

Penerapan Data Mining Untuk Klasifikasi Status Gizi Balita Dengan Algoritma *K-Nearest Neighbor*

Application of Data Mining for Classification of Toddler Nutritional Status Using the K-Nearest Neighbor Algorithm

Dwi Febrianti¹, Endang Etriyanti², Veradilla Amalia³

Sistem Informasi, STMIK Bina Nusantara Jaya Lubuklinggau,

Jl. Yos Sudarso No.97 A Kel. Jawa Kanan, Kota Lubuklinggau, Sumatera Selatan

E-mail : dwifebrianti080204@gmail.com³, endang.etriyanti@gmail.com², veradillaamalia@gmail.com³

Abstract

Advances in information technology have driven digital transformation in many sectors, including healthcare. Community health centers (Puskesmas), as primary healthcare institutions in Indonesia, play a crucial role in providing promotive and preventive services, particularly in monitoring the nutritional status of children under five. The L. Sidoharjo Community Health Center faces challenges in accurately determining the nutritional status of toddlers. Although an assessment system is available, field practice shows that assessment results often do not reflect the true condition. This is due to individual factors, such as genetics, which influence a toddler's body shape. A child who is genetically thin does not necessarily suffer from malnutrition, nor does an obese child necessarily suffer from overnutrition. This situation highlights the need for an assessment system that relies not only on visual observation but also on a more objective, data-driven approach. Therefore, it is necessary to develop a nutritional status classification system that can account for individual variations to produce more accurate and reliable evaluations to support optimal child nutrition monitoring.

Keywords: Data Mining, Classification, Toddler Nutritional Status, K-Nearest Neighbor, Community Health Center.

Abstrak

Kemajuan dalam teknologi informasi telah mendorong perubahan digital di banyak sektor, termasuk bidang kesehatan. Puskesmas sebagai lembaga kesehatan dasar di Indonesia memiliki peranan krusial dalam memberikan pelayanan promotif dan preventif, terutama dalam mengawasi kondisi gizi anak-anak di bawah lima tahun. Puskesmas L. Sidoharjo menghadapi tantangan dalam penentuan status gizi balita yang akurat. Meskipun sistem penilaian telah tersedia, praktik di lapangan menunjukkan bahwa hasil penilaian sering kali tidak mencerminkan kondisi sebenarnya. Hal ini disebabkan oleh faktor individual, seperti genetik, yang memengaruhi bentuk tubuh balita. Seorang anak yang secara genetik kurus belum tentu mengalami gizi buruk, begitu pula anak gemuk belum tentu mengalami kelebihan gizi. Kondisi ini menunjukkan perlunya sistem penilaian yang tidak hanya mengandalkan observasi visual, melainkan pendekatan yang lebih objektif dan berbasis data. Dengan demikian, diperlukan pengembangan sistem klasifikasi status gizi yang mampu memperhitungkan variasi individual guna menghasilkan evaluasi yang lebih akurat dan dapat diandalkan untuk mendukung pemantauan gizi anak secara optimal.

Kata kunci: Data Mining, Klasifikasi, Status gizi balita, K-Nearest Neighbor, Puskesmas

1. Pendahuluan

Kemajuan teknologi informasi di era digital saat ini telah mengakibatkan perubahan besar di berbagai bidang, termasuk di bidang kesehatan. Penerapan sistem komputerisasi menjadi sangat penting guna meningkatkan efisiensi dan akurasi dalam pengolahan data. Salah satu dampak nyata dari kemajuan ini adalah penerapan sistem digital di fasilitas pelayanan kesehatan

seperti rumah sakit, klinik, hingga pusat kesehatan masyarakat (Puskesmas).

Sesuai dengan Peraturan Menteri Kesehatan Republik Indonesia No. 43 Tahun 2019 mengenai Pusat Kesehatan Masyarakat, Puskesmas adalah fasilitas pelayanan kesehatan yang diselenggarakan oleh pemerintah mencakup berbagai jenis layanan, termasuk upaya promotif, preventif, kuratif, dan rehabilitatif[1].

Sebagai bagian vital dari sistem kesehatan nasional, puskesmas berada di bawah naungan Kementerian

Kesehatan dan diawasi langsung oleh pemerintah daerah. Dalam menjalankan tugasnya, puskesmas masih menghadapi berbagai tantangan, salah satunya adalah masalah gizi pada bayi. Dalam hal ini puskesmas L. Sidoharjo yang mengalami permasalahan terutama dalam hal pemantauan gizi balita. Meskipun terdapat sistem penilaian status gizi balita yang digunakan di puskesmas, dalam praktiknya, penentuan status gizi seringkali tidak mencerminkan kondisi sebenarnya dari balita. Hal ini disebabkan oleh faktor individu, seperti faktor genetik, yang bisa membuat seorang anak terlihat lebih kurus atau lebih gemuk dari rata-rata, padahal mereka tidak mengalami masalah gizi atau kelebihan gizi. Sebagai contoh, balita yang secara genetik kurus belum tentu menderita gangguan gizi, dan sebaliknya, balita yang terlihat gemuk juga belum tentu mengalami kelebihan gizi.

