

Penerapan Algoritma C4.5 Pada Klasifikasi Status Gizi Balita

Yusuf Ramadhan Nasution¹, Armansyah², Mhd. Furqan³, Toibatur Rahma Matondang⁴

^{1,2,3,4}Ilmu Komputer, Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara

¹ramadhannst@uinsu.ac.id, ²armansyah@uinsu.ac.id, ³mfurqan@uinsu.ac.id, ⁴toibaturrahmamtd@gmail.com *

Abstract

The study aims to classify the nutritional status of the child using the C4.5 algorithm. The secondary data used is derived from the assessment of the nutrition status of a child in Puskesmas Promji and Pukesmas Suka Makmur. A classification model is constructed using the C4.5 algorithm based on a number of predictor factors that have been determined. The research methodology includes data collection, data preprocessing, model development with C4.5 algorithms, model evaluation, and results analysis. Model evaluation is done using measurements such as accuracy. In addition, the significance of predictor variables in affecting the nutritional status of infants was also evaluated through data analysis. This research contributed to the development of a method of classifying the nutritional status of infants using the C4.5 algorithm approach. The implication of this study is that the classification model developed can be used as a tool to support early identification and intervention against nutritional problems in infants. Furthermore, based on testing using the confusion matrix technique with the 80:20 data division of a total of 502 datasets, consisting of 402 training data and 100 testing data, an accuracy rate of 80 percent was obtained.

Keywords: Machine Learning, C4.5 Algorithm, Nutritional Status of Toddlers, Classification, RapidMiner

Abstrak

Penelitian ini bertujuan untuk mengklasifikasikan status gizi balita dengan menggunakan algoritma C4.5. Data sekunder yang digunakan berasal dari penilaian status gizi balita di Puskesmas Janji dan Puskesmas Suka Makmur. Sebuah model klasifikasi dibangun dengan menggunakan algoritma C4.5 berdasarkan sejumlah faktor prediktor yang telah ditentukan. Metodologi penelitian mencakup pengumpulan data, prapemrosesan data, pengembangan model dengan algoritma C4.5, evaluasi model, dan analisis hasil. Temuan penelitian menunjukkan bahwa model kategorisasi menggunakan algoritma C4.5 mampu memprediksi status gizi balita dengan akurat. Evaluasi model dilakukan menggunakan ukuran-ukuran seperti akurasi. Selain itu, signifikansi variabel prediktor dalam mempengaruhi status gizi balita juga dievaluasi melalui analisis data. Penelitian ini memberikan kontribusi dalam pengembangan metode klasifikasi status gizi balita dengan pendekatan algoritma C4.5. Implikasi dari penelitian ini adalah bahwa model klasifikasi yang dikembangkan dapat digunakan sebagai alat untuk mendukung identifikasi dan intervensi dini terhadap masalah gizi pada balita. Selanjutnya, berdasarkan pengujian menggunakan teknik confusion matrix dengan pembagian data 80:20 dari total 502 dataset, terdiri dari 402 data training dan 100 data testing, diperoleh tingkat akurasi sebesar 80%.

Kata kunci: Machine Learning, Algoritma C4.5, Status Gizi Balita, Klasifikasi, RapidMiner.

©This work is licensed under a Creative Commons Attribution - ShareAlike 4.0 International License

1. Pendahuluan

Kesehatan adalah aspek yang vital dalam kehidupan manusia, dan telah dilakukan berbagai penelitian dan studi dalam ranah ini, salah satunya adalah mengenai gizi. Pentingnya peran gizi dalam kehidupan manusia tidak bisa diabaikan, karena asupan nutrisi yang cukup berperan dalam membentuk masyarakat yang sehat dan bermutu. Di Indonesia, masalah gizi masih menjadi isu nasional yang signifikan, terutama di kalangan balita. Istilah "balita" mengacu pada anak-anak berusia 1-3 tahun (batita) dan anak-anak prasekolah (3-5 tahun). [1]. Pada rentang usia 0-60 bulan, fase awal kehidupan memiliki dampak yang signifikan terhadap tahap perkembangan selanjutnya bagi setiap individu. Masalah gizi pada anak-anak dalam periode ini dapat menghambat pertumbuhan dan perkembangan, bahkan

berpotensi menyebabkan penyakit serius atau kematian [2]. Selain itu, untuk menilai kondisi kesehatan seseorang, kita dapat mengamati status gizinya. Status gizi merupakan indikator keberhasilan dalam memenuhi kebutuhan nutrisi anak, yang diukur melalui berat badan dan tinggi badan. Status gizi juga mencerminkan kesehatan yang tercapai dari keseimbangan antara asupan nutrisi dan kebutuhan tubuh.

Klasifikasi status gizi pada balita dapat dibagi menjadi empat kategori, yaitu kelebihan gizi (overweight), gizi cukup, kekurangan gizi (underweight), dan gizi buruk [3].

Berdasarkan informasi yang terdokumentasi dalam Buku Saku Hasil Survei Status Gizi Indonesia (SSGI) 2022, disampaikan bahwa status gizi balita mencakup

kondisi stunting (gangguan pertumbuhan dan perkembangan), wasting (berat badan rendah dibandingkan tinggi badan), underweight (berat badan terlalu rendah), serta overweight (berat badan berlebihan). Data ini diperoleh dari penelitian yang melibatkan 334.848 sampel bayi dan balita dari 486 kabupaten/kota di 33 provinsi Indonesia. Perbandingan antara tahun 2021 dan 2022 menunjukkan penurunan angka stunting sebesar 2,8%, kenaikan wasting sebesar 0,6%, peningkatan underweight sebesar 17,1%, dan penurunan overweight sebesar 0,3%. Secara spesifik di Kabupaten Labuhanbatu, data SSGI 2022 menggambarkan bahwa prevalensi stunting pada balita mencapai 23,9%, sementara overweight mencapai 1,3% [4].

