

Klasifikasi Status Gizi Balita Menggunakan Algoritma Decision Tree dalam RapidMiner di Desa Gombang

Putri Aryaningsih¹, Mutiara Dwi Triapanca², Tommy Dwi Putra³

^{1,2,3} Program Studi Sistem Informatika, STMIK Amikom Surakarta

e-mail: putri.10316@mhs.amikomsolo.ac.id, mutiara.10317@mhs.amikomsolo.ac.id,
tommy@dosen.amikomsolo.ac.id

Abstrak - Status gizi merupakan indikator penting dalam menilai kesehatan dan pertumbuhan balita. Penelitian ini bertujuan untuk membangun dan menguji model klasifikasi status gizi balita di Desa Gombang, Kabupaten Boyolali, menggunakan metode Decision Tree pada platform RapidMiner. Data yang digunakan merupakan data primer dari Pos Kesehatan Desa Gombang sebanyak 275 data balita yang dikumpulkan pada bulan April–Mei 2025. Atribut yang digunakan meliputi berat badan, tinggi badan, lingkaran lengan atas, lingkaran lengan kanan, indeks massa balita (IMB), dan status gizi sebagai label. Model dibangun melalui proses pelatihan dan pengujian data, serta dievaluasi menggunakan confusion matrix. Hasil klasifikasi menunjukkan bahwa algoritma Decision Tree mampu mengelompokkan status gizi ke dalam kategori Normal, Gizi Kurang, dan Gizi Lebih secara akurat. Evaluasi model menunjukkan kinerja yang sangat baik, dengan akurasi pada data pelatihan mencapai tingkat yang tinggi dan performa serupa pada data pengujian. Atribut yang paling berpengaruh dalam klasifikasi adalah IMB, diikuti oleh berat badan, tinggi badan, dan lingkaran lengan atas. Temuan ini membuktikan bahwa metode Decision Tree efektif digunakan dalam klasifikasi status gizi balita dan dapat diimplementasikan sebagai alat bantu pengambilan keputusan bagi tenaga kesehatan desa dalam upaya deteksi dini gizi buruk dan pencegahan stunting.

Kata kunci: Status Gizi, Decision Tree, RapidMiner, Klasifikasi, Desa Gombang

Abstract - Nutritional status was a crucial indicator used to assess the health and development of toddlers, especially in rural areas where access to healthcare services was often limited. This study aimed to develop and evaluate a classification model for the nutritional status of toddlers in Gombang Village, Boyolali Regency, using the Decision Tree method on the RapidMiner platform. The research used primary data obtained from the Gombang Village Health Post, consisting of 275 toddler records collected during April–May 2025. The attributes included body weight, height, mid-upper arm circumference, right arm circumference, toddler body mass index (IMB), and nutritional status as the classification label. The model was developed through training and testing processes and was evaluated using a confusion matrix. The results showed that the Decision Tree algorithm successfully classified nutritional status into Normal, Undernutrition, and Overnutrition categories with high accuracy. The evaluation indicated that the model performed well on both training and testing datasets. The most influential attributes in the classification process were IMB, followed by body weight, height, and mid-upper arm circumference. These findings confirmed that the Decision Tree method was effective in classifying toddler nutritional status and could be implemented as a decision-support tool for local health workers to assist in the early detection of malnutrition and the prevention of stunting.

Keywords: *Nutritional Status, Decision Tree, RapidMiner, Classification, Gombang Village*

I. PENDAHULUAN

Salah satu indikator penting untuk menilai kesehatan dan kesejahteraan anak usia dini, terutama pada balita (anak usia bawah lima tahun) adalah status gizi mereka. Balita dengan gangguan gizi, baik kurang atau lebih gizi, rentan terhadap berbagai gangguan pertumbuhan yang dapat berdampak jangka panjang. Sebagai negara berkembang, Indonesia masih menghadapi masalah gizi yang signifikan, terutama di daerah pedesaan seperti Desa Gombang, Boyolali. Seperti yang ditunjukkan oleh berbagai survei kesehatan, balita dengan status gizi buruk masih cukup banyak dan sering kali disembunyikan dari data agregat nasional, yang tidak secara akurat menunjukkan kondisi lokal.

Indikator antropometri seperti berat badan menurut umur (BB/U), tinggi badan menurut umur (TB/U), dan berat badan menurut tinggi badan (BB/TB) biasanya digunakan untuk mengukur dan mengklasifikasikan status gizi balita [1]. Namun, pengolahan data antropometri secara manual seringkali memakan waktu, tidak efisien, dan rentan terhadap kesalahan. Pendekatan teknologi data mining dan kecerdasan buatan, seperti algoritma Decision Tree, sangat penting dalam hal ini. Karena mampu membuat model interpretatif, visual, dan mudah dipahami oleh tenaga kesehatan dan pemangku kepentingan di tingkat desa, Decision Tree adalah metode klasifikasi yang populer.

Penelitian ini menggunakan platform RapidMiner karena kelebihan, seperti kemudahan integrasi, antarmuka visual, dan ketersediaan berbagai algoritma data mining yang dapat disesuaikan. Dalam penelitian ini, metode Decision Tree digunakan untuk klasifikasi. Metode ini mirip dengan struktur pohon, dengan setiap node menunjukkan atribut yang akan diuji dan setiap cabang menunjukkan hasil dari atribut tes tersebut [2]. Dalam klasifikasi status gizi, Decision Tree adalah metode yang paling umum digunakan, khususnya algoritma C4.5, karena keunggulannya dalam membuat model yang mudah diinterpretasikan oleh pengguna non-teknis, seperti kader posyandu atau petugas kesehatan di tingkat desa.