Situasi ini telah meningkatkan urgensi untuk menciptakan sistem klasifikasi status gizi yang tidak hanya bergantung pada persepsi visual atau penilaian manual, tetapi juga pada pendekatan komputasi yang objektif. Salah satu metode yang bisa digunakan dalam proses pengelompokan ini adalah sistem data mining dimana metode ini bekerja dengan melakukan pengambilan dan analisis data untuk menemukan pola-pola tersembunyi yang dapat mendukung pengambilan keputusan yang lebih baik. Salah satu algoritma yang efektif dan mudah diterapkan untuk tugas klasifikasi adalah algoritma *K-Nearest Neighbor* (*K-NN*).

Penerapan sistem berbasis data mining, khususnya algoritma *K-NN*, dalam pemantauan gizi balita di Puskesmas L. Sidoharjo memiliki manfaat yang signifikan, yaitu meningkatkan akurasi dan objektivitas penentuan status gizi, serta mengurangi kesalahan interpretasi akibat faktor subjektif atau perbedaan genetic dan juga dengan adanya sistem pemantauan gizi yang lebih baik, kesadaran masyarakat tentang pentingnya gizi seimbang pada balita juga dapat meningkat.

2. Tinjauan Pustaka

2.1 Analisa Penelitian Terdahulu

Pada penelitian sebelumnya yang telah dilakukan oleh Wahyudi et al. algoritma *K-NN* digunakan untuk mengklasifikasikan status gizi anak-anak di posyandu desa Bluto, yang membantu dalam mengidentifikasi status gizi balitadengan Tingkat akurasi yang baik[2]. Penerapan Algoritma *K-Nearest Neighbors* Pada Klasifikasi Penentuan Gizi Balita ini bertujuan untuk menghitung nilai akurasi dari algoritma *SVM* dalam mengklasifikasi status gizi balita pada kecamatan Bangun Purba dan untuk memberikan informasi serta pengetahuan yang dapat dimanfaatkan oleh berbagai pihak, salah satunya posyandu kecamatan Bangun Purba, Rokan Hulu, Riau[3]. Dari Penerapan Algoritma C.45 Pada Klasifikasi Status Gizi Balita di Posyandu Desa Sukalilah Cibatub Kabupaten Garut Jawa Barat

dapat disimpulkan bahwa algoritma C4.5 ini dapat diterapkan untuk mengklasifikasikan status gizi balita dengan kategori baik dan buruk. Penerapan algoritma pada Posyandu disini dengan cara mengelompokkan data status gizi balita[4]. Dalam penelitian Klasifikasi Status Gizi Balita Menggunakan Naïve Bayes dan K Nearest Neighbor Berbasis Web, menyatakan bahwa Klasifikasi Satus Gizi Balita dapat mempercepat menentukan keputusan untuk status gizi balita. Untuk perbandingan akuasi, recall, dan precision Naïve Bayes memperoleh nilai akurasi 80.60%, recall 80.60%, precision 79.66% sedangkan untuk KNN memperoleh nilai akurasi 91.79%, recall 91.79%, precision 91.17% [5].

2.2 Data Mining

Data mining merupakan aktivitas untuk menemukan pola dan hubungan dengan cara pengumpulan dan penggunaan data sebelumnya dalam jumlah data yang besar[5]. Data tersebut dapat disimpan dalam berbagai tempat, seperti *database*, *data warehouse*, atau penyimpanan informasi lainnya. Aktivitas ini melibatkan berbagai bidang ilmu pengetahuan, seperti sistem basis data, pengolahan data dalam gudang data (*data warehousing*), statistik, *machine learning*, serta pengambilan informasi, dan komputasi tingkat tinggi.

2.3 Klasifikasi

Klasifikasi adalah salah satu teknik dalam penambangan data yang bertujuan untuk mengelompokkan data ke dalam kategori atau kelas yang telah ditentukan sebelumnya[6]. klasifikasi merupakan salah satu metode dalam pengolahan data yang merupakan pengelompokan sekumpulan data sesuai dengan karakteristik atau ciri-ciri data itu sendiri satu sama lain[7]. Klasifikasi dilakukan dengan mengumpulkan sekumpulan data latih dengan beberapa kelas yang telah ditentukan.

2.4 Status Gizi

Status gizi adalah representasi dari apakah asupan nutrisi sejalan dengan kebutuhan gizi atau tidak. yang diperlukan untuk berbagai proses biologis dalam tubuh[6]. status gizi adalah kondisi fisik seseorang yang dapat diprediksi melalui asupan makanan dan zat-zat yang dicerna oleh tubuh, yang selanjutnya dapat memengaruhi tingkat kognitif individu[8]. Berdasarkan analisis penulis, status gizi merupakan keseimbangan antara asupan zat gizi dan kebutuhan, yang berfungsi dalam proses biologis tubuh serta memengaruhi kondisi fisik dan kognitif seseorang. Salah satu cara untuk menilai status gizi anak adalah dengan metode antropometri.

2.5 Antropometri

Antropometri merupakan teknik yang dipakai untuk mengukur ukuran, proporsi, dan komposisi fisik manusia. Standar Antropometri Anak harus dijadikan

pedoman oleh tenaga medis, pengelola program, serta semua pihak terkait dalam menilai kondisi gizi dan perkembangan anak. Di Indonesia, Standar Antropometri Anak merujuk pada Pedoman Pertumbuhan Anak WHO untuk anak usia 0-5 tahun. Standar tersebut memperlihatkan bagaimana pertumbuhan anak dapat dicapai apabila memenuhi syarat-syarat tertentu[9].

