

# Klasifikasi Status Gizi Balita Menggunakan Metode *K-Nearest Neighbor*

**Sarah Syifani<sup>\*1</sup>, Anindita Septiarini<sup>2</sup>, Medi Taruk<sup>3</sup>, Masna Wati<sup>4</sup>, Andi Tejawati<sup>5</sup>**  
<sup>1,2,3,4,5</sup>Program Studi Informatika, Fakultas Teknik, Universitas Mulawarman, Samarinda  
e-mail: <sup>1</sup>sarahsyifa34@gmail.com, <sup>2</sup>anindita@unmul.ac.id, <sup>3</sup>meditaruk@unmul.ac.id,  
<sup>4</sup>masnawati@fkti.unmul.ac.id, <sup>5</sup>anditejawati@unmul.ac.id

## **Abstrak**

Status gizi balita merupakan indikator penting dalam menilai tingkat kesehatan anak yang dapat diketahui melalui pemeriksaan antropometri. Berdasarkan data Survei Status Gizi Indonesia (SSGI) tahun 2022, prevalensi gizi kurang di Provinsi Kalimantan Timur mencapai 23,9%, khususnya di Kota Samarinda sebesar 25,3%, yang menunjukkan bahwa masalah gizi masih perlu mendapatkan perhatian. Penelitian ini bertujuan untuk mengklasifikasikan status gizi balita menggunakan metode *K-Nearest Neighbor* (KNN) dengan variasi nilai *K* dan rumus jarak yang berbeda guna memperoleh performa terbaik. Data yang digunakan merupakan data sekunder dari Puskesmas Pasundan, Kota Samarinda, sebanyak 760 data balita usia 0–60 bulan. Tahapan penelitian meliputi pengumpulan data, perancangan data melalui proses preprocessing (data selection, penanganan outlier, dan transformasi data menggunakan min-max normalization serta encoding), implementasi metode KNN dengan variasi nilai *K* (1, 3, 5, 7, 9, 11) dan rumus jarak (Euclidean, Manhattan, Minkowski), serta evaluasi model menggunakan Confusion Matrix Multiclass. Berdasarkan hasil pengujian, akurasi tertinggi diperoleh sebesar 89,24% dengan nilai presisi 66,29%, recall 63,70%, dan F1-score 63,12% menggunakan nilai *K* = 1 dan rumus jarak Euclidean. Hasil ini menunjukkan bahwa metode KNN mampu memberikan performa klasifikasi yang baik dalam menentukan status gizi balita berdasarkan atribut usia, jenis kelamin, berat badan, dan tinggi badan.

**Kata kunci** — Status Gizi Baita, *K-Nearest Neighbor*, Klasifikasi, Data Mining

## 1. PENDAHULUAN

Status gizi balita merupakan tolak ukur terpenuhinya asupan gizi balita yang dapat diketahui melalui pemeriksaan klinis, pengukuran antropometri, analisis biokimia, dan riwayat gizi [1]. Status gizi yang baik atau buruk menggambarkan tingkat Kesehatan pada balita (pangumbaran). Balita dengan kondisi gizi yang buruk akan mengalami gangguan perkembangan, mudah sakit, bahkan kematian [2]. Dampak lain dari gizi buruk yaitu kondisi psikososial, tidur, pernapasan yang terganggu, dan dapat menimbulkan masalah penyakit degenerative saat dewasa [3].

Berdasarkan peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 Tentang Standar Antropometri Anak, penilaian status gizi balita menggunakan parameter seperti umur, jenis kelamin, berat badan, dan tinggi atau panjang badan [4]. Berdasarkan data dari SSGI (Survei Status Gizi Indonesia) tahun 2022 angka prevalensi di Kalimantan Timur mencapai 23,9% khususnya di kota Samarinda mencapai 25,3% yang menunjukkan bahwa masalah gizi masih menjadi isu kesehatan yang perlu diperhatikan [5].

Pada era teknologi saat ini, data mining dapat digunakan untuk mengetahui status gizi balita dengan menerapkan metode *K-Nearest Neighbor* [6]. Metode KNN dapat menganalisis

data berukuran besar dengan baik, namun metode KNN bergantung pada pemilihan fitur dan atribut yang relevan [7]. Untuk mempelajari dan mengenali pola dari data yang besar, dapat menggunakan teknik data mining [8]. Data mining adalah proses menemukan pola, kesamaan, dan tren yang bermakna dari data berukuran besar [9]. Hasil informasi yang didapatkan dari data mining dapat digunakan untuk membantu pengambilan keputusan dan prediksi di berbagai bidang, salah satunya bidang Kesehatan [10].

