 3 yang terdiri dari C1 (Gizi Lebih), C2 (Gizi Baik), C3 (Gizi Kurang).

1. Sample data

Tabel 1. Sampel Data

No	BB Lahir	TB Lahir	Berat	Tinggi
1	3.1	50	50	9
2	3.2	48	48	5.7
237	2.9	48	48	10.7
238	3.1	48	48	10
239	2.6	49	49	9.5
240	3.2	49	49	9.9
241	2.7	48	48	7.8
245	3	48	48	12.4
246	2.9	48	48	10.7

Pada Tabel 1. merupakan sampel data yang akan digunakan untuk proses *k-means* dimana data terdiri dari 246 sample data dengan fitur NIK, Nama, BB Lahir, TB Lahir, Berat dan Tinggi.

2. Penentuan pusat awal cluster

Tabel 2. Tabel *centroid* awal

	BB Lahir	TB Lahir	Berat	Tinggi
Centroid-1	3.1	50	9	66
Centroid-2	2.8	49	12.6	92.6
Centroid-3	2.9	48	12.9	100.1

Pada Tabel 2 pusat awal cluster atau *centroid* dilakukan secara acak,

3. Penentuan jarak pusat cluster

Untuk mengukur jarak antar data dengan *centroid* digunakan *euclidian distance*. Jarak antar data ke *centroid* dapat dihitung menggunakan persamaan berikut :

$$D = \sqrt{(3.1 - 3.1)^2 + (50 - 50)^2 + (9 - 9)^2 + (66 - 66)^2} = 0$$

Hasil Data No.1

$$D = \sqrt{(3.2 - 3.1)^2 + (48 - 50)^2 + (5.7 - 9)^2 + (55 - 66)^2} = 11.65761554$$

Hasil Data No.2

$$D = \sqrt{(3.3 - 3.1)^2 + (50 - 50)^2 + (12.7 - 9)^2 + (100.2 - 66)^2} = 34.40014535$$

Hasil Data No.237

$$D = \sqrt{(2.9 - 3.1)^2 + (48 - 50)^2 + (10.7 - 9)^2 + (91.7 - 66)^2} = 25.83447309$$

Hasil Data No.238

$$d = \sqrt{(3.1 - 3.1)^2 + (48 - 50)^2 + (10.7 - 9)^2 + (88.5 - 66)^2} = 22.61083811$$

Hasil Data No.239

$$D = \sqrt{(2.6 - 3.1)^2 + (49 - 50)^2 + (9.5 - 9)^2 + (85.3 - 66)^2} = 19.33882106$$

Hasil Data No.240

$$D = \sqrt{(3.2 - 3.1)^2 + (49 - 50)^2 + (9.9 - 9)^2 + (85.5 - 66)^2} = 19.54661096$$

Hasil Data No.241

$$D = \sqrt{(3 - 3.1)^2 + (49 - 50)^2 + (11.7 - 9)^2 + (94 - 66)^2} = 28.14782407$$

Hasil Data No.245

$$D = \sqrt{(3.2 - 3.1)^2 + (48 - 50)^2 + (8.6 - 9)^2 + (65 - 66)^2} = 2.2737634$$

Hasil Data No.10

$$D = \sqrt{(2.6 - 3.1)^2 + (49 - 50)^2 + (9.5 - 9)^2 + (85.3 - 66)^2} = 33.83090303$$

Hasil Data No.246

Perhitungan tersebut merupakan persamaan matematika untuk C1 (Gizi Lebih), untuk menentukan jarak C2 (Gizi Lebih) dan C3 (Gizi Baik) proses persamaan matematika tersebut akan diulangi tetapi nilai *centroid* disesuaikan dengan nilai *centroid* yang akan dihitung.

