



## Klasifikasi Status Gizi Balita Menggunakan Algoritma K-Nearest Neighbor (KNN)

Hafsah. HS<sup>1</sup>, Nurul Azmi<sup>2</sup>, Hazriani<sup>3</sup>, Yuyun<sup>4</sup>

<sup>1,2</sup>Sistem Komputer, Program Pascasarjan, Universitas Handayani Makassar

<sup>3,4</sup>Badan Riset dan Inovasi Nasional (BRIN)

[Hafsah.hsulaeman@gmail.com](mailto:Hafsah.hsulaeman@gmail.com)

### Abstract

Nutrition is an important component contained in food which includes carbohydrates, protein, vitamins, minerals, fat and water which are needed by the body for growth, development and maintenance which are used directly by the body to repair body tissue. Nutritional needs are an important factor in the growth and development of children, especially children aged under five years (toddlers), because what happens in the first five years determines their growth and development year after year. To achieve good growth and development, strong nutrition is needed. Assessment of the nutritional status of toddlers can be determined through human body measurements known as anthropometry. The reference standards for toddler nutritional status are Body Weight according to Age (WW/U) which describes the child's relative weight for age, Body Weight according to Height (WW/TB) which describes whether the child's weight is in accordance with his height growth and Height according to Age (TB/U) describes a child's height growth based on age. The method used to determine the nutritional status of toddlers is the KNN method, which is to find the closest distance between the evaluated data and a number of  $K$  neighbors closest to the test data. Toddler nutrition data uses 4 classifications, namely insufficient, normal, poor and more. The amount of data used in this research was 170 data with a data composition of 90% consisting of 150 training data and 10% testing data totaling 20 data. Data that is normalized using Z-Score gets an accuracy of 95%, class precision of 98.08% and class recall of 93.75%, data normalized using the min-max technique gets an accuracy of 85%, class precision of 95.00% and class recall of 79.17%, Meanwhile, KNN modeling without normalization produces an accuracy of 80%, precision of 82.4% and a class recall value of 75%.

**Keywords:** *K-Neares Neioghbour, Klasifikasi, Gizi balita*

### Abstrak

Gizi merupakan komponen penting yang terkandung dalam makanan yang meliputi karbohidrat, protein, vitamin, mineral, lemak dan air yang diperlukan oleh tubuh untuk pertumbuhan, perkembangan, dan pemeliharaan yang dimanfaatkan secara langsung oleh tubuh untuk memperbaiki jaringan tubuh. Kebutuhan gizi merupakan salah satu faktor penting dalam tumbuh kembang anak, terutama anak usia bawah lima tahun (balita), karena apa yang terjadi pada lima tahun pertama sangat menentukan tahun demi tahun pertumbuhan dan perkembangannya. Untuk mencapai tumbuh kembang yang baik diperlukan nutrisi yang kuat. Penilaian status gizi balita dapat ditentukan melalui pengukuran tubuh manusia yang dikenal dengan istilah antropometri. Standar acuan status gizi balita adalah Berat Badan menurut Umur (BB/U) yang menggambarkan berat badan relative anak dengan umur, Berat Badan menurut Tinggi Badan (BB/TB) yang menggambarkan apakah berat badan anak sesuai dengan pertumbuhan tinggi badannya dan Tinggi Badan menurut Umur (TB/U) menggambarkan pertumbuhan tinggi badan anak berdasarkan umurnya. Metode yang digunakan dalam penentuan status gizi balita ini adalah metode KNN yaitu mencari jarak paling terdekat diantara data yang terevaluasi dengan sejumlah  $K$  tetangga (*neighbor*) paling dekat dengan data uji. Data gizi balita menggunakan 4 klasifikasi yaitu kurang, normal, buruk dan lebih. Jumlah data yang digunakan dalam penelitian ini adalah 170 data dengan komposisi data 90% merupakan data training sebanyak 150 dan 10% data testing sebanyak 20 data. Data yang dinormalisasi menggunakan Z-Score mendapatkan akurasi sebesar akurasi 95% presisi kelas 98.08% dan recall kelas 93.75%, data yang dinormalisasi menggunakan teknik *min-max* mendapatkan akurasi sebesar 85% presisi kelas 95.00% dan recall kelas 79.17%, sedangkan pemodelan KNN tanpa dinormalisasi menghasilkan akurasi sebanyak 80%, presisi 82.4% dan nilai recall kelas 75%.