Dalam menghadapi masalah status gizi balita, pemerintah Kabupaten Labuhanbatu berkomitmen untuk terus mengurangi angka gizi buruk pada balita. Salah satu kunci untuk mencapai masyarakat yang sehat adalah melalui pelayanan kesehatan yang memadai. Puskesmas di Kecamatan Bilah Barat, seperti Puskesmas Janji dan Puskesmas Suka Makmur, menjadi titik fokus dalam penyediaan layanan kesehatan. Mereka memiliki tenaga ahli, termasuk ahli gizi, yang siap memberikan konsultasi dan perawatan gizi kepada masyarakat. Namun, meskipun layanan ini tersedia, masih ditemukan kasus gizi buruk pada balita di Kecamatan Bilah Barat, yang mungkin disebabkan oleh kurangnya pemahaman masyarakat tentang gizi balita dan kesalahan dalam penanganannya. Penilaian status gizi balita dapat dilakukan melalui pengukuran tubuh manusia yang dikenal sebagai "Anthropometri". Ini mencakup parameter seperti Umur (U), Berat Badan (BB), Tinggi Badan (TB), dan Lingkar Kepala (LK) [5].

Di Puskesmas Janji dan Puskesmas Suka Makmur, pencatatan dan penilaian status gizi balita masih dilakukan secara manual melalui pencatatan dalam buku-buku. Ada sejumlah kriteria dan pertimbangan yang harus dipertimbangkan untuk mengklasifikasikan status gizi balita dengan cepat dan efisien agar proses pendataan bisa berjalan lancar. Namun, karena pendekatan manual, proses ini seringkali memakan waktu yang cukup lama. Oleh karena itu, peneliti memilih untuk melakukan studi kasus di Puskesmas yang berada di Kecamatan Bilah Barat, termasuk Puskesmas Janji dan Puskesmas Suka Makmur, untuk mengevaluasi kebutuhan akan sistem yang dapat membantu dalam proses penilaian status gizi balita secara lebih efisien dan cepat. Pihak Puskesmas di Kecamatan Bilah Barat melakukan perhitungan status gizi balita secara manual, yang menghabiskan waktu dan berpotensi menyebabkan kesalahan pencatatan data, yang pada gilirannya membuat proses tidak efisien dan mungkin menghasilkan informasi yang tidak akurat. Karena pentingnya untuk menentukan status gizi balita dengan cepat dan akurat, diperlukan adanya sistem yang dapat membantu dalam mengklasifikasikan status gizi balita, sehingga

mempermudah penanganan jika ditemukan kasus gizi buruk pada balita.

Teknik klasifikasi digunakan untuk memisahkan atau mengategorikan data ke dalam kelompok atau kelas tertentu berdasarkan kriteria atau label yang telah ditentukan [6]. Penelitian yang dilakukan oleh [7] Model klasifikasi yang telah dikembangkan dengan menggunakan algoritma C4.5 dapat membagi status gizi balita menjadi enam kategori berdasarkan indeks BB/TB. Kategori-kategori tersebut mencakup gizi normal, gizi kurang, gizi buruk, risiko kelebihan gizi, kelebihan gizi, dan obesitas. Tingkat akurasi model mencapai 90%, sementara nilai precision, recall, dan f1-score adalah 87%..

Berdasarkan masalah yang dihadapi dan penelitian sebelumnya, studi ini akan difokuskan pada klasifikasi status gizi balita di Puskesmas Janji dan Puskesmas Suka Makmur di Kecamatan Bilah Barat. Pendekatan yang akan digunakan adalah Algoritma C4.5. Pemilihan algoritma ini didasarkan pada keunggulannya dalam menghasilkan model berupa struktur pohon atau aturan yang mudah dipahami, tingkat akurasi yang dapat diterima, serta kemampuannya dalam menangani variabel diskrit dan numerik. Studi ini bertujuan untuk menyelidiki klasifikasi status gizi balita menggunakan Algoritma C4.5. Dengan menerapkan Algoritma C4.5, diharapkan dapat membantu dalam mengklasifikasikan status gizi balita, sehingga memberikan informasi yang berharga bagi pemerintah setempat untuk mendukung kebijakan program-program puskesmas. Peneliti akan melakukan penelitian skripsi yang berjudul "**Penerapan Algoritma C4.5 Pada Klasifikasi Status Gizi Balita**".

2. Metode Penelitian

2.1. Prosedur Kerja

penelitian ini bertujuan untuk mengembangkan suatu aplikasi berbasis android yang bisa dimanfaatkan untuk melakukan klasifikasi status gizi balita menggunakan Algoritma C4.5. Metode penelitian yang digunakan dalam penelitian ini adalah metode penelitian dan pengembangan (R&D). Prosedur kerja penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Kerja Penelitian

2.2. Identifikasi Masalah

Tahapan ini mengacu kepada permasalahan umum yang berasal dari fenomena yang terjadi khususnya pada Puskesmas yang ada di Kecamatan Bilah Barat yaitu Puskesmas Janji dan Puskesmas Suka Makmur . Adapun fenomena yang diamati selama ini adalah, adanya ditemukan balita yang mengalami ketidakseimbangan gizi ataupun gizi buruk dan lemahnya sistem perhitungan dalam menentukan status gizi balita.