Penelitian sebelumnya telah menerapkan metode *K-Nearest Neighbor* untuk klasifikasi status gizi balita, namun sebagian besar hanya menggunakan satu nilai K atau satu rumus jarak. Oleh karena itu, penelitian ini bertujuan untuk mengklasifikasikan status gizi balita berdasarkan atribut usia, jenis kelamin, berat badan, dan tinggi badan menggunakan metode *K-Nearest Neighbor* dengan variasi nilai K dan rumus jarak untuk mendapatkan performa terbaik.

## 2. METODE PENELITIAN

### 2.1 Pengumpulan Data

Pengumpulan data dilakukan dengan metode studi dokumentasi, yaitu dengan mengambil data status gizi balita yang telah tersedia di Puskesmas Pasundan, Kota Samarinda, Provinsi Kalimantan Timur. Data yang digunakan yaitu data status gizi balita usia 0-60 bulan dengan jumlah data sebanyak 760 data.

### 2.2 Perancangan Data

Pada tahap ini dilakukan *preprocessing* untuk mempersiapkan data sebelum diolah untuk mendapatkan hasil yang optimal dan efisien [11]. Tahapan yang dilakukan yaitu data *selection* dengan mengurangi atribut yang berlebihan, pengecekan dan penanganan *outlier* dengan menghapus data yang memiliki nilai yang terlalu jauh dari data lainnya, dan transformasi data dengan mengubah data menjadi format yang sesuai sehingga data siap diolah untuk keperluan penelitian. Pengecekan dan penanganan *outlier* menggunakan metode persentil yang membagi data menjadi 100 kelompok dan menentukan batas persentil yang bersifat subjektif [12]. Pada tahap transformasi melakukan normalisasi menggunakan metode min-max dan encoding. Normalisasi menggunakan metode min-max dinyatakan dengan persamaan (1) [13]:

$$x' = \frac{\min_R + (x - x_{\min})(\max_R - \min_R)}{x_{\max} - x_{\min}} \quad (1)$$

Keterangan:

$x'$  : data hasil normalisasi

$x$  : data asli

$x_{\min}$  : nilai minimum pada atribut

$x_{\max}$  : nilai maksimum pada atribut

$\min_R$  : minimum *range*

$\max_R$  : maksimum *range*

### 2.3 Implementasi Metode

Pada tahap ini akan diterapkan proses klasifikasi dengan menggunakan *K-Nearest Neighbor*. Metode KNN melakukan klasifikasi data baru berdasarkan mayoritas kategori pada tetangga terdekatnya [14]. Tahapan implementasi meliputi pemilihan nilai K (3, 5, 7, 9, 11), penggunaan rumus jarak *Euclidean Distance*, *Manhattan Distance*, dan *Minkowski Distance*, pembagian data dengan *Stratified K-Fold Cross Validation*, menghitung dan mengurutkan hasil perhitungan jarak, dan menentukan kelas data uji [15]. Menghitung jarak *Euclidean* antara data *training* dan *testing* dinyatakan dengan persamaan (2) [16]:

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

Keterangan:

$d$  : jarak antara x dan y

$x_i$  : data *training* ke -i

$y_i$  : data *testing* ke -i

Menghitung jarak *Manhattan* antara data training dan testing dinyatakan dengan persamaan (3) [17]:

$$D(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (3)$$

Keterangan:

$d$  : jarak antara x dan y

$x_i$  : data *training* ke -i

$y_i$  : data *testing* ke -i

Menghitung jarak *Minkowski* antara data training dan testing dinyatakan dengan persamaan (4) [17]:

$$D(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^r \right)^{\frac{1}{r}} \quad (4)$$

Keterangan:

$d$  : jarak antara x dan y

$x_i$  : data *training* ke -i

$y_i$  : data *testing* ke -i

$r$  : parameter yang dapat disesuaikan

#### 2.4 Evaluasi Metode

Evaluasi metode dilakukan dengan menggunakan *Confusion Matrix Multiclass* yang memberikan informasi mengenai *True Positive* (TP), *False Positive* (FP), *False Negative* (FN) untuk mengetahui akurasi, presisi, *recall*, dan *F1-Score* [18]. Menghitung metrik evaluasi menggunakan persamaan berikut:

$$Akurasi = \left( \frac{TP}{N} \right) \times 100\% \quad (5)$$

$$Presisi = \left( \frac{TP}{TP+FP} \right) \times 100\% \quad (6)$$

$$Recall = \left( \frac{TP}{TP+FN} \right) \times 100\% \quad (7)$$

$$F1 - Score = \left( 2 \times \frac{precision \times recall}{precision + recall} \right) \times 100\% \quad (8)$$