**Kata kunci:** *K-Neares Neioghbour, Klasifikasi, Gizi balita*

### 1. Pendahuluan

Gizi merupakan komponen penting yang terkandung dalam makanan yang meliputi karbohidrat, protein, vitamin, mineral, lemak dan air yang diperlukan oleh tubuh untuk pertumbuhan, perkembangan, dan

pemeliharaan yang dimanfaatkan secara langsung oleh tubuh untuk memperbaiki jaringan tubuh. Gizi sangat diperlukan oleh setiap orang khususnya pada masa balita karena gizi berfungsi untuk mempertinggi derajat kesehatan. [1]

Status gizi anak balita adalah cerminan ukuran terpenuhinya kebutuhan gizi anak balita yang didapatkan dari asupan dan penggunaan zat gizi oleh tubuh. Status gizi dapat ditentukan dengan pemeriksaan klinis, pengukuran antropometri, analisis biokimia, dan riwayat gizi.

Kebutuhan gizi merupakan salah satu faktor penting dalam tumbuh kembang anak, terutama anak usia bawah lima tahun (balita), karena apa yang terjadi pada lima tahun pertama sangat menentukan tahun demi tahun pertumbuhan dan perkembangannya. Untuk mencapai tumbuh kembang yang baik diperlukan nutrisi yang kuat. Makanan yang kurang baik secara kualitas maupun kuantitas akan menyebabkan gizi kurang. Keadaan gizi kurang dapat menyebabkan gangguan pertumbuhan dan perkembangan.

Salah satu cara memenuhi kebutuhan gizi tersebut adalah dengan cara mengkonsumsi gizi yang cukup dan sesuai untuk tubuh. Selain itu orang tua juga perlu mengetahui tingkat kesehatan si balita yang dapat dilihat dari status gizinya. Penilaian status gizi balita dapat ditentukan melalui pengukuran tubuh manusia yang dikenal dengan istilah antropometri. Standar acuan status gizi balita adalah Berat Badan menurut Umur (BB/U) yang menggambarkan berat badan relative anak dengan umur, Berat Badan menurut Tinggi Badan (BB/TB) yang menggambarkan apakah berat badan anak sesuai dengan pertumbuhan tinggi badannya dan Tinggi Badan menurut Umur (TB/U) menggambarkan pertumbuhan tinggi badan anak berdasarkan umurnya.

Oleh karena itu, maka perlu menerapkan metode klasifikasi untuk menentukan status gizi anak karena metode ini memungkinkan kita untuk mengorganisir data gizi anak-anak ke dalam kategori yang lebih mudah dipahami dan digunakan oleh para profesional kesehatan, peneliti dan pembuat kebijakan, hal ini membantu dalam menyajikan informasi yang lebih terstruktur dan mudah diinterpretasikan. Pada sistem ini pengguna akan menginputkan data balita yang telah diketahui. Alasan menggunakan metode KNN karena metode ini merupakan salah satu pendekatan yang digunakan dalam pengklasifikasian secara mudah dan efisien.

Konsep mendasar dari algoritma KNN yaitu mencari jarak paling terdekat diantara data yang terevaluasi dengan sejumlah K tetangga (*neighbor*) paling dekat dengan data uji. KNN bekerja dengan cara membandingkan data uji dan data training. KNN mencari jarak data training yang paling mendekati dengan data uji. Perhitungan jarak menggunakan rumus *Euclidean Distance*.

Hasil dari penelitian ini berupa status gizi balita menggunakan standar acuan Berat Badan menurut

Umur (BB/U). Klasifikasi dari status gizi balita ini, yaitu gizi buruk, gizi kurang, gizi baik dan gizi lebih.