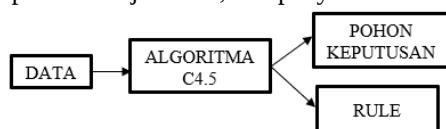
2.3. Pengumpulan Data

Teknik pengumpulan data yang pertama dilakukan dalam penelitian ini yaitu studi pustaka referensi dan dasar teori pada penelitian ini diambil dari berbagai jurnal, ebook dan buku, kemudian observasi pada Puskesmas di Kecamatan Bilah Barat yaitu Puskesmas Janji dan Puskesmas Suka Makmur untuk mengamati langsung dan mencatat secara sistematis terhadap balita yang dicek status gizi nya yang menjadi sasaran pengamatan. Selanjutnya pengumpulan data dengan wawancara dengan narasumber (Kepala Puskesmas, bidan atau staff pengurus status gizi balita) untuk mendapatkan informasi berupa jumlah data asli klasifikasi status gizi balita yang terdiri dari jenis kelamin, umur, berat badan, tinggi badan dan status gizi.

2.4. Algoritma C4.5

Algoritma C4.5 dikenal sebagai teknik pohon keputusan yang terkenal karena berbagai keunggulannya. Salah satunya adalah kemampuannya untuk memproses data numerik dan diskrit, serta mampu menangani nilai-nilai atribut yang hilang. Algoritma ini menghasilkan aturan-aturan yang mudah dimengerti dan merupakan salah satu yang paling cepat dibandingkan dengan algoritma lainnya. Kelebihan lainnya adalah akurasi prediksi yang tinggi, di mana model dapat memprediksi label kelas dengan baik untuk data baru yang belum diketahui sebelumnya. C4.5 juga dapat melakukan prediksi yang akurat bahkan jika data mengandung nilai-nilai atribut yang hilang. Kemampuan skalabilitasnya memungkinkan pembangunan model secara efisien dalam data yang besar, dengan model yang dihasilkan memiliki tingkat interpretabilitas yang signifikan sehingga mudah dipahami [8].

Algoritma C4.5 adalah salah satu teknik klasifikasi yang memanfaatkan pembangunan pohon keputusan. Dalam proses ini, setiap cabang dari pohon mengarah ke node lain atau keputusan akhir. Langkah-langkah pembentukan pohon keputusan meliputi transformasi data dari format tabel menjadi model pohon, konversi model pohon menjadi rule, dan penyederhanaan rule.



Gambar 2. Alur Pohon Keputusan Algoritma C4.5

Pada proses pengolahan data menggunakan algoritma C4.5, data yang sudah diseleksi sesuai dengan kebutuhan penelitian akan diolah dengan menghitung menggunakan rumus-rumus yang terkait dengan algoritma C4.5. Adapun rumus – rumus algoritma C4.5 sebagai berikut:

a. Entropy

Entropy (S) adalah jumlah bit yang diperlukan untuk mengekstraksi kelas tertentu (+ atau -) dari sejumlah data acak dalam ruang sampel S. Entropy bisa dianggap sebagai kebutuhan bit untuk merepresentasikan kelas tersebut. Entropy digunakan untuk menilai ketidakmurnian S. Untuk menghitung nilai Entropy, langkah-langkahnya sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (1)$$

Keterangan:

S = Himpunan kasus

n = Jumlah partisi S

p_i = probabilitas yang didapatkan dari jumlah (ya/tidak) dibagi total kasus

b. Gain

Gain (S, A) adalah informasi yang diperoleh dari atribut A relatif terhadap data output S. Informasi ini diukur dari output data atau variabel dependen S yang dikelompokkan berdasarkan atribut A, dan dinotasikan sebagai gain (S, A). Pemilihan atribut akan berdasarkan nilai gain tertinggi dari semua atribut yang tersedia. Rumusnya adalah sebagai berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} Entropy(S_i) \quad (2)$$

Keterangan :

S = himpunan kasus

A = atribut

n = jumlah partisi A

$|S_i|$ = jumlah kasus pada partisi ke-i

$|S|$ = jumlah kasus dalam S

c. Split Info

Split Info adalah rumus yang digunakan untuk membagi sejumlah data yang dihasilkan dari nilai gain dalam perhitungannya, seperti yang dijelaskan berikut:

$$SplitInfo = \sum_{i=1}^k p(V_i|S). \log_2 p(V_i|S) \quad (3)$$

Keterangan:

S = Jumlah data

S_i = Jumlah masing-masing pada setiap atribut.

d. Gain ratio

Gain ratio adalah rumus yang digunakan untuk menentukan dimana perhitungan selanjutnya yang dihitung ulang lagi dan rumusnya sebagai berikut:

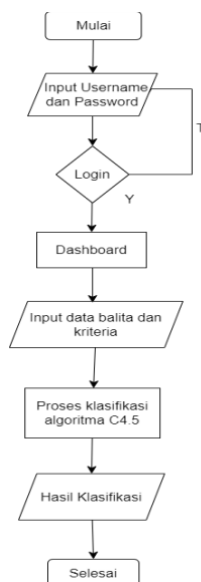
$$Gain\ Ratio(S, j) = \frac{Gain(S, j)}{Split\ Info(S, j)} \quad (4)$$

Setelah nilai tersebut diperoleh, langkah selanjutnya adalah membuat pohon keputusan berdasarkan hasil

perhitungan tersebut [9]. Pohon keputusan merupakan hasil proses dari perhitungan nilai entropy dan information gain yang dilakukan secara berulang-ulang hingga semua atribut memiliki kelas dan tidak bisa dilakukan proses perhitungan lagi [10].

2.5 Perancangan Sistem

Flowchart sistem berfungsi untuk memberi gambaran bagaimana sistem aplikasi klasifikasi status gizi balita akan dibangun. dapat dilihat pada gambar 3 dibawah ini.