Setelah menghitung semua jarak, berikut merupakan langkah perkemudian didapatkan matrik jarak seperti Tabel 3 berikut :

Tabel 3. Jarak matrik

No	C1 (Gizi Lebih)	C2 (Gizi Baik)	C3 (Gizi Kurang)	Hasil
1	0	26.8628	34.3811	Gizi Lebih
2	11.65762	38.24304	45.67209	Gizi Lebih
239	22.61084	4.965884	11.95868	Gizi Baik
240	19.33882	7.933473	15.22137	Gizi Baik
245	28.14782	1.676305	6.297619	Gizi Baik
246	33.8309	7.076722	0.714143	Gizi Kurang

Pada Tabel 3 diperoleh hasil dari perhitungan jarak matriks yaitu dua Gizi Lebih, dua Gizi Baik dan dua Gizi Kurang.

4. Hitung *centroid* baru

Tabel 4. Centroid baru

	BB Lahir	TB Lahir	Berat	Tinggi
Centro id-1	3.005434 783	49.14130 435	8.630434 783	67.56630 435
Centro id-2	2.9784	49.36	10.8512	87.3856
Centro id-3	2.996551 724	49.10344 828	12.57931 034	98.99655 172

Pada Tabel 4 hasil dari pusat *centroid* baru yang ditentukan dengan menghitung nilai rata-rata dari data yang ada pada *centroid* sama.

5. Hasil akhir *k-means*.

Tabel 5. Hasil akhir *k-means* iterasi 8

No	C1 (Gizi Lebih)	C2 (Gizi Baik)	C3 (Gizi Kurang)	Hasil
1	2.109673379	16.29980707	29.02935814	Gizi Lebih
2	9.664026507	27.6391018	40.37753937	Gizi Lebih
239	24.34464861	6.410082368	6.785054713	Gizi Baik
240	21.10045213	3.146119333	9.904934966	Gizi Baik
245	29.95409382	11.86820144	0.947354356	Gizi Kurang
246	35.61305906	17.56252332	4.895895012	Gizi Kurang

Tabel 5 merupakan hasil akhir dari *k-means* setelah dilakukan melakukan perhitungan jarak dengan *euclidean distance*, titik konvergen atau nilai cluster tidak berubah setelah iterasi 8.

3.7. Perancangan Metode K-medoids

Untuk data sampel pada metode *k-medoids* sama dengan *k-means*, dengan nilai *k* yang juga sama terdiri dari terdiri dari C1 (Gizi Lebih), C2 (Gizi Baik), C3 (Gizi Kurang).

1. Penentuan pusat awal cluster

Tabel 6. Tabel *centroid* awal

	BB Lahir	TB Lahir	Berat	Tinggi
Medoids-1	3.1	51	11	73.3
Medoids-2	2.9	50	11.6	88
Medoids-3	2.7	51	10.5	87.9

Tabel 6 menentukan pusat awal cluster atau *medoid* dilakukan secara acak.

2. Penentuan jarak pusat cluster dan total cost

Untuk mengukur jarak antar data dengan *medoid* digunakan *euclidian distance*. Jarak antar data ke *centroid* dapat dihitung menggunakan persamaan berikut :

$$D = \sqrt{(3.1 - 3.1)^2 + (50 - 51)^2 + (9 - 11)^2 + (66 - 73.3)^2} = 7.634788799$$

Hasil Data No.1

$$D = \sqrt{(3.2 - 3.1)^2 + (48 - 51)^2 + (5.7 - 11)^2 + (55 - 73.3)^2} = 19.28704228$$

Hasil Data No.2

$$D = \sqrt{(3.3 - 3.1)^2 + (50 - 51)^2 + (12.7 - 11)^2 + (100.2 - 73.3)^2} = 26.97294941$$

Hasil Data No.237

$$D = \sqrt{(2.9 - 3.1)^2 + (48 - 51)^2 + (10.7 - 11)^2 + (91.7 - 73.3)^2} = 18.64644738$$

Hasil Data No.238

$$D = \sqrt{(3.1 - 3.1)^2 + (48 - 51)^2 + (10.7 - 11)^2 + (88.5 - 73.3)^2} = 15.52546296$$

Hasil Data No.239

$$D = \sqrt{(2.6 - 3.1)^2 + (49 - 51)^2 + (9.5 - 11)^2 + (85.3 - 73.3)^2} = 12.26784415$$