## 2. Metode Penelitian

### 2.1. Pengumpulan Data

Pengumpulan data ini bertujuan untuk mendapatkan data yang digunakan untuk penelitian. Data diperoleh melalui Url

<https://www.kaggle.com/datasets/mjalaluddinassuyuti/baby-nutrition-classification> data yang dikumpulkan merupakan data Gizi Balita Pada Tahun 2019.

### 2.2. Pemodelan K-Nearest Neighbor

Langkah selanjutnya adalah menyiapkan dataset, data training, dan data testing dengan menggunakan metode split data 90% dan 10%. Model yang dipilih dalam penelitian ini adalah K-Nearest Neighbor (KNN) dengan menggunakan tiga bentuk pemodelan yaitu pemodelan pertama menggunakan data yang tidak dinormalisasi, yang kedua menggunakan data yang dinormalisasi menggunakan metode *Z-Score*, dan yang ketiga menggunakan data yang dinormalisasikan menggunakan metode *min-max*. Tujuannya adalah untuk menentukan seberapa besar pengaruh normalisasi terhadap kinerja algoritma. Langkah-langkah dalam pemodelan KNN adalah sebagai berikut:

Penentuan nilai K. K adalah jumlah data tetangga terdekat, tidak ada standar baku di dalamnya menentukan nilai K.

Hitung jarak dari data uji dengan setiap baris data pelatihan. Perhitungannya menggunakan rumus jarak. Rumus jarak yang digunakan dalam penelitian ini adalah jarak Euclidean dengan Persamaan berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (1)$$

Dimana:

$d(x, y)$  = Jarak antara data pengujian baru dan data pelatih

$x_1$  = Data latih

$y_1$  = Data uji

Tabel 1. Table data uji normalisasi Z-Score

	UMUR (BLN)	BERAT(KG)	TINGGI (CM)
Mean	32.11176	8.773529	73.96765
standar deviasi	9.203446	2.146954	10.9519

$i$  = Jumlah baris data dari data uji

Urutkan hasil pengukuran jarak dari nilai terkecil ke besar.

Ambil himpunan kelas terbanyak dari kelima data latih dengan jarak terpendek sebagai kelas uji yang baru data.

Evaluasi

Evaluasi mengukur kinerja pemodelan yang telah dibuat dalam penelitian ini. Pengukuran evaluasi menggunakan *confussion matrix*, dan dari *confussion matrix*, *matrix* dianalisis seberapa besar akurasi, rata-rata kelas presisi, dan rata-rata kelas *recall*. Semakin besar nilai ini, semakin baik pemodelan dilakukan.

Akurasi adalah nilai yang menggambarkan seberapa akurat model mengklasifikasikan dengan benar atau rasio keberhasilan model dalam memprediksi data sebenarnya. Dengan kata lain, akurasi adalah seberapa dekat hasil prediksi model dengan data sebenarnya. Proses ini disebut proses pengukuran ketepatan dan kesalahan. Dapat dilihat persamaan berikut:

$$\text{Akurasi} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total data}} \times 100\% \quad (2)$$

Presisi adalah gambaran tingkat akurasi antara data yang diminta dengan hasil yang diberikan oleh model, yaitu rasio hasil prediksi positif yang sebenarnya terhadap semua hasil prediksi positif.

$$\text{Presisi} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \times 100\% \quad (3)$$

Recall atau sensitivitas adalah nilai keberhasilan model dalam menemukan kembali informasi atau membandingkan jumlah prediksi benar positif dengan jumlah data benar positif. Recall dapat dirumuskan dengan

	UMUR (BLN)	BERAT(KG)	TINGGI (CM)
min (x)	4	3.4	50
max (x)	54	14	98

Tabel 2. Table data uji normalisasi *Recall*

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False negative}} \times 100\% \quad (4)$$