Gambar 3. Flowchart Metode C4.5

2.6 Pengujian Sistem

Tahap ini adalah tahap pengujian terhadap sistem yang telah dibuat:

- Pengujian dilakukan dengan merujuk pada data balita yang telah disusun pada tahap perancangan. Pengujian ditujukan untuk menguji fungsi-fungsi khusus dari aplikasi yang dirancang. Aplikasi berbasis Android yang diuji dilihat berdasarkan keluaran yang dihasilkan dari data dan dilakukan pencocokan dengan perhitungan manual. Dari output, kemampuan program dalam memenuhi kebutuhan dapat dinilai sembari dapat diketahui kesalahan-kesalahannya.
- Pengujian black box merupakan pengujian yang bertujuan untuk menunjukkan fungsi perangkat lunak tentang cara beroperasinya, apakah masukan data dan keluaran telah berjalan sebagaimana yang telah di harapkan sebelumnya.

2.7 Penerapan

Penerapan atau penggunaan pada sistem ini ialah pihak Puskesmas yang ada di Kecamatan Bilah Barat yaitu Puskesmas Janji dan Puskesmas Suka Makmur untuk dapat mengklasifikasikan status gizi balita berdasarkan atribut yang sudah sesuai dengan antropometri gizi balita dengan sistem yang sudah dibangun menggunakan metode Algoritma C4.5.

3. Hasil dan Pembahasan

Adapun data atau kebutuhan informasi yang dibutuhkan untuk penelitian ini adalah dataset dengan parameter utama yakni usia (bulan), berat badan (kg), dan tinggi badan (cm). Penelitian ini dilakukan di Puskesmas yang terdapat di Kecamatan Bilah Barat yaitu Puskesmas Janji dan Puskesmas Suka Makmur dengan total 502 sampel data balita. 270 berjenis kelamin laki-laki dan 232 perempuan, 407 balita dan 95 balita, 434 gizi normal, 48 gizi kurang, 5 gizi buruk, 15 gizi lebih.

Tabel 1. berikut ini merupakan 30 sampel dari 502 data status gizi balita pada Puskesmas di Kecamatan Bilah Barat yaitu Puskesmas Janji Dan Puskesmas Suka Makmur.

Tabel 1. 30 Sampel Data Status Gizi Balita

No	Nama	JK	Usia	BB	TB	Status
1.	Ezra	L	40	11,3	89	Kurang
2.	Zaka	L	10	7	60	Kurang
3.	Fatif	L	9	13	73	Lebih
4.	Arsila	P	46	15	103	Normal
5.	Arvino	L	55	18	103	Normal
6.	Azlan	L	9	8,6	65	Normal
7.	Zenia	P	7	8	65	Normal
8.	Rakhel	P	22	10	81	Normal
9.	Arshaka	L	9	7	67	Kurang
10.	Alrifki	L	6	10	70	Lebih
11.	Reyhan	L	11	7,8	69	Normal
12.	Rafasya	L	21	9,5	81	Normal
13.	M. Alfais	L	10	9,5	71	Normal
14.	Fajar	L	46	17	102	Normal
15.	Ziovan	L	12	11	80	Normal
16.	Salwa	P	16	9,5	74	Normal
17.	Dede Ayu	P	9	7,6	71	Normal
18.	Syalwa	P	30	6	50	Buruk
19.	Hafizah	P	3	5,7	58	Normal
20.	Akbar	L	49	14	98	Normal
21.	Byan	L	5	6,7	64	Normal
22.	Zulfikar	L	10	7,6	72	Normal
23.	Azan	L	30	9,5	78	Lebih
24.	Utami	P	37	12	90	Normal
25.	Utama	L	37	13,5	92	Normal
26.	Prabu	L	7	7,2	70	Normal
27.	Faiz	L	7	7,1	67	Normal
28.	Mera	P	45	14	99	Normal
29.	Amel	P	47	15	104	Normal
30.	Wahyudi	L	5	8	74	Normal

3.1 Perhitungan Entropy

Perhitungan nilai *entropy* menggunakan rumus pada persamaan (1) yang dapat dilihat sebagai berikut.

a) Entropy Total

$$\begin{aligned} \text{Entropy Total} = & \left(-\frac{\text{Jlh kasus Normal}}{\text{jlh semua kasus}} \right) \\ & * \log \frac{\text{jlh kasus Normal}}{\text{jlh semua kasus}} \\ & + \left(-\frac{\text{Jlh kasus Kurang}}{\text{jlh semua kasus}} \right) \\ & * \log \frac{\text{jlh kasus Kurang}}{\text{jlh semua kasus}} \\ & + \left(-\frac{\text{Jlh kasus Buruk}}{\text{jlh semua kasus}} \right) \\ & * \log \frac{\text{jlh kasus Buruk}}{\text{jlh semua kasus}} \\ & + \left(-\frac{\text{Jlh kasus Lebih}}{\text{jlh semua kasus}} \right) \\ & * \log \frac{\text{jlh kasus Lebih}}{\text{jlh semua kasus}} \end{aligned}$$

$$\begin{aligned} \text{Entropy Total} = & \left(-\frac{434}{502} * \log \frac{434}{502} \right) \\ & + \left(-\frac{48}{502} * \log \frac{48}{502} \right) \\ & + \left(-\frac{5}{502} * \log \frac{5}{502} \right) \\ & + \left(-\frac{15}{502} * \log \frac{15}{502} \right) \\ \text{Entropy Total} = & 0.722929 \end{aligned}$$

b) Entropy Berat Badan

1. Kriteria 1 (1-4 Kg)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{1}{2} * \log \frac{1}{2} \right) + \left(-\frac{1}{2} * \log \frac{1}{2} \right) \\ & + \left(-\frac{0}{2} * \log \frac{0}{2} \right) + \left(-\frac{0}{2} * \log \frac{0}{2} \right) \\ \text{Entropy} = & 1 \end{aligned}$$