Hasil Data No.240

$$D = \sqrt{(3.2 - 3.1)^2 + (49 - 51)^2 + (9.9 - 11)^2 + (85.5 - 73.3)^2} = 12.41209088$$

Hasil Data No.241

$$D = \sqrt{(3 - 3.1)^2 + (49 - 51)^2 + (11.7 - 11)^2 + (94 - 73.3)^2} = 20.80841176$$

Hasil Data No.245

$$D = \sqrt{(2.6 - 3.1)^2 + (49 - 51)^2 + (9.5 - 11)^2 + (85.3 - 73.3)^2} = 26.50773472$$

Hasil Data No.246

Perhitungan tersebut merupakan persamaan matematika untuk M1 (Gizi Lebih), untuk menentukan jarak M2 (Gizi Lebih) dan M3 (Gizi Baik) proses persamaan matematika tersebut akan diulangi tetapi nilai *medoid* disesuaikan dengan nilai *medoid* yang akan dihitung.

Setelah menghitung semua jarak, kemudian didapatkan matrik jarak dan total cost seperti Tabel 7 berikut :

Tabel 7. Jarak matrik dan total cost

No	M1 (Gizi Lebih)	M2 (Gizi Baik)	M3 (Gizi Kurang)	Hasil
1	7.634788799	22.15400641	21.7784297	Gizi Lebih
2	19.28704228	33.58422249	33.19005875	Gizi Lebih
237	26.97294941	12.25601893	12.74558747	Gizi Baik
238	18.64644738	4.301162634	5.00799361	Gizi Baik
240	12.26784415	3.576310948	3.281767816	Gizi Kurang
241	12.41209088	3.198437118	3.07408523	Gizi Kurang
Total Cost			1647.664	

Pada Tabel 7 diperoleh hasil yaitu dua Gizi Lebih, dua Gizi Baik dan dua Gizi Kurang dengan total cost 1647.664.

3. Menentukan *medoid* baru

Tabel 8. *Medoid* baru

	BB Lahir	TB Lahir	Berat	Tinggi
Medoids-1	3.1	49	5.9	56
Medoids-2	3	50	11.7	92.5
Medoids-3	2.8	50	9.5	88.4

Tabel 8 merupakan tabel *medoid* baru untuk perhitungan jarak pada iterasi 2 yang ditentukan secara random.

4. Penentuan jarak pusat cluster dan total cost iterasi ke-2

Untuk mengukur jarak antar data dengan *medoid* digunakan *euclidian distance*, kemudian didapatkan matrik jarak dan total cost seperti Tabel 9 berikut :

Tabel 9. Jarak matrik dan total cost iterasi-2

No	M1 (Gizi Lebih)	M2 (Gizi Baik)	M3 (Gizi Kurang)	Hasil
1	10.51712889	26.63737975	22.407588	Gizi Lebih
2	1.431782106	38.03011964	33.67729205	Gizi Lebih
237	44.73164428	7.770456872	12.23642105	Gizi Baik
238	36.03567677	2.376972865	4.042276586	Gizi Baik
240	29.52456604	7.605261337	3.263433774	Gizi Kurang
241	29.77011925	7.299315036	3.119294792	Gizi Kurang
Total Cost			1562.04394	

Pada Tabel 9 diperoleh hasil yaitu dua Gizi Lebih, dua Gizi Baik dan dua Gizi Kurang dengan total cost 1562.04394.

5. Menghitung selisih cost

Karena sudah terdapat iterasi-1 dan iterasi-2 maka selanjutnya adalah mencari selisih cost, berikut merupakan hasil selisih cost antar iterasi-1 dan iterasi-2:

$$\text{Selisih cost} = 1562.044 - 1647.664$$

$$\text{Selisih cost} = -85.62033838$$

Syarat iterasi berhenti apabila nilai selisih cost bernilai lebih dari 0, karena selisih cost pada perhitungan cost iterasi-2 dikurangi iterasi-1 adalah -85.62033838, maka iterasi akan dilanjutkan.