2.6 Balita

Balita didefinisikan sebagai anak yang berusia 12 bulan (1 tahun) hingga 59 bulan (5 tahun). Dapat dikatakan bahwa balita adalah anak-anak yang dibawah lima tahun, berasal dari istilah "balita" yang merupakan singkatan dari "bawah lima tahun". Pada usia tersebut, seorang balita mengalami pertumbuhan dan perkembangan yang sangat pesat. Proses ini ditandai oleh berbagai perubahan yang terjadi memerlukan asupan zat gizi dalam jumlah yang lebih banyak dan dengan kualitas yang lebih tinggi[1].

2.7 Algoritma K-Nearest Neighbor

K-Nearest Neighbor (K-NN) adalah teknik untuk mengklasifikasikan objek berdasarkan data pembelajaran yang paling dekat dengan objek tersebut. Data pembelajaran digambarkan dalam ruang multidimensi, di mana setiap dimensi mewakili suatu ciri atau fitur dari data[10]. Klasifikasi ditentukan oleh mayoritas label kelas dari tetangga terdekat. Kelas yang paling sering muncul di antara tetangga terdekat akan menjadi hasil klasifikasi. Berikut adalah langkah-langkah dalam menghitung algoritma K-NN:

1. Menentukan nilai K.
2. Menghitung jarak antara data pelatihan dan data pengujian dengan menggunakan rumus Jarak Euclidean.
3. Mengurutkan objek-objek dalam data pelatihan berdasarkan jarak terkecil.
4. Menetapkan kelas, di mana kelas yang dipilih adalah kelas yang memiliki jumlah suara terbanyak dari K tetangga terdekat pada data pengujian.

2.8 Jarak Euclidean Distance

Euclidean Distance adalah perhitungan yang dipakai untuk mencari jarak antara 2 titik dalam Euclidean space. Perhitungan metode ini berfungsi sebagai alat untuk mengukur dan menginterpretasikan kedekatan antara dua objek yang direpresentasikan[5]. kita dapat menggunakan rumus Euclidean Distance dengan persamaan sebagai berikut:

$$D = \sqrt{(x1 - y1)^2 + (x2 - y2)^2} \quad (i)$$

Keterangan :

D : Jarak Euclidean Distance

x1 : Data baru yang di cari

y1 : Data lama

x2 : Data baru yang di cari

y2 : Data lama

2.9 Teknik Slovin

Rumus Slovin digunakan untuk menentukan ukuran sampel minimal yang diperlukan dalam penelitian, terutama ketika jumlah populasi sangat besar[11]. Rumus ini membantu peneliti untuk mendapatkan sampel yang representatif dari populasi, sehingga hasil penelitian dapat digeneralisasikan. Berikut ini adalah rumus slovin :

$$n = \frac{N}{(1 + N) \cdot (e)^2} \quad (ii)$$

Keterangan :

n : ukuran sampel yang diperlukan

N = total populasi

e = margin of error (dalam desimal, misalnya 5% menjadi 0.05)

3. Metodologi Penelitian

3.1 Teknik Pengumpulan Data

Penulis melakukan pengumpulan data yang digunakan pada penelitian kali ini adalah :

1. Wawancara

Wawancara, merupakan teknik pengumpulan data yang melibatkan interaksi langsung antara peneliti dan partisipan penelitian yang bertujuan untuk mendapatkan pemahaman mendalam tentang pengalaman, pandangan, dan perspektif individu terkait fenomena yang diteliti. Jenis wawancara yang dipilih adalah wawancara terstruktur, di mana peneliti telah memiliki pemahaman yang mendalam mengenai informasi yang ingin didapatkan[12].

2. Observasi

Observasi merupakan teknik pengumpulan data yang melibatkan pengamatan langsung terhadap partisipan dan konteks yang terlibat dalam fenomena penelitian. Melalui observasi, peneliti memiliki kesempatan untuk melihat interaksi sosial, perilaku, dan konteks yang berkaitan dengan fenomena yang sedang dianalisis. Hasil dari pengamatan ini bisa mencakup sikap, tindakan, percakapan, dan interaksi interpersonal[12].

3. Studi Pustaka

Studi pustaka adalah suatu langkah untuk mengumpulkan data atau sumber-sumber yang berkaitan dengan topik yang dibahas dalam sebuah penulisan. Metode kajian pustaka mencakup serangkaian kegiatan yang berhubungan dengan pengumpulan informasi dari referensi tertulis, membaca, mencatat, serta

mengatur bahan yang akan diterapkan dalam penulisan.[12].

3.2 Metode Pengembangan Sistem

Dalam studi ini, penulis menggunakan pendekatan pengembangan sistem yang dinamakan waterfall. Metode waterfall merupakan suatu cara dalam pengembangan perangkat lunak yang memungkinkan proses pembuatan sistem dilakukan secara terencana dan berurutan berdasarkan siklus pengembangan yang telah ditentukan[13]. Berikut adalah tahapan-tahapan yang terdapat dalam model *waterfall* :

1) Requirement

Tahap ini memerlukan komunikasi antara pengembang dan pengguna untuk memahami perangkat lunak yang diinginkan serta batasan-batasan yang ada. Data bisa dikumpulkan lewat wawancara, diskusi, atau survei secara langsung. Data yang telah dikumpulkan kemudian dianalisis untuk mendapatkan informasi yang diperlukan oleh pengguna.