### 3. HASIL DAN PEMBAHASAN

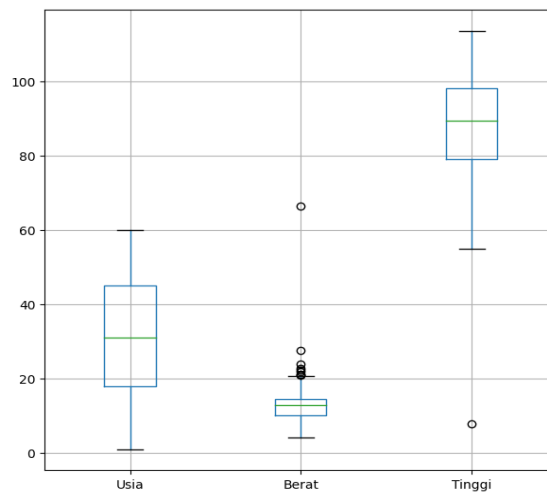
#### 3.1 Perancangan Data

Penelitian ini menggunakan data status gizi balita yang masih mentah dan perlu melalui tahapan perancangan data untuk membantu proses analisis. Dataset awal terdiri dari 24 kolom yang masih mengandung atribut yang tidak diperlukan, sehingga perlu dilakukan data *selection*. Adapun atribut yang digunakan dalam penelitian adalah usia, jenis kelamin, berat badan, tinggi badan, dan BB/TB sehingga data yang tidak dibutuhkan akan dihapus atau di *drop*. Data yang sudah dibersihkan dari atribut yang tidak diperlukan dapat dilihat pada Tabel 1.

Tabel 1 Hasil Data *Selection*

No.	Nama	JK	Usia	Berat	Tinggi	BB/TB
1.	Balita 1	P	58	16.2	106	Normal
2.	Balita 2	L	59	15.8	103.5	Normal
3.	Balita 3	P	56	14.5	101	Normal
4.	Balita 4	L	55	15.4	105	Normal
5.	Balita 5	P	59	16.8	107.5	Normal
⋮	⋮	⋮	⋮	⋮	⋮	⋮
756.	Balita 756	P	34	13	92.4	Normal
757.	Balita 757	L	10	8.8	71.4	Normal
758.	Balita 758	L	5	6	62	Normal
759.	Balita 759	L	4	6.6	65	Normal
760.	Balita 760	L	13	8.8	78	Normal

Selanjutnya dilakukan pengecekan dan penanganan *outlier*, hasil pengecekan *outlier* dapat dilihat pada Gambar 1.



Gambar 1 Hasil Pengecekan Outlier

Hasil pengecekan *outlier* menunjukkan adanya data *outlier* pada atribut berat dan tinggi. Data *outlier* akan dihapus menggunakan metode persentil dengan batas 5%-95%, dimana data yang berada di bawah persentil ke-5 atau diatas persentil ke-95 dianggap sebagai *outlier*. Total data *outlier* yang dihapus sebanyak 100 data, sehingga data yang akan digunakan untuk analisis selanjutnya sebanyak 660 data. Selanjutnya dilakukan transformasi data dengan melakukan normalisasi dan encoding. Contoh untuk mendapatkan hasil normalisasi menggunakan data Balita 2 dengan atribut Usia menggunakan persamaan (1).

$$x' = \frac{0 + (59 - 2)(1 - 0)}{59 - 2} = 1$$

Setelah dilakukan normalisasi pada semua atribut numerik seperti usia, berat badan, dan tinggi badan, selanjutnya dilakukan encoding pada atribut kategorikal yaitu jenis kelamin. Hasil encoding dapat dilihat pada Tabel 2.