6. Hasil akhir *k-medoids*

Tabel 10. Hasil akhir *k-medoids*

No	M1 (Gizi Lebih)	M2 (Gizi Baik)	M3 (Gizi Kurang)	Hasil
1	3.229551052	28.83522152	19.03943276	Gizi Lebih
2	8.497646733	40.15084059	30.28200786	Gizi Lebih
237	37.58377309	6.044832504	15.525785	Gizi Baik
238	28.98637611	3.573513677	6.854195795	Gizi Baik
240	22.46953493	9.87724658	0.7	Gizi Kurang

No	M1 (Gizi Lebih)	M2 (Gizi Baik)	M3 (Gizi Kurang)	Hasil
241	22.69449272	9.539392014	0.538516481	Gizi Kurang
Total Cost			1562.04394	

Pada Tabel 10 perhitungan *k-medoids*, nilai sudah dikatakan bisa konvergen pada iterasi ke-3 karena pada iterasi ke-4 nilai total cost bernilai > 0.

4. HASIL DAN PEMBAHASAN

4.1. Implementasi Software

Hasil website saat ini adalah pembuatan halaman Train, dan Perbandingan.

1. Tampilan Halaman Train

Gambar 6. Tampilan Train

Pada Gambar 6 Melihat tentang jumlah cluster, total iterasi serta menampilkan tabel yang berisi nama anak, berat lahir, tinggi lahir, berat, tinggi dan hasil label dari *k-means* dan *k-medoids*.

2. Tampilan Halaman Perbandingan

Gambar 7. Tampilan Perbandingan

Pada Gambar 7. Halaman ini menampilkan hasil perbandingan *k-means* dan *k-medoids* yang tabelnya terdiri dari nama anak, *k-means* label, *k-medoids* label dan label aktual.

4.2. Pengujian Black Box

Pengujian *black box* dilakukan untuk menguji fitur-fitur yang ada pada website Hasil dari pengujian bisa dilihat pada Tabel 11.

Tabel 11. Pengujian *Black Box*

No	Pengujian	Harapan Hasil Uji	Hasil Uji
1	Halaman login	Username, password sesuai dan berhasil	Berhasil
2	Halaman Dashboard	Menampilkan grafik dari data anak, data orangtua, data gizi	Berhasil

No	Pengujian	Harapan Hasil Uji	Hasil Uji
3	Halaman Data Orangtua	Menampilkan, menambah, mengedit, dan menghapus data orangtua dan disimpan dalam database	Berhasil
4	Halaman Data Anak	Menampilkan, menambah, mengedit, dan menghapus data anak dan disimpan dalam database	Berhasil
5	Halaman Data Gizi	Menampilkan, menambah, mengedit, dan menghapus data gizi dan disimpan dalam database	Berhasil
6	Halaman Train	Melakukan proses klasifikasi dengan algoritma clustering <i>k-means</i> dan <i>k-medoids</i> kemudian label hasil ditampilkan bersamaan dengan nama anak, berat lahir, tinggi lahir, berat, dan tinggi, serta menampilkan jumlah cluster, total iterasi <i>k-means</i> dan <i>k-medoids</i> , dan waktu training <i>k-means</i> dan <i>k-medoids</i>	Berhasil
7	Halaman Perbandingan	Menampilkan <i>k-means</i> label, <i>k-medoids</i> label dan label aktual, serta sistem melakukan perhitungan total true positive, true negative, false positive, dan false negative serta menampilkan hasil dari <i>confusion matrix accuracy</i> , <i>sensitivity</i> dan <i>specificity</i> .	Berhasil

Tabel 11 pengujian black box, dapat disimpulkan semua halaman pada website berhasil dan berjalan baik.

Pengujian user dilakukan bertujuan untuk mengetahui apakah sistem sudah berjalan dengan baik atau belum. Hasil dari pengujian bisa dilihat pada Tabel 12.