### 3. Hasil dan Pembahasan

#### 3.1. Data Preparation

Untuk pemodelan menggunakan normalisasi, diadakan proses normalisasi sebelum proses split. Teknik normalisasi yang digunakan adalah Z-Score dan min-max. untuk proses normalisasi Z-score dapat dilihat pada persamaan berikut:

$$x_{stand} = \frac{x - \text{mean}(x)}{\text{stdev}(x)} \quad (5)$$

Table berikut adalah nilai mean dan standar deviasi setiap atribut

Table 1. nilai Mean dan standar Deviasi setiap atribut

Setelah menentukan nilai rata-rata dan standar deviasi setiap atribut, langkah selanjutnya adalah menghitung nilai normalisasi setiap atribut dari setiap record sampai akhir dengan menggunakan rumus Z-Score. Contoh untuk mendapatkan hasil normalisasi menggunakan Z-score adalah sebagai berikut:

$$x_1 = \frac{24 - 32.11176}{9.203446} = -0.881383457$$

Proses diatas dilakukan pada setiap atribut untuk semua baris data training.

Untuk normalisasi menggunakan *min-max* dapat dilihat pada persamaan berikut ini:

$$x_{norm} = \frac{x' - \min(x)}{\max(x) - \min(x)} \quad (6)$$

Table berikut adalah nilai *min-max* setiap atribut:

Table 2. nilai *min-max* setiap atribut

Setelah menentukan *min-max* setiap atribut, langkah selanjutnya adalah menghitung nilai normalisasi setiap atribut dari setiap record sampai akhir dengan menggunakan rumus *min-max*. Contoh untuk mendapatkan hasil normalisasi menggunakan *min-max* adalah sebagai berikut

$$x_{norm} = \frac{24 - 4}{54 - 4} = 0.4$$

Proses diatas dilakukan pada setiap atribut untuk semua baris data training.

Untuk pemodelan tanpa normalisasi data, Data langsung dibagi tanpa proses normalisasi. Sedangkan model kedua dan ketiga data dibagi setelah proses normalisasi. Proses perpecahan ini menghasilkan dua data yang berbeda yaitu data training dan data testing, dengan perbandingan 9:1, 90% (150 baris) untuk data training dan 10% (20 baris) untuk menguji data.

#### 3.2. Algoritma KNN

KNN (K-Nearest Neighbors) adalah salah satu algoritma yang digunakan dalam klasifikasi dan regresi dalam pembelajaran mesin. Algoritma KNN memprediksi klasifikasi suatu data baru berdasarkan klasifikasi data tetangga terdekatnya.

Pada dasarnya, algoritma KNN bekerja dengan cara mencari k-nearest neighbors (tetangga terdekat) dari data baru yang akan diprediksi. "k" dalam KNN mengacu pada jumlah tetangga terdekat yang akan dipertimbangkan dalam pengambilan keputusan. Setelah tetangga terdekat ditemukan, algoritma KNN melakukan mayoritas voting atau penghitungan rata-rata untuk memprediksi label klasifikasi atau nilai regresi data baru tersebut. Berikut langkah-langkah KNN

Tentukan jumlah tetangga.

Untuk setiap data uji yang akan diprediksi, hitung jarak antara data tersebut dengan setiap data latihan dalam dataset menggunakan persamaan jarak *Euclidian*. Sebagai contoh, mencari jarak antara atribut data ke 170 dengan atribut data pertama pada data yang ternormalisasi menggunakan Z-Score

$$\begin{aligned} d(x, y) &= \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \\ &= \sqrt{(1.291715712 - (-0.881383457))^2 + ((-0.03424824) - (-1.384998821))^2 + (0.185570732 - (-0.818820783))^2} \\ &= 2.748761 \end{aligned}$$

untuk menghitung jarak ecludien pada data non ternormalisasi dan data ternormalisasi menggunakan *min-max* dilakukan cara yang sama pada data ternormalisasi *Z-Score* Urutkan data berdasarkan jarak Eucliden terkecil kemudian tentukan 5 jarak terkecil.