Tabel 2 Hasil Encoding

No.	JK	Usia	Berat	Tinggi	BB/TB	
1.	Balita 2	0	1	0.894737	0.9625	Normal
2.	Balita 3	1	0.947368	0.757895	0.9	Normal
3.	Balita 4	0	0.929825	0.852632	1	Normal
4.	Balita 9	1	0.947368	0.863158	0.95	Normal

No.		JK	Usia	Berat	Tinggi	BB/TB
5.	Balita 10	0	0.947368	0.768421	0.86	Normal
⋮	⋮	⋮	⋮	⋮	⋮	⋮
656.	Balita 754	1	0	0.757895	0.87	Normal
657.	Balita 755	0	0.105263	0.063158	0.025	Normal
658.	Balita 756	1	0.561404	0.6	0.685	Normal
659.	Balita 757	0	0.140351	0.157895	0.16	Normal
660.	Balita 760	0	0.192982	0.157895	0.325	Normal

Data yang sudah melewati proses perancangan data akan digunakan untuk implementasi metode KNN.

### 3.2 Implementasi Metode KNN

Proses implementasi metode KNN menggunakan data status gizi balita sebanyak 660 data yang sudah melewati tahap perancangan data. Adapun tahapan-tahapan dalam proses implementasi dengan metode KNN sebagai berikut:

1. Pemilihan nilai K

Klasifikasi status gizi balita menggunakan metode KNN dilakukan dengan menentukan nilai K terlebih dahulu. Adapun nilai K yang digunakan yaitu K=1, 3, 5, 7, 9, dan 11.

2. Penentuan rumus jarak

Penelitian ini menggunakan beberapa rumus perhitungan jarak yaitu *Euclidean Distance*, *Manhattan Distance*, dan *Minkowski Distance*.

3. Membagi dataset menjadi data *training* dan data *testing*

Pembagian data pada penelitian ini menggunakan metode *Stratified K-Fold Cross Validation* dengan *10-fold* sehingga setiap iterasi *fold* terdapat 66 data yang akan digunakan sebagai data *testing*, sedangkan sisanya akan digunakan sebagai data *trainig*.

4. Menghitung jarak antara data *training* dan data *testing*

Proses perhitungan jarak menggunakan beberapa rumus perhitungan jarak yang telah ditentukan. Adapun perhitungan jarak antara data *testing* pertama dan data *training pertama* menggunakan persamaan *Euclidean Distance* sebagai berikut:

$$D(x_1, y_1) = \sqrt{(0 - 1)^2 + (1 - 0.7017)^2 + (0.8947 - 0.7052)^2 + (0.9625 - 0.7975)^2}$$

$$D(x_1, y_1) = \sqrt{(-1)^2 + (0.2982)^2 + (0.1894)^2 + (0.165)^2}$$

$$D(x_1, y_1) = 1.07334$$

Selanjutnya menghitung jarak antara data *testing* pertama dan data *training* pertama menggunakan persamaan *Manhattan Distance* sebagai berikut:

$$D(x_1, y_1) = |0 - 1| + |1 - 0.7017| + |0.8947 - 0.7052| + |0.9625 - 0.7975|$$

$$D(x_1, y_1) = |-1| + |0.2982| + |0.1894| + |0.1650|$$

$$D(x_1, y_1) = 1.6527$$

Tahap terakhir menghitung jarak antara data *testing* pertama dan data *training* pertama menggunakan persamaan *Minkowski Distance* sebagai berikut:

$$D(x_1, y_1) = (|0 - 1|^3 + |1 - 0.7017|^3 + |0.8947 - 0.7052|^3 + |0.9625 - 0.7975|^3)^{\frac{1}{3}}$$

$$D(x_1, y_1) = (|-1|^3 + |0.2982|^3 + |0.1894|^3 + |0.165|^3)^{\frac{1}{3}}$$

$$D(x_1, y_1) = 1.0124$$

5. Mengurutkan hasil perhitungan jarak dan menentukan kategori klasifikasi berdasarkan nilai K

Hasil perhitungan jarak diurutkan dari jarak terkecil hingga terbesar. Hasil klasifikasi dengan persamaan *Euclidean Distance* dapat dilihat pada Tabel 3.

Tabel 3 Hasil Klasifikasi *Euclidean Distance*

Nama	Nilai Jarak	Ranking	Kelas Gizi
Balita 617	0.04435	1	Normal
Balita 336	0.04747	2	Normal

<b>Nama</b>	<b>Nilai Jarak</b>	<b>Ranking</b>	<b>Kelas Gizi</b>
Balita 397	0.081402	3	Normal
Balita 602	0.093156	4	Normal
Balita 662	0.107146	5	Normal
Balita 407	0.107609	6	Normal
Balita 655	0.112848	7	Normal
Balita 340	0.11835	8	Normal
Balita 713	0.123392	9	Normal
Balita 682	0.135132	10	Normal
Balita 732	0.137591	11	Normal

Hasil klasifikasi dengan persamaan *Manhattan Distance* dapat dilihat pada Tabel 4.