4.3. Pengujian User

Tabel 12. Pengujian User

No	Pengujian	Harapan Hasil Uji Hasil Uji		
		Sangat Setuju (SS)	Setuju (S)	Kurang Setuju (KS)
1	Tampilan User interface dari website klasifikasi status gizi anak dengan <i>k-means</i> dan <i>k-medoids</i> terlihat Menarik	6 (28.6%)	15 (71.4%)	0 (0%)
2	Fitur yang ada pada aplikasi website klasifikasi status gizi anak dengan <i>k-means</i> dan <i>k-medoids</i> berfungsi dengan baik.	7 (33.3%)	14 (66.7%)	0 (0%)
3	Website klasifikasi status gizi anak dengan <i>k-means</i> dan <i>k-medoids</i> mudah digunakan	8 (38.1%)	13 (61.9%)	0 (0%)
4	Informasi yang ada pada website klasifikasi status gizi anak dengan <i>k-means</i> dan <i>k-medoids</i> mudah di mengerti	8 (42.9%)	12 (57.1%)	0 (0%)
5	Website klasifikasi status gizi anak dengan <i>k-means</i> dan <i>k-medoids</i> dapat membantu dalam mengklasifikasikan status gizi anak	11 (52.4%)	10 (47.6%)	0 (0%)
Rata-rata jawaban		40 (39%)	64 (61%)	0 (0%)

Berdasarkan pada Tabel 12 dapat dilihat jawaban sangat setuju mendapatkan persentase 39%, jawaban setuju mendapatkan persentase 61%, dan jawaban kurang setuju 0% sehingga dapat dikatakan website sudah berjalan dengan baik.

4.4. Hasil perbandingan *K-means* dengan *K-medoids*

Untuk hasil perbandingan, *k-means* dengan *k-medoids* menghasilkan dua perbandingan dari segi *confusion matrix* dan implementasi metode pada sistem seperti tabel 13:

Tabel 13. Perbandingan *confusion matrix*

	<i>Accuracy</i>	<i>Specificity</i>	<i>Sensitivity</i>
Selisih <i>K-Means</i>	0%	0%	0%
Selisih <i>K-Medoids</i>	-0.82%	-0.66%	-1.27%
Rata-rata <i>K-Means</i>	77.8%	82.88%	67.70%
Rata-rata <i>K-Medoids</i>	68.97%	76.16%	55.56%

Pada Tabel 13 terlihat apabila selisih untuk *k-means* lebih besar yaitu 0% pada *accuracy*, *specificity*, dan *sensitivity*, sedangkan selisih *k-medoids* yaitu -0.82% pada *accuracy*, -0.66 pada *specificity* dan -1.27% pada *sensitivity*. apabila Rata-rata untuk *k-means* lebih besar yaitu 77.8% pada *accuracy*, 82.88% pada *specificity* dan 67.70% pada *sensitivity*, sedangkan rata-rata pada *k-medoids* lebih kecil yaitu 68.97% pada *accuracy*, 76.16% pada *specificity* dan 55.56% pada *sensitivity*.

Tabel 14. Perbandingan sistem

	<i>K-Means</i>	<i>K-Medoids</i>
Jumlah Looping (Iterasi)	8	3
Waktu Train Metode	0.182 detik	0.406 detik
Baris Implementasi Metode	62 baris	65 baris

Pada Tabel 14 dapat disimpulkan bahwa dari segi implementasi metode pada sistem pada saat training dilakukan, *k-means* lebih baik dari waktu train metode dan baris implementasi metode dibandingkan *k-medoids*.