Tentukan kelas dari data testing dengan mengambil kelas mayoritas dari kumpulan kelas sejumlah K. Table berikut adalah table data uji ternormalisasi *Z-Score*

Tabel 3. Table data uji normalisasi *Z-Score*

N O	NAM A BALITA	Z SCORE UMUR (BLN)	Z SCORE BERAT (KG)	Z SCORE TINGGI (CM)	Stats Gizi	K=5	Ket
1	-	-	-	-	-	-	-
5	sinca	0.0121	1.037	0.5448	GIZI_N	Nor	Sesuai
1	nk	43789	0367	95824	ORMAL	mal	
1	-	-	-	-	-	-	-
5	0.9900	1.502	0.2768	GIZI_N	Nor		
2	dila	38415	8128	79051	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
5	0.2051	1.897	1.1899	GIZI_BU	bur		
3	dikae	66127	3525	62246	RUK	uk	Sesuai
1	-	-	-	-	-	-	-
5	0.2051	0.453	0.1855	GIZI_KU	Kur		
4	mita	66127	4467	70732	RANG	ang	Sesuai
1	-	-	-	-	-	-	-
5	0.9900	0.478	0.0942	GIZI_N	Nor		
5	sika	38415	1054	62412	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
5	1.4003	0.080	0.8188	GIZI_N	Nor		
6	ika	7067	8258	20783	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
5	0.2051	0.105	0.8188	GIZI_N	Nor		
7	safa	66127	4846	20783	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
5	0.9900	0.571	0.5448	GIZI_N	Nor		
8	jina	38415	2606	95824	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
5	0.2051	1.037	0.1855	GIZI_N	Nor		
9	june	66127	0367	70732	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
6	nadir	0.5311	1.968	1.3725	GIZI_N	Nor	
0	a	31003	5888	78885	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
6	syakir	1.1830	1.502	0.8188	GIZI_LE	Lebi	
1	a	60753	8128	20783	BIH	h	Sesuai
1	-	-	-	-	-	-	-
6	1.2917	0.431	0.9101	GIZI_N	Nor		
2	aira	15712	5278	29102	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
6	0.4224	0.478	1.6405	GIZI_LE	Lebi		
3	kayla	76044	1054	95658	BIH	h	Sesuai
1	-	-	-	-	-	-	-
6	0.2051	0.453	1.8232	GIZI_LE	Nor		
4	ardi	66127	4467	12297	BIH	mal	tidak sesuai
1	-	-	-	-	-	-	-
6	0.0121	0.012	0.8188	GIZI_N	Nor		
5	bita	43789	3294	20783	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
6	1.1830	0.384	0.1855	GIZI_N	Nor		
6	gita	60753	9502	70732	ORMAL	mal	Sesuai
1	-	-	-	-	-	-	-
6	3.0544	0.058	0.3681	GIZI_KU	Kur		
7	suci	82625	907	87371	RANG	ang	Sesuai
1	-	-	-	-	-	-	-
6	0.9900	0.384	1.6405	GIZI_LE	Lebi		
8	dama	38415	9502	95658	BIH	h	Sesuai
1	-	-	-	-	-	-	-
6	0.0965	0.965	0.6362	GIZI_KU	Kur		
9	indah	11169	8004	04144	RANG	ang	Sesuai
1	-	-	-	-	-	-	-
7	1.2917	0.034	0.1855	GIZI_N	Nor		
0	dera	15712	2482	70732	ORMAL	mal	Sesuai

Berdasarkan hasil uji normalisasi *Z-Score* maka terdapat satu data uji yang tidak sesuai dengan data actual dan data prediksi dimana hasil prediksi menunjukkan gizi Normal sedangkan data aktualnya adalah Gizi Lebih.