2) Design

Pada fase ini, para pengembang menciptakan rancangan sistem yang bertujuan untuk mengidentifikasi kebutuhan perangkat keras serta spesifikasi sistem, sekaligus membantu menetapkan struktur arsitektur sistem secara keseluruhan.

3) Implementation

Pada tahap ini, sistem awalnya dibuat di sebuah program kecil yang dikenal dengan istilah unit, yang kemudian diintegrasikan pada langkah berikutnya. Setiap unit dirancang dan diuji untuk memastikan fungsinya, yang dinamakan tes unit.

4) Verification

Pada fase ini, sistem akan melalui proses pengecekan dan pengujian untuk menentukan apakah sistem sepenuhnya atau sebagian besar memenuhi standar yang ditetapkan. Pengujian ini dapat dibagi menjadi beberapa kategori, yaitu pengujian unit (yang dilakukan pada bagian tertentu dari kode), pengujian sistem (untuk mengamati respons sistem ketika semua modul telah diintegrasikan), dan pengujian penerimaan (yang dilakukan bersama atau atas nama pelanggan untuk memastikan bahwa semua keinginan pelanggan terpenuhi).

5) Maintenance

Ini adalah tahap terakhir dari pendekatan waterfall. Software yang telah selesai dikembangkan dioperasikan dan akan dipelihara. Pemeliharaan ini mencakup perbaikan atas error yang tidak terdeteksi pada langkah-langkah sebelumnya.

3.3 Analisis

Saat ini, sistem yang diterapkan di puskesmas L. Sidoharjo untuk menilai status gizi anak balita sudah berbasis digital, namun dalam praktiknya, hasil penentuan status gizi sering kali tidak mencerminkan kondisi sebenarnya dari seorang balita. Hal ini

disebabkan oleh faktor individual, seperti kondisi genetik, yang membuat seorang anak tampak lebih kurus atau gemuk dari rata-rata, meskipun mereka sebenarnya tidak mengalami kekurangan gizi atau kelebihan gizi.

Untuk mengatasi masalah keterbatasan dalam penyampaian dan transparansi hasil penilaian status gizi balita di Puskesmas L. Sidoharjo, sistem yang diusulkan adalah sistem klasifikasi status gizi balita berbasis *data mining* menggunakan algoritma *K-Nearest Neighbor (K-NN)*. Sistem ini dirancang dengan tujuan utama untuk mengklasifikasikan status gizi balita yaitu :

- 1) Gizi buruk (*severely wasted*)
- 2) Gizi kurang (*wasted*)
- 3) Gizi baik (*normal*)
- 4) Berisiko gizi lebih (*possible risk of overweight*)
- 5) Gizi lebih (*overweight*)
- 6) Obesitas (*obese*)

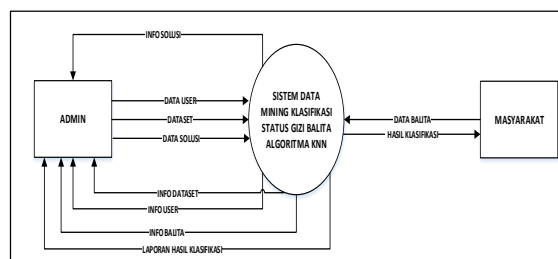
berdasarkan data antropometri dimana terdapat dua parameter utama yang menjadi acuan, yaitu :

- 1) Berat badan
- 2) Panjang/tinggi badan,

3.4 Model Perancangan

1. DFD Level 0

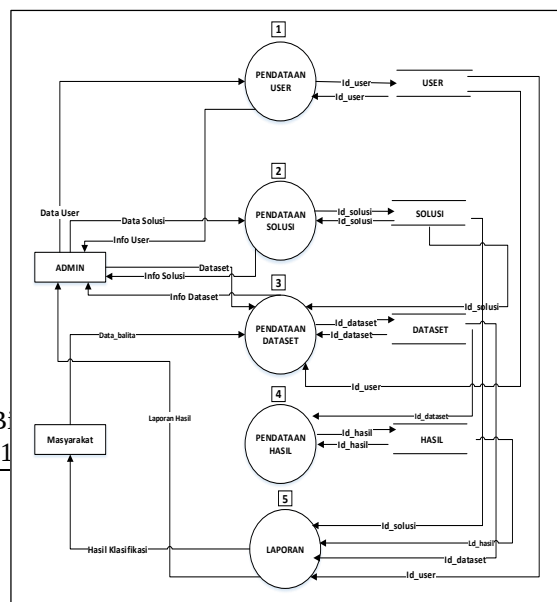
Dalam *Data Flow Diagram (DFD)* level 0, terdapat dua entitas yang berperan, yaitu admin dan masyarakat. Entitas admin bertanggung jawab untuk menginput data user, data solusi, dan dataset. Sementara itu, masyarakat hanya fokus pada penginputan data balita.