5. KESIMPULAN DAN SARAN

Berdasarkan pengujian black box, pengujian fungsional yang telah dilakukan mendapatkan hasil semua halaman pada website klasifikasi status gizi anak dengan metode *k-means* dan *k-medoids* berhasil dan berjalan baik. Hasil pengujian user, perhitungan persentase yang telah dilakukan dapat disimpulkan jawaban sangat setuju mendapatkan persentase 39%, jawaban setuju mendapatkan persentase 61%, dan jawaban kurang setuju 0% sehingga dapat dikatakan website sudah berjalan dengan baik. Kesimpulan akhir untuk perbandingan metode *k-means* dan *k-medoids* dapat dikatakan bahwa metode *k-means* lebih baik dibandingkan *k-medoids* dalam segi *confusion matrix* (*accuracy*, *specificity*, *sensitivity*) dan perbandingan implementasi metode pada sistem. Penelitian selanjutnya menggunakan *feature* yang lebih banyak seperti umur, jenis kelamin, lingkaran lengan atas dengan *feature* bersifat *categorical* bisa dilakukan *label encoder* atau *one-hot encoder* agar berbentuk data numerik. serta bisa menggunakan metode atau algoritma yang memang dikhususkan untuk klasifikasi seperti *decision tree*, *k-nearest neighbour*, *naïve bayes*, *svm*, *ensemble model*, dan menggunakan metode *deep learning* seperti *cnn* atau *rnn*.

DAFTAR PUSTAKA

- [1] N. Rizki Octaviyani and R. Mayasari, "Implementasi Algoritma K-Means Clustering Status Gizi Balita," *Jurnal Ilmiah Wahana Pendidikan*, vol. 2022, no. 13, pp. 370–381, doi: 10.5281/zenodo.6962588.
- [2] E. Irfiani, S. Sulistia Rani, S. Nusa Mandiri Jl Kramat Raya No, and J. Pusat, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," vol. 6, no. 4, pp. 17–27, 2018.
- [3] W. Mega, "CLUSTERING MENGGUNAKAN METODE K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA," 2015.
- [4] H. Sastypratiwi, R. Dwi, and N. #2, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Data Artikel Sistem Pakar Menggunakan Metode Systematic Review".
- [5] S. Nurajizah and A. Salbinda, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *Jurnal Teknik Komputer AMIK BSI*, vol. 7, no. 2, 2021, doi: 10.31294/jtk.v4i2.
- [6] K. Singgahan, K. N. Tuban Tri Dian Mustika, and M. Wahini, "POLA ASUH MAKAN ANTARA IBU BEKERJA DAN TIDAK BEKERJA DAN FAKTOR YANG MEMPENGARUHI STATUS GIZI ANAK USIA SEKOLAH DASAR (Kasus di Desa Tingkis)," 2015.
- [7] R. Setiadi, J. Kebidanan, P. Kemenkes Kalimantan Timur, J. Wolter Monginsidi No, J. Keperawatan, and J. Wolter, "FAKTOR-FAKTOR YANG MEMPENGARUHI STATUS GIZI BALITA: SYSTEMATIC REVIEW," 2020.
- [8] N. Nuraliza and S. Herlina, "Hubungan Pengetahuan dan Pendidikan Ibu terhadap Status Gizi Balita," *Jurnal Kesmas Asclepius*, vol. 1, no. 2, pp. 106–115, Dec. 2019, doi: 10.31539/jka.v1i2.578.
- [9] C. A. Chuzaini, "KLASIFIKASI KELOMPOK BELAJAR PPTQ MENGGUNAKAN METODE K-MEANS CLUSTERING (STUDI KASUS DI PONDOK PESANTREN MABADI'UL HUDA DESA BANJARARUM KECAMATAN SINGOSARI KABUPATEN MALANG)," 2017.
- [10] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," vol. 15, no. 2.
- [11] G. Putri, I. Rani, A. Aziz, M. Priyono, and T. Sulistyono, "IMPLEMENTASI EUCLIDEAN DAN CHEBYSHEV DISTANCE PADA K-MEDOIDS CLUSTERING," 2022.
- [12] M. Khoncita Dasriana Bau, Y. Setyawan, M. Titah Jatipaningrum, J. Statistika, F. Sains Terapan, and I. AKPRIND Yogyakarta, "PERBANDINGAN METODE ALGORITMA K-MEANS DAN K-MEDOIDS PADA PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI NUSA TENGGARA TIMUR BERDASARKAN DIMENSI INDEKS PEMBANGUNAN MANUSIA TAHUN 2020," *Jurnal Statistika Industri dan Komputasi*, vol. 08, no. 1, pp. 48–57, 2023, [Online]. Available: <https://ntt.bps.go.id>.