2. Kriteria 2 (4-9 Kg)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{75}{85} * \log \frac{75}{85} \right) + \left(-\frac{7}{85} * \log \frac{7}{85} \right) \\ & + \left(-\frac{2}{85} * \log \frac{2}{85} \right) \\ & + \left(-\frac{1}{85} * \log \frac{1}{85} \right) \\ \text{Entropy} = & 0.658651 \end{aligned}$$

3. Kriteria 3 (10-18 Kg)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{358}{415} * \log \frac{358}{415} \right) \\ & + \left(-\frac{40}{415} * \log \frac{40}{415} \right) \\ & + \left(-\frac{3}{415} * \log \frac{3}{415} \right) \\ & + \left(-\frac{14}{415} * \log \frac{14}{415} \right) \\ \text{Entropy} = & 0.725543 \end{aligned}$$

c) Entropy Usia

1. Kriteria 1 (1-12 Bulan)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{243}{276} * \log \frac{243}{276} \right) + \left(-\frac{22}{276} * \log \frac{22}{276} \right) \\ & + \left(-\frac{3}{276} * \log \frac{3}{276} \right) \\ & + \left(-\frac{8}{110} * \log \frac{8}{110} \right) \\ \text{Entropy} = & 0.671597 \end{aligned}$$

2. Kriteria 2 (13-36 Bulan)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{109}{131} * \log \frac{109}{131} \right) + \left(-\frac{13}{131} * \log \frac{13}{131} \right) \\ & + \left(-\frac{2}{131} * \log \frac{2}{131} \right) \\ & + \left(-\frac{7}{131} * \log \frac{7}{131} \right) \\ \text{Entropy} = & 0.869382 \end{aligned}$$

3. Kriteria 3 (37-60 Bulan)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{82}{95} * \log \frac{82}{95} \right) \\ & + \left(-\frac{13}{95} * \log \frac{13}{95} \right) \\ & + \left(-\frac{0}{95} * \log \frac{0}{95} \right) \\ & + \left(-\frac{0}{95} * \log \frac{0}{95} \right) \\ \text{Entropy} = & 0.575908 \end{aligned}$$

d) Entropy Tinggi Badan

1. Kriteria 1 (40-70 cm)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{182}{210} * \log \frac{182}{210} \right) \\ & + \left(-\frac{20}{210} * \log \frac{20}{210} \right) \\ & + \left(-\frac{4}{210} * \log \frac{4}{210} \right) \\ & + \left(-\frac{4}{210} * \log \frac{4}{210} \right) \\ \text{Entropy} = & 0.719687 \end{aligned}$$

2. Kriteria 2 (70-90 cm)

$$\begin{aligned} \text{Entropy} = & \left(-\frac{177}{207} * \log \frac{177}{207} \right) + \left(-\frac{20}{207} * \log \frac{20}{207} \right) \\ & + \left(-\frac{1}{207} * \log \frac{1}{207} \right) \\ & + \left(-\frac{9}{207} * \log \frac{9}{207} \right) \\ \text{Entropy} = & 0.752742 \end{aligned}$$

3. Kriteria 3 (91-120 cm)

$$\begin{aligned}
 Entropy &= \left(-\frac{75}{85} * \log \frac{75}{85} \right) \\
 &+ \left(-\frac{8}{85} * \log \frac{8}{85} \right) \\
 &+ \left(-\frac{0}{85} * \log \frac{0}{85} \right) \\
 &+ \left(-\frac{2}{85} * \log \frac{2}{85} \right) \\
 Entropy &= 0.607492
 \end{aligned}$$

e) Entropy Jenis Kelamin

1. Kriteria 1 (L = Laki-laki)

$$\begin{aligned}
 Entropy &= \left(-\frac{218}{270} * \log \frac{218}{270} \right) \\
 &+ \left(-\frac{40}{270} * \log \frac{40}{270} \right) \\
 &+ \left(-\frac{4}{270} * \log \frac{4}{270} \right) \\
 &+ \left(-\frac{8}{270} * \log \frac{8}{270} \right) \\
 Entropy &= 0.897773
 \end{aligned}$$

2. Kriteria 2 (P = Perempuan)

$$\begin{aligned}
 Entropy &= \left(-\frac{216}{232} * \log \frac{216}{232} \right) + \left(-\frac{8}{232} * \log \frac{8}{232} \right) \\
 &+ \left(-\frac{1}{232} * \log \frac{1}{232} \right) \\
 &+ \left(-\frac{7}{232} * \log \frac{7}{232} \right) \\
 Entropy &= 0.449760
 \end{aligned}$$

3.2 Perhitungan Gain

Perhitungan nilai *gain* menggunakan rumus pada persamaan (2) yang dapat dilihat sebagai berikut.