Gambar 1 DFD Level 0

2. DFD Level 1

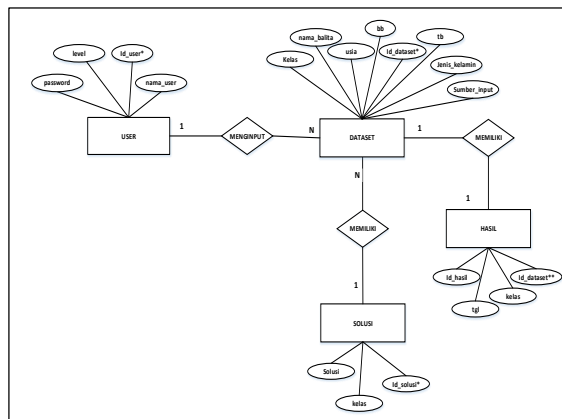
Data Flow Diagram (DFD) Level 1 digunakan untuk menggambarkan secara lebih rinci lagi yang ada dalam *Data Flow Diagram (DFD)* Level 0.



Gambar 2 DFD Level 1

3. ERD

Entity Relationship Diagram (ERD) menggambarkan suatu diagram yang menunjukkan suatu entitas dengan relasinya.



Gambar 3 ERD

4. Hasil dan Pembahasan

4.1 Proses Perhitungan Model K-NN

1. Persiapan Dataset

Pada penelitian ini, data yang digunakan merupakan hasil pengumpulan dari BLUD UPT Puskesmas L Sidoharjo yang berkaitan dengan informasi balita dan status gizinya. Jumlah total data yang berhasil dikumpulkan sebanyak 1.256 data ini diperoleh melalui metode pengumpulan data wawancara dan observasi.

2. Penentuan Dataset

Teknik pengambilan sampel dilakukan dengan metode *random sampling* selanjutnya Dari seluruh data populasi tersebut, sejumlah data akan dipilih secara acak menggunakan *tools* seperti *Excel (RANDBETWEEN)*, Hasil dari pemilihan acak ini akan digunakan sebagai dataset untuk klasifikasi. Setelah mengetahui jumlah populasi seluruh data sudah diacak, teknik *slovin* digunakan untuk menentukan ukuran sampel minimum berikut adalah hasilnya.

Rumus *Slovin*:

$$n = \frac{1245}{(1+1245) \cdot (0,05)^2} = 302.75 \text{ (303)}$$

Keterangan :

n = jumlah sampel

1256 = jumlah populasi

Clasifiation error = 5% atau 0.05

Tabel 1 Dataset

No	Nama	Usia	TB	BB	jk	Kelas
1	TALITA SABRINA	2 Tahun - 8 Bulan - 3 Hari	10	15, 2	P	Gizi Baik
2	SHABIRA UZELIN PUTRI	2 Tahun - 2 Bulan - 8 Hari	91	13, 3	L	Gizi Baik
3	ANANDA PUTRI	1 Tahun - 5 Bulan - 5 Hari	77	10, 2	P	Gizi Baik
4	KHALISA IMEL	1 Tahun - 8 Bulan - 8 Hari	87, 3	11, 6	P	Gizi Baik
5	SALISATUN HUMAIROH	4 Tahun - 6 Bulan - 1 Hari	94, 6	13, 1	L	Gizi Baik
6	ASRAF RIFATUL	4 Tahun - 9 Bulan - 13 Hari	10, 8	17, 5	P	Gizi Baik
7	ANIN MUHAMAD ARYA DZAKY PRADANA	2 Tahun - 4 Bulan - 4 Hari	88	12, 2	L	Gizi Baik
8	HANUM	2 Tahun - 3 Bulan - 6 Hari	83, 5	10, 4	P	Gizi Baik
9	MEIRA CITRA	3 Tahun - 5 Bulan - 27 Hari	98	13, 3	P	Gizi Baik
10	ALSIAZARA H HAFLA	4 Tahun - 11 Bulan - 20 Hari	10, 2,7	14, 3	P	Gizi Baik
...	...	...	...	...	...	...
30	ZIAN DEVANO	2 Tahun - 3 Bulan - 25 Hari	92, 5	14, 5	L	Gizi Baik
1	ARKANANT A	2 Tahun - 2 Bulan - 15 Hari	86	10, 8	P	Gizi Baik
30	RAINA ZAKIYA SAIRA	3 Tahun - 11 Bulan - 7 Hari	98, 3	13, 2	L	Gizi Baik
3	AZZA REZILIA					

3. Proses Klasifikasi

1) Input Data Uji

Data ini belum diketahui status gizinya, dan bertujuan untuk dianalisis menggunakan algoritma *K-Nearest Neighbor (KNN)* untuk menentukan status gizi berdasarkan data pembandingan dataset.