Tabel 4. hasil data uji ternormalisasi *min-max*

N O	NAM A BALITA	Min - max UMUR (BLN)	Min - max BERAT (KG)	Min - max TINGGI (CM)	STAT US GIZI	K=5	Ket
15	sinca	0.56	0.7169	6.0943	GIZI_NOR	Normal	Sesuai
1	nk	-	811	396	MAL		
15	dila	0.38	0.8113	6.9433	GIZI_NOR	Normal	Sesuai
2	-	-	208	962	MAL		
15	dika	0.6	0.1226	7.8867	GIZI_BUR	Buruk	Sesuai
3	ee	-	415	925	UK		
15	mita	0.6	0.4150	6.8490	GIZI_KUR	Normal	Tidak Sesuai
4	-	-	943	566	ANG		
15	sika	0.38	0.6037	6.7547	GIZI_NOR	Normal	Sesuai
5	-	-	736	17	MAL		
15	ika	0.82	0.4905	5.8113	GIZI_NOR	Normal	Sesuai
6	-	-	66	208	MAL		
15	safa	0.6	0.5283	5.8113	GIZI_NOR	Normal	Sesuai
7	-	-	019	208	MAL		
15	jina	0.38	0.6226	6.0943	GIZI_NOR	Normal	Sesuai
8	-	-	415	396	MAL		
15	june	0.6	0.7169	6.8490	GIZI_NOR	Normal	Sesuai
9	-	-	811	566	MAL		
16	nadir	0.66	0.9056	8.0754	GIZI_NOR	Normal	Sesuai
0	a	-	604	717	MAL		
16	syaki	0.78	0.8113	5.8113	GIZI_LEBIH	Normal	Tidak Sesuai
1	ra	-	208	208			
16	aira	0.8	0.5943	5.7169	GIZI_NOR	Normal	Sesuai
2	-	-	396	811	MAL		
16	kayla	0.64	0.6037	4.9622	GIZI_LEBIH	Lebih	Sesuai
3	-	-	736	642			
16	ardi	0.6	0.4150	4.7735	GIZI_LEBIH	Normal	Tidak Sesuai
4	-	-	943	849			
16	bita	0.56	0.5094	5.8113	GIZI_NOR	Normal	Sesuai
5	-	-	34	208	MAL		
16	gita	0.78	0.5849	6.8490	GIZI_NOR	Normal	Sesuai
6	-	-	057	566	MAL		
16	suci	0	0.5188	7.0377	GIZI_KUR	Kurang	Sesuai
7	-	-	679	358	ANG		
16	dam	0.38	0.5849	4.9622	GIZI_LEBIH	Lebih	Sesuai
8	ai	-	057	642			
16	inda	0.58	0.3113	6	GIZI_KUR	Kurang	Sesuai
9	h	-	208		ANG		
17	dera	0.8	0.5	6.8490	GIZI_NOR	Normal	Sesuai
0	-	-		566	MAL		

Berdasarkan hasil uji normalisasi *Min-max* maka terdapat 3 data uji yang tidak sesuai dengan data actual dan data prediksi dimana satu hasil prediksi

menunjukkan gizi Normal sedangkan data aktualnya adalah Gizi Kurang serta dua diantaranya adalah hasil prediksi menunjukkan gizi normal sedangkan data aktualnya adalah gizi lebih.

Tabel 5. Hasil data uji tanpa normalisasi

	NAM A BALI TA	UM UR (BLN )	BERA T(KG)	TIN GG I (C M)	STATUS GIZI	K=5	Ket
15	sinca nk	32	11	68	GIZI_NOR MAL	Normal	sesu ai
15	dila	23	12	77	GIZI_NOR MAL	Normal	sesu ai
15	dikae	34	4.7	87	GIZI_BURU K	Buruk	sesu ai
15	mita	34	7.8	76	GIZI_KURA NG	Normal	tidak sesu ai
15	sika	23	9.8	75	GIZI_NOR MAL	Normal	sesu ai
15	ika	45	8.6	65	GIZI_NOR MAL	Normal	sesu ai
15	safa	34	9	65	GIZI_NOR MAL	Normal	sesu ai
15	jina	23	10	68	GIZI_NOR MAL	Normal	sesu ai
15	june	34	11	76	GIZI_NOR MAL	Normal	sesu ai
16	nadir a	37	13	89	GIZI_NOR MAL	Kurang	tidak sesu ai
16	syaki ra	43	12	65	GIZI_LEBIH	Normal	tidak sesu ai
16	aira	44	9.7	64	GIZI_NOR MAL	Normal	sesu ai
16	kayla	36	9.8	56	GIZI_LEBIH	Lebih	sesu ai
16	ardi	34	7.8	54	GIZI_LEBIH	Lebih	sesu ai
16	bita	32	8.8	65	GIZI_NOR MAL	Normal	sesu ai
16	gita	43	9.6	76	GIZI_NOR MAL	Normal	sesu ai
16	suci	4	8.9	78	GIZI_KURA NG	Kurang	sesu ai
16	dam ai	23	9.6	56	GIZI_LEBIH	Lebih	sesu ai
16	inda h	33	6.7	67	GIZI_KURA NG	Normal	tidak sesu ai
17	dera	44	8.7	76	GIZI_NOR MAL	Normal	sesu ai