$$\begin{aligned}
 Gain &= Entropy Total \\
 &- \left(\left(\frac{jlh \text{ kriteria } 1}{jlh \text{ kasus kriteria } 1} \right) \right. \\
 &\quad * entropy \text{ kriteria } 1 \left. \right) \\
 &+ \left(\left(\frac{jlh \text{ kriteria } 2}{jlh \text{ kasus kriteria } 2} \right) \right. \\
 &\quad * entropy \text{ kriteria } 2 \left. \right) \\
 &+ \left(\left(\frac{jlh \text{ kriteria } 3}{jlh \text{ kasus kriteria } 3} \right) \right. \\
 &\quad * entropy \text{ kriteria } 3 \left. \right)
 \end{aligned}$$

a) *Gain* Berat Badan

$$\begin{aligned}
 Gain &= 0.722929 \\
 &- \left(\left(\frac{2}{502} * 1 \right) \right. \\
 &\quad + \left(\frac{85}{502} * 0.6586 \right) \\
 &\quad \left. + \left(\frac{415}{502} * 0.7255 \right) \right)
 \end{aligned}$$

$$Gain = 0.00761$$

b) *Gain* Usia

$$\begin{aligned}
 Gain &= 0.722929 \\
 &- \left(\left(\frac{276}{502} * 0.671597 \right) \right. \\
 &\quad + \left(\frac{131}{502} * 0.869382 \right) \\
 &\quad \left. + \left(\frac{95}{502} * 0.575908 \right) \right)
 \end{aligned}$$

$$Gain = 0.017826$$

c) *Gain* Tinggi Badan

$$\begin{aligned}
 Gain &= 0.722929 \\
 &- \left(\left(\frac{210}{502} * 0.719678 \right) \right. \\
 &\quad + \left(\frac{207}{502} * 0.752742 \right) \\
 &\quad \left. + \left(\frac{85}{502} * 0.607492 \right) \right)
 \end{aligned}$$

$$Gain = 0.008608$$

d) *Gain* Jenis Kelamin

$$\begin{aligned}
 Gain &= 0.722929 \\
 &- \left(\left(\frac{270}{502} * 0.897773 \right) \right. \\
 &\quad \left. + \left(\frac{232}{502} * 0.449760 \right) \right)
 \end{aligned}$$

$$Gain = 0.032205$$

3.3 Perhitungan Split Info

Perhitungan nilai *split info* menggunakan rumus pada persamaan (3) yang dapat dilihat sebagai berikut.

$$\text{Split Info} = - \left(\left(\frac{\text{jlh kriteria 1}}{\text{jlh kasus kriteria 1}} \right) * \log_2 \left(\frac{\text{jlh kriteria 1}}{\text{jlh kasus kriteria 1}} \right) \right. \\ + \left(\frac{\text{jlh kriteria 2}}{\text{jlh kasus kriteria 2}} \right) * \log_2 \left(\frac{\text{jlh kriteria 2}}{\text{jlh kasus kriteria 2}} \right) \\ + \left(\frac{\text{jlh kriteria 3}}{\text{jlh kasus kriteria 3}} \right) * \log_2 \left(\frac{\text{jlh kriteria 3}}{\text{jlh kasus kriteria 3}} \right) \left. \right)$$

a) *Split Info* Berat Badan

$$\text{Split Info} = - \left(\left(\frac{2}{502} * \log_2 \left(\frac{2}{502} \right) \right) \right. \\ + \left(\frac{85}{502} * \log_2 \left(\frac{85}{502} \right) \right) \\ + \left(\frac{415}{502} * \log_2 \left(\frac{415}{502} \right) \right) \left. \right)$$

$$\text{Split info} = 2.356713$$

b) *Split Info* Usia

$$\text{Split Info} = - \left(\left(\frac{276}{502} * \log_2 \left(\frac{276}{502} \right) \right) \right. \\ + \left(\frac{131}{502} * \log_2 \left(\frac{131}{502} \right) \right) \\ + \left(\frac{95}{502} * \log_2 \left(\frac{95}{502} \right) \right) \left. \right)$$

$$\text{Split info} = 0.892526$$

c) *Split Info* Tinggi Badan

$$\text{Split Info} = - \left(\left(\frac{210}{502} * \log_2 \left(\frac{210}{502} \right) \right) \right. \\ + \left(\frac{207}{502} * \log_2 \left(\frac{207}{502} \right) \right) \\ + \left(\frac{85}{502} * \log_2 \left(\frac{85}{502} \right) \right) \left. \right)$$

$$\text{Split info} = 0.835542$$

d) *Split Info* Jenis Kelamin

$$\text{Split Info} = - \left(\left(\frac{270}{502} * \log_2 \left(\frac{270}{502} \right) \right) \right. \\ + \left(\frac{232}{502} * \log_2 \left(\frac{232}{502} \right) \right) \left. \right)$$

$$\text{Split info} = 0.995862$$

3.4 Perhitungan Gain Ratio

Perhitungan nilai *gain ratio* menggunakan rumus pada persamaan (4) yang dapat dilihat sebagai berikut.

$$\text{Gain Ratio} = \frac{\text{Gain Kelas } x}{\text{Split Info Kelas } x}$$

a) *Gain Ratio* Berat Badan

$$\text{Gain Ratio} = \frac{0.007617}{2.356713}$$

$$\text{Gain Ratio} = 0.003232$$

b) *Gain Ratio* Usia

$$\text{Gain Ratio} = \frac{0.0178268}{0.892562}$$

$$\text{Gain Ratio} = 0.019973$$

c) *Gain Ratio* Tinggi Badan

$$\text{Gain Ratio} = \frac{0.008608}{0.835542}$$

$$\text{Gain Ratio} = 0.010302$$

d) *Gain Ratio* Jenis Kelamin

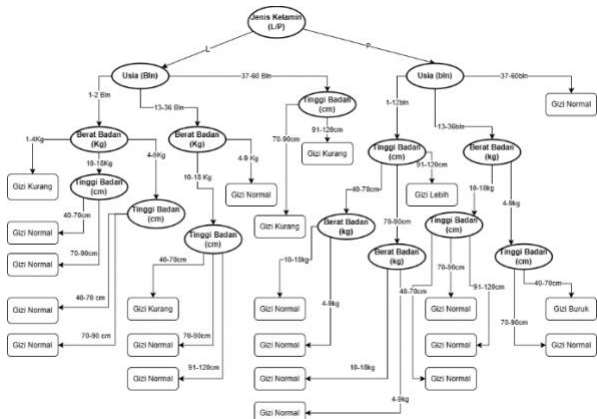
$$\text{Gain Ratio} = \frac{0.032205}{0.995862}$$

$$\text{Gain Ratio} = 0.0323391$$

berdasarkan hasil perhitungan tersebut diperoleh nilai tertinggi pada parameter Jenis Kelamin dengan nilai *gain ratio* adalah 0.032339. Parameter dengan nilai *gain ratio* tertinggi akan dijadikan prioritas dalam struktur pohon Keputusan. Dengan begitu Jenis Kelamin akan dijadikan sebagai akar (*root*) dari struktur pohon keputusan.