GALINA	1 Tahun -				
AIRIN	2 Bulan -	75.9	8.9	P	
ZIVANA	12 Hari				

2) Hitung Jarak Ke Semua Data (Menggunakan Rumus *Euclidean*)

$$\begin{aligned}
 D1 &= \sqrt{(75.9 - 102)^2 + (8.9 - 15.8)^2} \\
 &= \sqrt{(-26.1)^2 + (-6.9)^2} \\
 &= \sqrt{680.67 + 47.61} \\
 &= \sqrt{728.28} \\
 &= 26.98 \\
 D4 &= \sqrt{(75.9 - 87.3)^2 + (8.9 - 11.6)^2} \\
 &= \sqrt{(-11.4)^2 + (-2.7)^2} \\
 &= \sqrt{129.96 + 7.29} \\
 &= \sqrt{137.25} \\
 &= 11.71 \\
 D2 &= \sqrt{(75.9 - 91)^2 + (8.9 - 13.3)^2} \\
 &= \sqrt{(-15.1)^2 + (-4.4)^2} \\
 &= \sqrt{228.01 + 19.36} \\
 &= \sqrt{247.37} \\
 &= 15.73 \\
 D5 &= \sqrt{(75.9 - 94.6)^2 + (8.9 - 13.1)^2} \\
 &= \sqrt{(-18.7)^2 + (-4.2)^2} \\
 &= \sqrt{349.69 + 17.64} \\
 &= \sqrt{367.33} \\
 &= 19.17 \\
 D3 &= \sqrt{(75.9 - 77)^2 + (8.9 - 10.2)^2} \\
 &= \sqrt{(-1.1)^2 + (-1.3)^2} \\
 &= \sqrt{1.21 + 1.69} \\
 &= \sqrt{2.90} \\
 &= 1.70 \\
 D6 &= \sqrt{(75.9 - 108)^2 + (8.9 - 17.5)^2} \\
 &= \sqrt{(-32.1)^2 + (-8.6)^2} \\
 &= \sqrt{1030.41 + 73.96} \\
 &= \sqrt{1104.37} \\
 &= 33.23 \\
 D302 &= \sqrt{(75.9 - 86)^2 + (8.9 - 10.8)^2} \\
 &= \sqrt{(-10.1)^2 + (-1.9)^2} \\
 &= \sqrt{102.01 + 3.61} \\
 &= \sqrt{105.62} \\
 &= 10.28 \\
 D303 &= \sqrt{(75.9 - 98.3)^2 + (8.9 - 13.2)^2} \\
 &= \sqrt{(-22.4)^2 + (-4.3)^2} \\
 &= \sqrt{501.76 + 18.49} \\
 &= \sqrt{520.25} \\
 &= 22.81
 \end{aligned}$$

Tabel 2 Hasil Perhitungan Jarak

No	TB	BB	JK	Kelas	Jarak
1	102	15.8	P	Gizi Baik	26.98
2	91	13.3	L	Gizi Baik	15.73
3	77	10.2	P	Gizi Baik	1.70
4	87.3	11.6	P	Gizi Baik	11.71
5	94.6	13.1	L	Gizi Baik	19.17
6	108	17.5	P	Gizi Baik	33.23
...	...	...	...	...	...
302	86	10.8	P	Gizi Baik	10.28
303	98.3	13.2	L	Gizi Baik	22.81

3) Ambil Nilai K Tetangga Terdekat

Apabila sudah diketahui jarak masing-masing dataset dengan data uji maka selanjutnya dilakukan proses pengurutan dari nilai jarak terkecil ke nilai jarak yang terbesar lalu ambil data sebanyak K yang di tentukan dari nilai jarak terkecil, dalam penelitian ini penulis memilih nilai K=3.

Tabel 3 Hasil Rangking Jarak K=3

No	TB	BB	JK	Kelas	Jarak
1	76	8.9	L	Gizi Baik	0.1
2	76	9.1	P	Gizi Baik	0.22
3	76	9.3	P	Gizi Baik	0.41

4) Voting atau Mayoritas Kelas

Pada langkah ini dilakukan pengambilan keputusan untuk menentukan status gizi balita baru. Dengan Menentukan kelas (status gizi) yang paling banyak muncul (mayoritas) dari K=3 data tetangga terdekat yang telah dipilih pada langkah sebelumnya.

5) Hasil Klasifikasi

Berdasarkan proses klasifikasi status gizi menggunakan algoritma *K-Nearest Neighbor* (*K-NN*) dengan nilai K=3, dilakukan perhitungan manual terhadap data balita atas nama Galina Airin Zivana yang memiliki berat badan 8,9 kg dan tinggi badan 75,9 cm. Hasil klasifikasi secara manual menunjukkan bahwa tiga data tetangga terdekat yang memiliki jarak terkecil dengan data uji tersebut didominasi oleh kelas "Gizi Baik". Oleh karena itu, balita atas nama Galina Airin Zivana termasuk dalam kategori "Gizi Baik".

4.2 Perbandingan Hasil Sistem dengan Model K-NN

Dalam penelitian ini dari 20 data uji yang diproses, dilakukan perbandingan antara hasil klasifikasi status gizi balita menggunakan sistem dan metode *K-Nearest Neighbor* (*KNN*) yang berbasis data historis. Tujuannya adalah untuk mengevaluasi konsistensi dan akurasi metode klasifikasi berbasis data mining (*KNN*) terhadap

metode perhitungan manual algoritma *K-NN*. Berikut ini adalah tabel hasil perbandingan hasil sistem dengan model *K-NN* :