Berdasarkan hasil uji tanpa normalisasi maka terdapat 4 data uji yang tidak sesuai dengan data actual dan data prediksi dimana satu hasil prediksi menunjukkan gizi Normal sedangkan data aktualnya adalah Gizi Kurang serta dua diantaranya adalah hasil prediksi menunjukkan gizi normal sedangkan data aktualnya adalah gizi Kurang, satu data uji menghasilkan prediksi gizi normal sedangkan data aktualnya adalah gizi lebih.

### 3.3. Evaluasi

Evaluasi menggunakan *confussion matrix*, diperoleh bahwa pemodelan KNN menggunakan data yang tidak ternormalisasi, dengan nilai K=5, hanya mampu mengklasifikasikan 16 data dengan benar dari 20 data uji dengan akurasi 80%, rata-rata presisi 82.4% dan

rata-rata nilai recall kelas 75% seperti yang terlihat pada table berikut

Table 6. evaluasi pemodelan tidak ternormalisasi

		Kelas Prediksi				Truth Overall	recall
		Kurang	Normal	Buruk	Lebih		
Kelas Aktual	kurang	1	2	0	0	3	33.33%
	normal	1	11	0	0	12	91.67%
	buruk	0	0	1	0	1	100.00%
	lebih	0	1	0	3	4	75.00%
Classification Overall		2	14	1	3	20	
presisi		50.00%	78.57%	100.00%	100.00%		
akurasi		80.00%					

Sementara itu, pemodelan dengan menggunakan normalisasi Z-Score mampu mengklasifikasikan 19 data dengan benar dari 20 data uji dengan akurasi 95% rata-rata presisi kelas 98.08% dan rata-rata recall kelas 93,75% seperti yang terlihat pada tabel 9

Table 7. evaluasi pemodelan normalisasi Z-Score

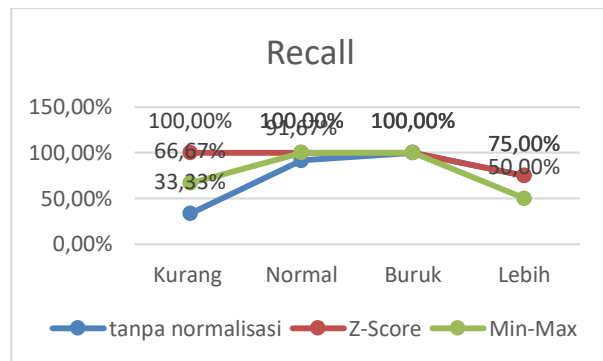
		Kelas Prediksi				Tru th Ov eral l	recall
		Kuran g	Norm al	Buruk	Lebih		
Kelas Aktual	kuran g	3	0	0	0	3	100.00%
	norm al	0	12	0	0	12	100.00%
	buruk	0	0	1	0	1	100.00%
	lebih	0	1	0	3	4	75.00%
Classificati on Overall		3	13	1	3	20	
presisi		100.00 %	92.31 %	100.00 %	100.00 %		
akurasi		95.00%					

Untuk pemodelan dengan menggunakan normalisasi min-max mampu mengklasifikasikan 17 data dengan benar dari 20 data uji dengan akurasi 85% rata-rata presisi kelas 95.00% dan rata-rata recall kelas 79,17% seperti yang terlihat pada table