Tabel 4 Hasil Klasifikasi *Manhattan Distance*

<b>Nama</b>	<b>Nilai Jarak</b>	<b>Ranking</b>	<b>Kelas Gizi</b>
Balita 617	0.070614	1	Normal
Balita 336	0.071667	2	Normal
Balita 662	0.125263	3	Normal
Balita 407	0.133816	4	Normal
Balita 397	0.135746	5	Normal
Balita 602	0.154254	6	Normal
Balita 655	0.169167	7	Normal
Balita 340	0.170395	8	Normal
Balita 713	0.180921	9	Normal
Balita 692	0.185965	10	Normal
Balita 594	0.198465	11	Normal

Hasil klasifikasi dengan persamaan *Minkowski Distance* dapat dilihat pada Tabel 5.

Tabel 5 Hasil Klasifikasi *Minkowski Distance*

<b>Nama</b>	<b>Nilai Jarak</b>	<b>Ranking</b>	<b>Kelas Gizi</b>
Balita 617	0.039146	1	Normal
Balita 336	0.042137	2	Normal
Balita 397	0.070329	3	Normal
Balita 602	0.080506	4	Normal
Balita 655	0.104817	5	Normal
Balita 662	0.105503	6	Normal
Balita 407	0.105556	7	Normal
Balita 340	0.109532	8	Normal
Balita 713	0.112405	9	Normal
Balita 732	0.122611	10	Normal
Balita 682	0.12556	11	Normal

Berdasarkan hasil klasifikasi dengan persamaan *Euclidean Distance*, *Manhattan Distance*, dan *Minkowski Distance* dapat dilihat bahwa parameter  $K=1, 3, 5, 7, 9, 11$  menunjukkan hasil mayoritas kelas gizi normal yang sesuai dengan data aktual pada data *testing* pertama yaitu balita dengan kelas gizi normal maka model benar mengklasifikasikan balita tersebut.

### 3.2 Evaluasi Metode KNN

Hasil evaluasi disajikan dalam tabel performance metrics dan confusion matrix. Hasil performance metrics dapat dilihat pada Tabel 6.

Tabel 6 Hasil Performance Metrics

<b>Nilai K</b>	<b>Perhitungan</b>	<b>Akurasi</b>	<b>Presisi</b>	<b>Recall</b>	<b>F1-Score</b>
1	Euclidean	89.09%	88.77%	89.09%	88.64%
3	Distance	87.12%	62.08%	46.36%	51.63%

Nilai K	Perhitungan	Akurasi	Presisi	Recall	F1-Score
5		86.21%	56.36%	42.38%	46.42%
7		85%	48.94%	30.84%	33.38%
9		85%	29.04%	26.99%	27.44%
11		84.69%	28.60%	26.70%	27.10%
1	Manhattan Distance	88.33%	65.23%	51.88%	56.69%
3		87.12%	62.82%	47.43%	52.71%
5		86.06%	53.18%	36.18%	39.98%
7		85.30%	49.50%	29.63%	31.31%
9		84.54%	28.63%	26.24%	26.71%
11		83.48%	26.75%	24.91%	25.12%
<b>1</b>	<b>Minkowski Distance</b>	<b>89.54%</b>	<b>84.05%</b>	<b>58.02%</b>	<b>65.36%</b>
3		87.77%	64.62%	47.51%	53.21%
5		85.90%	63.07%	40.59%	45.60%
7		85.75%	50.16%	34.24%	38.08%
9		85%	48.94%	29.13%	30.76%
11		85%	29.35%	27.41%	27.87%

Berdasarkan Tabel 6 diketahui bahwa akurasi tertinggi yaitu sebesar 89.54% dengan presisi sebesar 84.05%, *recall* sebesar 58.02% , dan *F1-Score* sebesar 65.36% menggunakan nilai K=1 dengan perhitungan *Minkowski Distance*. Tabel *confusion matrix* menggunakan nilai K=1 dengan perhitungan *Minkowski Distance* dapat dilihat pada Tabel 7.