3.5 Pembentukan Pohon Keputusan dan Rule

Pada tahap ini akan dilakukan perancangan basis pengetahuan menggunakan *software data mining* RapidMiner yang menghasilkan pengetahuan yang dibutuhkan oleh sistem agar menghasilkan kesimpulan dari aturan yang ditetapkan. Adapun fakta-fakta atau aturan-aturan tersebut diperoleh berdasarkan *mining* 502 dataset yang diperoleh dari puskesmas. Hasil *decision tree* berdasarkan *software* RapidMiner dapat dilihat pada gambar 4 dan tabel 2.



Gambar 4. Hasil Pembentukan Pohon Keputusan

Tabel 2. Tabel Analisis Jumlah Kasus dengan Interval

Paramet er	Interv al	Jumlah Kasus				Total Kasus
		Normal	Kurang	Buruk	Lebih	
Jenis Kelamin	Laki-laki	218	40	4	8	270
	Perempuan	216	8	1	7	232
Usia	1-12 bulan	243	22	3	8	276
	13-36 bulan	109	13	2	7	131
	37-60 bulan	82	13	0	0	95
	61-120 bulan	82	13	0	0	95
Berat Badan	1-4 kg	1	1	0	0	2
	4-9 kg	75	7	2	1	85
	10-18 kg	358	40	3	14	415
	40-70 kg	182	20	4	4	210
Tinggi Badan	70-90 cm	177	20	1	9	207
	91-120 cm	75	8	0	2	85

3.6 Perancangan Interface

Perancangan antarmuka input/output ini bertujuan untuk menciptakan struktur data yang efisien, terorganisir, dan dapat diakses dengan mudah guna mendukung kebutuhan spesifik suatu organisasi atau sistem.



Gambar 5. Login Page

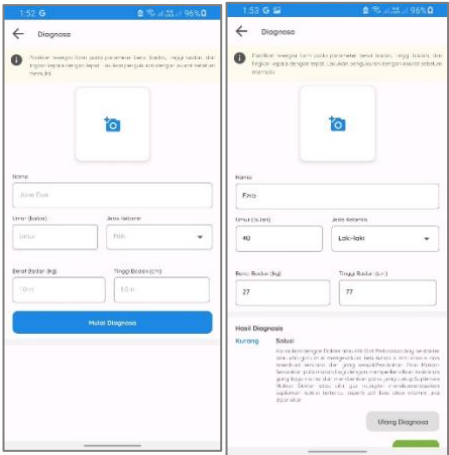
Login Page merupakan tampilan yang pertama kali saat aplikasi dijalankan. jika ingin masuk maka masukkan email dan password apabila sudah punya akun, jika

belum maka pengguna harus membuat akun terlebih dahulu dibagian Register.



Gambar 6. Beranda

Gambar 6. Merupakan Fitur-fitur yang terdapat dalam tampilan halaman yaitu: diagnosis, pengaturan, feed, edukasi, account, tentang developer.



Gambar 7. Diagnosa

Pada Gambar 7 dapat dilihat terdapat button mulai diagnosa, setelah user mulai diagnosis user akan menambahkan parameter umur, jenis kelamin, berat badan, tinggi badan untuk mendapatkan hasil berupa diagnosis serta penanganannya menggunakan rule berdasarkan algoritma C4.5.

3.7 Pengujian Akurasi

Untuk menguji Tingkat akurasi dari sistem, maka akan dibuat 20 skenario dengan mengambil data pada dataset secara acak. Tabel skenario pengujian dapat dilihat pada Tabel 2.

Tabel 3. Skenario Pengujian

No.	Nama	Umur	JK	BB (kg)	TB (cm)	Hasil	Fakta
1	Ezra	40	L	11,3	89	Kurang	Kurang
2	Zaka	10	L	7	60	Normal	Kurang

3	Arsila	46	P	15	103	Nor mal	Nor mal
4	Arvin o	55	L	18	103	Nor mal	Nor mal
5	Azlan	9	L	8,6	65	Nor mal	Nor mal
6	Zenia	7	P	8	65	Nor mal	Nor mal
7	Rakhe l	22	P	10	81	Nor mal	Nor mal
8	Arsha ka	9	L	7	67	Nor mal	Kura ng
9	Rifki	6	L	10	70	Nor mal	Lebi h
10	Reyha n	11	L	7,8	69	Nor mal	Nor mal
11	Rafas ya	21	L	9,5	81	Nor mal	Nor mal
12	M. Alfais	10	L	9,5	71	Nor mal	Nor mal
13	Fajar	46	L	17	102	Nor mal	Nor mal
14	Ziova n	12	L	11	80	Nor mal	Nor mal
15	Salwa	16	P	9,5	74	Nor mal	Nor mal
16	Dede Ayu	9	P	7,6	71	Nor mal	Nor mal
17	Hafiz ah	3	P	5,7	58	Nor mal	Nor mal
18	Akbar	49	L	14	98	Nor mal	Nor mal
19	Byan	5	L	6,7	64	Nor mal	Nor mal
20	Zulfik ar	10	L	7,6	72	Nor mal	Nor mal