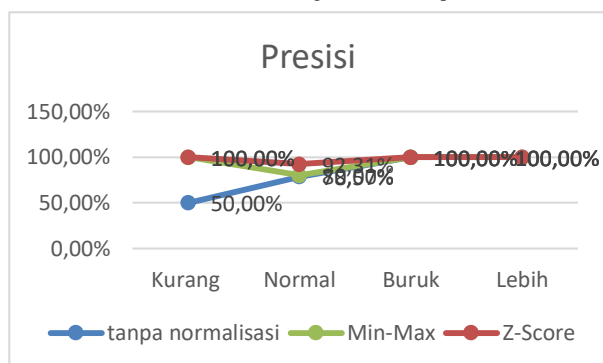
Table 8. evaluasi pemodelan normalisasi min-max

		Kelas Prediksi				True Overall	recall
		Kurang	Normal	Buruk	Lebih		
Kelas Aktual	kurang	2	1	0	0	3	66.67%
	normal	0	12	0	0	12	100.00%
	buruk	0	0	1	0	1	100.00%
	lebih	0	2	0	2	4	50.00%
Classification Overall		2	15	1	2	20	
presisi		100.00%	80.00%	100.00%	100.00%		
akurasi	85.00%						

Perbandingan model menggunakan data yang tidak dinormalisasi, data yang normalisasi Z-Score dan data yang ternormalisasi menggunakan min-max seperti pada gambar, terlihat bahwa Normalisasi data menggunakan teknik Z-Score adalah pemodelan yang terbaik dengan akurasi 95% recall 93,75% dan presisi 98.08% terlihat pada gambar 1 dan gambar 2.



Gambar 1. Perbandingan recall setiap kelas



Gambar 2. Perbandingan presisi setiap kelas

#### 4. Kesimpulan

Berdasarkan hasil dan pembahasan mengenai penentuan status gizi balita, penelitian ini

menyimpulkan bahwa pemodelan KNN menggunakan data yang dinormalisasi menggunakan teknik Z-Score dan teknik min-max menghasilkan akurasi yang lebih tinggi daripada pemodelan KNN tanpa normalisasi. Data yang dinormalisasi menggunakan Z-Score mendapatkan akurasi sebesar akurasi 95% presisi kelas 98.08% dan recall kelas 93,75%, data yang dinormalisasi menggunakan teknik min-max mendapatkan kurasi sebesar akurasi 85% presisi kelas 95.00% dan recall kelas 79,17%, sedangkan pemodelan KNN tanpa dinormalisasi menghasilkan akurasi sebanyak 80%, presisi 82.4% dan nilai recall kelas 75%. Dengan demikian untuk mendapatkan hasil yang lebih baik maka dataset harus dinormalisasi

#### Daftar Rujukan

- [1] S. Anraeni, E. R. Melani, and H. Herman, "Ripeness Identification of Chayote Fruits using HSI and LBP Feature Extraction with KNN Classification," *Ilk. J. Ilm.*, vol. 14, no. 2, pp. 150–159, 2022, doi: 10.33096/ilkom.v14i2.1153.150-159.
- [2] D. Fitrianiingsih, M. Bettiza, and A. Uperiati, "Klasifikasi Status Gizi Pada Pertumbuhan Balita Menggunakan K-Nearest Neighbor (Knn)," *Student Online J.*, vol. 2, no. 1, pp. 106–111, 2021.
- [3] M. H. Wahyudi, "Sistem Pendukung Keputusan Penentuan Status Gizi Balita Menggunakan Metode Naive Bayes," *Semin. Nas. Teknol. Inf. dan Multi Media*, no. 1, pp. 25–30, 2018.
- [4] D. A. N. Wulandari and A. Prasetyo, "Sistem Penunjang Keputusan Untuk Menentukan Status Gizi Balita Menggunakan Metode Fuzzy Tsukamoto," *J. Inform.*, vol. 5, no. 1, pp. 22–33, 2018, doi: 10.31311/ji.v5i1.2440.
- [5] A. Yudhana, I. Riadi, and M. R. Djou, "Determining eligible villages for mobile services using k- NN algorithm," *Ilk. J. Ilm.*, vol. 15, no. 1, pp. 11–20, 2023.
- [6] M. Yunus and N. K. A. Pratiwi, "Prediksi Status Gizi Balita Dengan Algoritma K-Nearest," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 221–231, 2023.