Tabel 4 Perbandingan Hasil

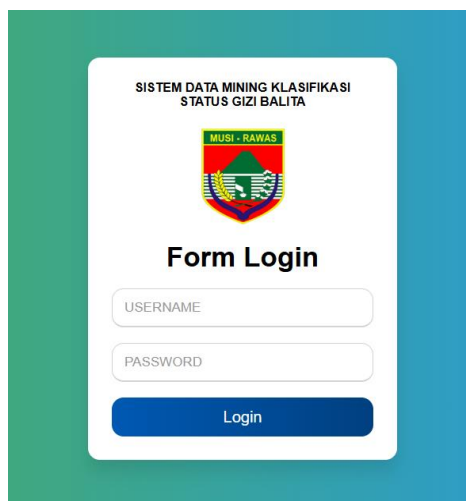
No	Nama	Jenis Kelamin	Tinggi Badan	Berat Badan	Hasil Klasifikasi Manual	Hasil Klasifikasi Sistem
1.	Galina Airin Zivana	P	75.9	8.9	Gizi Baik	Gizi Baik
2.	KYRA FRISTA	P	103	14.1	Gizi Baik	Gizi Baik
3.	ZEANISA ARFANA	P	90	16	Beresiko Gizi Lebih	Beresiko Gizi Lebih
4.	ZEEY ZYAN SAPUTRO	L	96	13.1	Gizi Baik	Gizi Baik
5.	UWAIS ALQORNI	P	98	17.7	Beresiko Gizi Lebih	Beresiko Gizi Lebih
6.	ELFATHAN DENANDRA D	L	89.1	10.9	Gizi Baik	Gizi Baik
7.	EVANO RIZKI	L	86.5	10.9	Gizi Baik	Gizi Baik
8.	ALINGGA ALZENA	P	91.7	17.9	Obesitas	Obesitas
9.	YUMNA H ALEA	P	80.1	9.5	Gizi Baik	Gizi Baik
10.	GISKA FAYOLA	P	80.1	9.5	Gizi Baik	Gizi Baik
11.	CILA KHAIRUNIS A	P	102.8	13.6	Gizi Baik	Gizi Baik
12.	AGUNG PUTRA P	L	107	22.7	Beresiko Gizi Lebih	Beresiko Gizi Lebih
13.	NEIRA FHATIMA	P	95.5	14	Gizi Baik	Gizi Baik
14.	DAYANA SHANUM	P	107	24.4	Beresiko Gizi Lebih	Beresiko Gizi Lebih
15.	ANWAR M AZAM	L	69.2	6.8	Gizi Baik	Gizi Baik
16.	RIYANSAH PUTRI	P	68.5	6.5	Gizi Kurang	Gizi Kurang
17.	KEYSA HUMAIRA	L	98	15.9	Gizi Baik	Gizi Baik
18.	RAIZAKI ADITAMA	P	100	14	Gizi Baik	Gizi Baik
19.	ANINDIYA FERLINDI	P	103	13.7	Gizi Baik	Gizi Baik
20.	SHAQUEEN A ADZANA	L	90	13	Gizi Baik	Gizi Baik
21.	M RAYAN ARSENIO ZAFEER A	L	83.3	14.5	Beresiko Gizi Lebih	Beresiko Gizi Lebih

4.3 Implementasi Sistem

Sistem data Mining Klasifikasi Status Gizi Balita ini dibuat berdasarkan rancangan sistem yang sudah dilakukan sebelumnya. Adapun hasil sistem data mining klasifikasi status gizi balita ini bisa dilihat pada tampilan halaman sebagai berikut :

1. Halaman Login Admin

Form login admin digunakan oleh admin untuk bisa mengakses sistem. Untuk dapat *login*, admin diwajibkan untuk memasukkan *username* dan *password* yang sudah terdaftar.



Gambar 4 Halaman Login Admin

2. Halaman Input Data Balita

Halaman input data balita adalah halaman yang dapat diakses oleh masyarakat untuk memasukkan data balita yang juga sebagai data uji untuk diklasifikasikan status gizinya. Saat data sudah diisi seluruhnya lalu di klik tombol cek maka hasil akan otomatis keluar.

Gambar 5 Halaman Input Data Balita

3. Halaman Hasil Klasifikasi

Halaman hasil klasifikasi adalah halaman yang menampilkan hasil klasifikasi dari data yang diinputkan masyarakat ke dalam tabel balita pada halaman sebelumnya.

No	ID Dataset	BB (kg)	TB (cm)	Kelas	Jarak
1	148	8.9	76	Gizi Baik	0.1000
2	135	9.1	76	Gizi Baik	0.2236
3	240	9.3	76	Gizi Baik	0.4103

Gambar 6 Halaman Hasil Klasifikasi

4. Halaman Cetak Hasil Klasifikasi

Halaman cetak hasil klasifikasi adalah halaman yang menampilkan hasil klasifikasi yang akan dicetak.



Gambar 7 Cetak Hasil Klasifikasi

5. Kesimpulan dan Saran

Berdasarkan uraian yang telah dijelaskan oleh penulis di bab-bab sebelumnya yaitu pada sistem *data mining* klasifikasi status gizi balita algoritma *K-NN* studi kasus BLUD UPT Puskesmas L.Sidoharjo. maka dapat ditarik kesimpulan sebagai berikut :

5.1 Simpulan

Adapun kesimpulan dari hasil penelitian ini, adalah sebagai berikut :

1. Sistem klasifikasi status gizi balita yang dibangun menggunakan *algoritma K-Nearest Neighbor (K-NN)* mampu menilai status gizi balita tidak hanya berdasarkan standar umum, tetapi juga mempertimbangkan kemiripan karakteristik dengan data lain yang telah diklasifikasikan sebelumnya. Sehingga sistem dapat memberikan hasil klasifikasi yang lebih objektif dan konsisten dari perbedaan genetik yang bersifat individual.
2. Penelitian ini berhasil merancang dan membangun sistem klasifikasi status gizi balita berbasis *data mining* dengan menggunakan algoritma *K-Nearest Neighbor (K-NN)*.