Tabel 7 Tabel *Confusion Matrix*

		<i>Predicted Class</i>				
		Beresiko Gizi Lebih	Gizi Kurang	Gizi Lebih	Normal	Obesitas
<i>Actual Class</i>	Beresiko Gizi Lebih	52	0	0	28	0
	Gizi Kurang	0	12	0	12	0
	Gizi Lebih	6	0	7	0	0
	Normal	15	5	0	519	0
	Obesitas	2	0	1	0	1

Selanjutnya menghitung metrik evaluasi berdasarkan *confusion matrix* nilai K=1 persamaan *Minkowski Distance*:

$$Akurasi = \left( \frac{52+12+7+519+1}{660} \right) = \frac{591}{660} \times 100\% = 89.54\%$$

$$Presisi = 0.0829 + 0.0242 + 0.0175 + 0.7584 + 0.0061$$

$$Presisi = 0.8405 \times 100\% = 84.05\%$$

$$Recall = 0.0818 + 0.0152 + 0.0121 + 0.4879 + 0.0015$$

$$Recall = 0.5802 \times 100\% = 58.02\%$$

$$F1 - Score = 0.0823 + 0.0186 + 0.0143 + 0.5674 + 0.0024$$

$$F1 - Score = 0.6536 \times 100\% = 65.36\%$$

Dari hasil perhitungan manual maka mendapatkan hasil *confusion matrix* pada K=1 menghasilkan akurasi sebesar 89.54%, presisi sebesar 84.05%, *recall* sebesar 58.02%, dan *F1-Score* sebesar 65.36%.

#### 4. KESIMPULAN

Berdasarkan hasil dan pembahasan mengenai penentuan status gizi balita, penelitian melakukan klasifikasi status gizi balita menggunakan metode *K-Nearest Neighbor* menggunakan data sebanyak 660 data, dan membagi data dengan metode *stratified k-fold cross validation*. Setelah membandingkan tingkat akurasi antara nilai K=1, 3, 5, 7, 9, dan 11, dengan rumus jarak

*Euclidean Distance*, *Manhattan Distance*, dan *Minkowski Distance*, diketahui bahwa tingkat akurasi tertinggi diperoleh dari nilai  $K=1$  dengan rumus jarak *Minkowski Distance* yaitu sebesar 89.54%.

## 5. SARAN

Berdasarkan penelitian yang telah dilakukan maka penulis dapat memberi saran untuk pengembangan selanjutnya yaitu menggunakan metode klasifikasi lain agar mendapatkan nilai akurasi yang lebih baik dalam kasus klasifikasi status gizi balita, seperti *Naive Bayes* dan SVM. Saran lainnya yaitu menggunakan nilai  $K$  dan rumus perhitungan jarak yang lebih variatif untuk mengetahui parameter terbaik dalam mengklasifikasikan status gizi balita.

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Fakultas Teknik, Universitas Mulawarman yang telah memberi dukungan terhadap penelitian ini.

## DAFTAR PUSTAKA

- [1] N. Azmi, H. Hazriani, Y. Yuyun, dan H. HS, “Klasifikasi status gizi balita menggunakan algoritma K-Nearest Neighbor (KNN),” *Prosiding SISFOTEK*, vol. 7, no. 1, hlm. 1–6, 2023. [Daring]. Tersedia: <http://www.seminar.iaii.or.id/index.php/SISFOTEK/article/view/396>
- [2] H. Saleh, M. Faisal, dan R. I. Musa, “Klasifikasi status gizi balita menggunakan metode K-Nearest Neighbor,” *Simtek: Jurnal Sistem Informasi dan Teknik Komputer*, vol. 4, no. 2, hlm. 120–126, 2019, doi: 10.51876/simtek.v4i2.60.
- [3] H. D. Herawati, H. K. Rahayu, R. K. Triastanti, dan R. Rusiyono, “Pencegahan malnutrisi pada anak prasekolah melalui pelatihan pengukuran status gizi pada guru PAUD,” *Media Karya Kesehatan*, vol. 6, no. 1, hlm. 157–168, 2023, doi: 10.24198/mkk.v6i1.40800.
- [4] Kementerian Kesehatan Republik Indonesia, *Peraturan Menteri Kesehatan Republik Indonesia tentang Standar Antropometri Penilaian Status Gizi Anak*, 2020.
- [5] Kementerian Kesehatan Republik Indonesia, *Hasil Survei Status Gizi Indonesia (SSGI) Tahun 2022*, Jakarta: Kemenkes RI, 2022. [Daring]. Tersedia: <https://ayosehat.kemkes.go.id>
- [6] S. N. Azizah dan Z. Fatah, “Implementasi metode K-Nearest Neighbor (K-NN) pada klasifikasi stunting balita,” *Gudang: Jurnal Multidisiplin Ilmu*, vol. 2, no. 10, hlm. 282–288, 2024, doi: 10.59435/gjmi.v2i10.1000.
- [7] F. Putra, H. F. Tahiyat, R. M. Ihsan, Rahmaddeni, dan L. Efrizoni, “Penerapan algoritma K-Nearest Neighbor menggunakan wrapper sebagai preprocessing untuk penentuan keterangan berat badan manusia,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, hlm. 273–281, 2024.
- [8] L. Suryadi, N. Ngajiyanto, N. E. Pratiwi, F. Ardhy, dan P. Riswanto, “Penerapan data mining prediksi penjualan mebel terlaris menggunakan metode K-Nearest Neighbor (K-NN) (studi kasus: Toko Zerita Meubel),” *JUSIM (Jurnal Sistem Informasi Musirawas)*, vol. 7, no. 2, hlm. 174–184, 2022, doi: 10.32767/jusim.v7i2.1697.