Berdasarkan skenario pengujian pada Tabel 2. dapat diketahui bahwa terdapat 3 hasil klasifikasi yang salah dan 17 klasifikasi benar. Berdasarkan hasil pengujian tersebut maka dapat dihitung tingkat akurasi sebagai berikut.

$$Akurasi = \frac{Jumlah\ pengujian\ benar}{Total\ pengujian} \times 100\%$$
$$Akurasi = \frac{17}{20} \times 100\% = 85\%$$

Dari hasil pengujian tersebut dapat diperoleh kesimpulan sistem berhasil mengklasifikasi dengan benar 17 skenario pengujian. Persentase akurasi sistem berdasarkan pengujian tersebut adalah 85%.

3.8 Confusion Matrix

Untuk menguji model *decision tree* yang dihasilkan dari 502 dataset status gizi balita, maka akan dilakukan evaluasi model menggunakan teknik *confusion matrix*. *Confusion matrix* sering dipakai dalam penelitian untuk menilai hasil serta mengevaluasi efektivitas suatu metode klasifikasi. Pengujian *confusion matrix* menggunakan software RapidMiner dapat dilihat pada gambar 8.

accuracy: 80.00%					
	true Kurang	true Lebih	true Normal	true Buruk	class precision
pred Kurang	2	0	0	0	100.00%
pred Lebih	0	0	0	0	0.00%
pred Normal	13	3	78	2	81.25%
pred Buruk	0	0	2	0	0.00%
class recall	13.33%	0.00%	87.50%	0.00%	

Gambar 8. Hasil Confusion Matrix

Berdasarkan gambar 8. dapat diketahui bahwa model *decision tree* dari 502 dataset status gizi balita yang diuji dengan teknik *confusion matrix* mendapatkan *accuracy* sebesar 80%. Pengujian dilakukan dengan membagi 20% dari total dataset yakni 100 data untuk diuji.

4. Kesimpulan

Berdasarkan penelitian yang dilakukan, algoritma C4.5 menunjukkan performa yang optimal dalam mendeteksi status gizi balita berdasarkan parameter jenis kelamin, berat badan, tinggi badan, dan usia yang dilaporkan. Kelebihan algoritma C4.5 terletak pada kemampuannya mengidentifikasi pola hubungan antara parameter-parameter tersebut secara sistematis. Selanjutnya berdasarkan pengujian yang telah dilakukan menggunakan teknik *confusion matrix* dan melakukan pembagian data dengan perbandingan 80:20 dari total 502 dataset yakni 402 data training dan 100 data testing. Maka dari tahap pengujian *confusion matrix* yang telah dilakukan diperoleh tingkat *accuracy* sebesar 80%.

Aplikasi ini memberikan kemudahan akses dan mobilitas bagi pengguna untuk memantau status gizi balita dari berbagai lokasi, sesuai dengan karakteristik kebutuhan pemantauan kontinu terhadap kesehatan balita. Dengan penelitian ini, diharapkan teknologi dapat memberikan kontribusi yang signifikan dalam membantu masyarakat dan memberikan solusi dalam deteksi status gizi balita berdasarkan parameter yang diukur.

Daftar Rujukan

[1] Setyawati, V. A. V., & Hartini, E.(2018). Buku Ajara Dasar Ilmu Gizi Kesehatan Masyarakat, Yogyakarta: Deepublish

[2] Numaliza, N., & Herlina, S. (2018). Hubungan Pengetahuan dan Pendidikan Ibu terhadap Status Gizi Balita. *KESMARS: Jurnal Kesehatan Masyarakat, Manajemen Dan Administrasi Rumah Sakit*, 1(1), 44–48. <https://doi.org/10.31539/kesmars.v1i1.171>

[3] Yuliansyah, M. R., B, M., & Franz, A. (2022). Perbandingan Metode K-Nearest Neighbors dan Naïve Bayes Classifier Pada Klasifikasi Status Gizi Balita di Puskesmas Muara Jawa Kota Samarinda. *Adopsi Teknologi Dan Sistem Informasi(ATASI)*, 1(1), 8–20.

-
- [4] Kementrian Kesehatan RI 2022. <https://upk.kemkes.go.id/new/kementerian-kesehatan-rilis-hasil-survei-status-gizi-indonesia-ssgi-tahun-2022>
 - [5] Hafizan, H. S. T. B., & Putri, A. N. S. T. B. (2020). Penerapan Metode Klasifikasi Decision Tree Pada Status Gizi. *Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, 1(2), 68–72.
 - [6] Wanto, A. dkk. (2020). Data Mining: Algoritma dan Implementasi. Yayasan Kita Menulis.
 - [7] Islam, H. I., Mulyadien, M. K., & Enri, U. (2022). Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita. 116-125.
 - [8] Dhika, H & F. Destiawati.(2018). Penerapan Algoritma C45 Untuk Penilaian Karyawan Pada Restoran ISSN(E): 2830-3083 221 SENTIMAS – 25 Agustus 2022 Cepat Saji.
 - [9] Batubara, Dinda Nabila, and Agus Perdana Windarto. 2019. “Analisa Klasifikasi Data Mining Pada Tingkat Kepuasan Pengunjung Taman Hewan Pematang Siantar Dengan Algoritma.” KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer) 3(1):588–92. doi: 10.30865/komik.v3i1.1664.
 - [10] Cynthia, Eka Pandu, and Edi Ismanto. 2018. “Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji.” Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika) 3(July):1. doi: 10.30645/jurasik.v3i0.60.