5.2 Saran

Adapun beberapa saran yang dapat disampaikan berdasarkan hasil penelitian ini, antara lain:

1. Untuk pengembangan sistem selanjutnya sebaiknya menggunakan data yang lebih lengkap dan bervariasi dari periode waktu yang berbeda dapat meningkatkan akurasi model dan memperkuat generalisasi sistem.
2. Perlu dilakukan pengembangan dalam bentuk aplikasi web responsif atau mobile akan sangat membantu dalam meningkatkan akses dan pemanfaatan sistem, khususnya bagi masyarakat dan tenaga kesehatan di lapangan.

3. Penggunaan algoritma K-NN dapat dibandingkan dengan algoritma klasifikasi lain untuk mengetahui apakah terdapat metode lain yang dapat menghasilkan akurasi lebih tinggi.

Daftar Rujukan

- [1] Kemenkes “PERATURAN MENTERI KESEHATAN REPUBLIK INDONESIA NOMOR 41 TAHUN 2014 TENTANG PEDOMAN GIZI SEIMBANG.pdf.” pp. 1–96, 2014.
- [2] R. Wahyudi, M. Orisa, and N. Vendyansyah, “Penerapan Algoritma K-Nearest Neighbors Pada Klasifikasi Penentuan Gizi Balita (Studi Kasus Di Posyandu Desa Bluto),” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 5, no. 2, pp. 750–757, 2021, doi: 10.36040/jati.v5i2.3738.
- [3] E. Ramon, A. Nazir, N. Novriyanto, Y. Yusra, and L. Oktavia, “Classification of Infant Nutritional Status at Bangun Purba Subdistrict Integrated Health Post Using Support Vector Machine (SVM) Algorithm,” *J. Sist. Inf. dan Inform.,* vol. 5, no. 2, pp. 143–150, 2022.
- [4] A. S. Millah, Apriyani, D. Arobiah, E. S. Febriani, and E. Ramdhani, “Analisis Data dalam Penelitian Tindakan Kelas,” *J. Kreat. Mhs.,* vol. 1, no. 2, pp. 140–153, 2023.
- [5] R. Setiawan and A. Triayudi, “Klasifikasi Status Gizi Balita Menggunakan Naïve Bayes dan K-Nearest Neighbor Berbasis Web,” *J. Media Inform. Budidarma,* vol. 6, no. 2, p. 777, 2022, doi: 10.30865/mib.v6i2.3566.
- [6] S. Larissa, Z. Lewoema, and P. T. Prasetyaningrum, “Implementasi Data Mining Pada Klasifikasi Status Gizi Bayi Dengan Metode Decision Tree CHAID (Studi Kasus : Puskesmas Godean 1 Yogyakarta),” vol. 5, no. 1, pp. 61–74, 2024, doi: 10.51519/journalita.v5i1.538.
- [7] A. S. Biyantoro and B. Prasetyo, “Penerapan Decision Tree untuk Klasifikasi Status Kesehatan dengan Perbandingan KNN dan Naive Bayes,” *Indones. J. Inform. Res. Softw. Eng.,* vol. 4, no. 1, pp. 47–55, 2024.
- [8] S. K. P. Loka and A. Marsal, “Perbandingan Algoritma K-Nearest Neighbor dan Naïve Bayes Classifier untuk Klasifikasi Status Gizi Pada Balita,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.,* vol. 3, no. 1, pp. 8–14, 2023, doi: 10.57152/malcom.v3i1.474.
- [9] Kemenkes, “PERATURAN MENTERI KESEHATAN REPUBLIK INDONESIA NOMOR 43 TAHUN 2019 TENTANG PUSAT KESEHATAN MASYARAKAT.pdf.” pp.1-96, 2019.
- [10] F. Sri, M. Latowa, A. Lahinta, M. Ramdhan, and A. Kaluku, “Sistem Informasi Monitoring Pajak Bumi Dan Bangunan Berbasis Web,” vol. 2, no. 2, 2020, doi: 10.37905/jji.v2i2.7150.
- [11] A. Santoso, “Rumus Slovin : Panacea Masalah Ukuran Sampel?,” *Suksma J. Psikol. Univ. Sanata Dharma,* vol. 4, no. 2, pp. 24–43, 2023, doi: 10.24071/suksma.v4i2.6434.
- [12] Ardiansyah et al., “Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif,” *J. IHSAN J. Pendidik. Islam,* vol. 1, no. 2, pp. 1–9, 2023, doi: 10.61104/ihsan.v1i2.57.
- [13] M. Badrul, R. Ardy, S. Nusa Mandiri Jl Jatiwaringin Raya No, and K. Cipinang Melayu Jakarta Timur, “Penerapan Metode Waterfall pada Perancangan Sistem Informasi Pendaftaran Siswa Baru,” *J. Sains Komput. Inform. (J-SAKTI),* vol. 5, no. 1, pp. 52–61, 2021.