- [9] S. K. P. Loka dan A. Marsal, "Perbandingan algoritma K-Nearest Neighbor dan Naïve Bayes Classifier untuk klasifikasi status gizi pada balita," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 1, hlm. 8–14, 2023, doi: 10.57152/malcom.v3i1.474.
- [10] Muttaqin et al., *Pengenalan Data Mining*, Medan: Yayasan Kita Menulis, 2023.
- [11] A. Harmain, P. Paiman, H. Kurniawan, K. Kusriani, dan D. Maulina, "Normalisasi data untuk efisiensi K-Means pada pengelompokan wilayah berpotensi kebakaran hutan dan lahan berdasarkan sebaran titik panas," *Teknimedia: Teknologi Informasi dan Multimedia*, vol. 2, no. 2, hlm. 83–89, 2022, doi: 10.46764/teknimedia.v2i2.49.
- [12] M. K. Najib dan S. Nurdiani, "Judul artikel," *Jambura Geoscience Review*, vol. 3, no. 1, hlm. 9–17, 2021, doi: 10.34312/jgeosrev.v3i1.8259.
- [13] S. Dwidianti dan D. A. Anggoro, "Analisis perbandingan algoritma bisecting K-Means dan fuzzy C-Means pada data pengguna kartu kredit," *Emitor: Jurnal Teknik Elektro*, vol. 22, no. 2, hlm. 110–117, 2022, doi: 10.23917/emitor.v22i2.15677.
- [14] M. A. Damayanti, D. Siskasari, S. Hidayatullah, R. A. Fernandi, dan H. Santoso, "Pengembangan aplikasi mobile Saffco Skin untuk edukasi perawatan wajah dengan sistem rekomendasi menggunakan algoritma KNN," *Indonesian Journal of Data Science*, vol. 3, no. 1, hlm. 41–50, 2025, doi: 10.30989/ijds.v3i1.1585.
- [15] F. Shidiq, "Penerapan metode K-Nearest Neighbor (KNN) untuk menentukan ikan cupang dengan ekstraksi fitur ciri bentuk dan Canny," *Innovative Research in Informatics*, vol. 3, no. 2, hlm. 39–46, 2021, doi: 10.37058/innovatics.v3i2.3093.
- [16] A. Awaliyah, A. Prasetyadi, dan A. Junaidi, "Sistem rekomendasi desain website berdasarkan tingkat kemiripan menggunakan Euclidean Distance," *Jurnal Dinda: Data Science, Information Technology, and Data Analytics*, vol. 2, no. 2, hlm. 75–81, 2022, doi: 10.20895/dinda.v2i2.543.
- [17] G. H. Martono dan N. Sulistianingsih, "Perbandingan matriks jarak pada algoritma K-NN untuk prediksi penyakit diabetes," *JoMI: Journal of Millennial Informatics*, vol. 2, no. 1, hlm. 1–6, 2024.
- [18] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, dan A. Doulamis, "Multiclass confusion matrix reduction method and its application on Net Promoter Score classification problem," *Technologies*, vol. 9, no. 4, Art. no. 81, 2021, doi: 10.3390/technologies9040081.