[3] Hairani et al. (2022) melakukan penelitian yang menunjukkan bahwa algoritma C4.5 efektif dalam pembuatan aplikasi berbasis web yang memungkinkan untuk mengklasifikasikan status gizi balita. Studi ini menghasilkan sistem klasifikasi yang sangat akurat dan memungkinkan untuk melihat aturan klasifikasi yang sederhana, membuatnya cocok untuk digunakan dalam sistem kesehatan masyarakat yang membutuhkan keterbacaan model yang baik (interpretability). Penelitian ini menjadi rujukan penting dalam mengembangkan sistem klasifikasi gizi karena memberikan contoh penerapan nyata dalam konteks pengguna akhir yang awam terhadap teknologi komputasi canggih.

Tujuan utama dari penelitian ini adalah untuk membangun dan menguji model klasifikasi status gizi balita di Desa Gombang. Tujuan khusus dari penelitian ini adalah untuk mengumpulkan dan memproses data tentang status gizi balita berdasarkan indikator antropometri dan faktor pendukung lainnya yang relevan di lingkungan desa. Selain itu, tujuan penelitian ini adalah untuk menghasilkan rekomendasi berbasis data yang akan membantu upaya untuk menangani gizi buruk dan mencegah stunting di tingkat desa. Oleh karena itu, penelitian ini diharapkan dapat membantu meningkatkan kualitas layanan

kesehatan anak, khususnya dalam hal mendeteksi masalah gizi dini di daerah pedesaan seperti Desa Gombang.

II. TINJAUAN TEORI

2.1 Status Gizi Balita

Menurut Syarifa 2021, gizi sangat penting untuk pertumbuhan dan perkembangan balita dan terkait erat dengan kesehatan dan kecerdasan [4]. Status Gizi bergantung pada data antropometri yang akurat. Indikator antropometri, yang dikembangkan oleh World Health Organization (WHO), adalah berat badan menurut umur (BB/U), tinggi badan menurut umur (TB/U), dan berat badan menurut tinggi badan (BB/TB) [5]. Indikator ini digunakan di seluruh dunia untuk menentukan apakah seorang balita termasuk dalam kategori gizi baik, gizi kurang, atau gizi buruk[5].

2.2 Data Mining dan Klasifikasi

Data mining adalah proses menemukan pola atau informasi penting dari sejumlah besar data [6]. Klasifikasi adalah cabang utama data mining, yaitu teknik untuk mengelompokkan data ke dalam kelas-kelas tertentu berdasarkan fitur atau karakteristik mereka [7]. Untuk masalah status gizi, klasifikasi berguna untuk membagi balita ke dalam kategori berdasarkan data antropometri dan faktor pendukung lainnya, seperti usia, jenis kelamin, riwayat kesehatan, status imunisasi, dan kondisi lingkungan. Algoritma klasifikasi menggunakan pola dari data sebelumnya (training data) untuk memprediksi atau menentukan kelas data baru.

2.3 RapidMiner

RapidMiner adalah platform perangkat lunak pembelajaran mesin dan ilmu data yang kuat yang menawarkan berbagai alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. Desainnya yang ramah pengguna memungkinkan pengguna tanpa pengalaman pemrograman membangun dan menguji berbagai model [8].

RapidMiner adalah program pengolahan data mining yang bersifat open source (terbuka) dan dibuat menggunakan bahasa Java. Ini bekerja dengan sistem operasi apa pun. Analisis teks, ekstraksi pola-pola dari kumpulan data yang besar, dan penggabungan metode statistika, kecerdasan buatan, dan database adalah bagian dari proses pemrosesan teks RapidMiner. Tujuan dari analisis teks ini adalah untuk mendapatkan informasi terbaik dari teks yang diolah[1].

2.4 Decision Tree

Decision Tree adalah diagram alur yang menyerupai struktur pohon. Setiap node internal menunjukkan pengujian terhadap suatu atribut, setiap cabang menunjukkan output dari pengujian, dan setiap node daun menunjukkan distribusi kelas atau kelas-kelas [9]. Untuk mencapai solusinya, *Decision Tree* biasanya menggunakan strategi pencarian dari atas ke bawah atau top-down. Nilai atribut diuji selama proses klasifikasi dari node akar ke node daun. Hanya setelah itu, kelas baru dapat diidentifikasi [10]. Keunggulan utama algoritma ini adalah bahwa hasil klasifikasi dapat diinterpretasikan dan divisualisasikan dalam bentuk pohon keputusan [11]. Ini membuatnya ideal untuk pengguna awam atau dalam bidang non-teknis seperti kesehatan Masyarakat .

III. METODE

3.1 Metode Teori Dasar

Decision Tree adalah salah satu algoritma data mining yang paling sering digunakan sebagai penyelesaian untuk mengklasifikasikan masalah. Selain itu, Decision Tree sangat populer sebagai metode prediksi dan klasifikasi yang sangat akurat [12]. Metode Pohon Keputusan, juga dikenal sebagai "Pohon Keputusan", memiliki struktur yang menyerupai diagram alir [13]. Dalam metode ini, setiap node internal menunjukkan pengujian pada suatu atribut, setiap cabang menunjukkan hasil dari pengujian, dan node daun menunjukkan distribusi kelas [2].

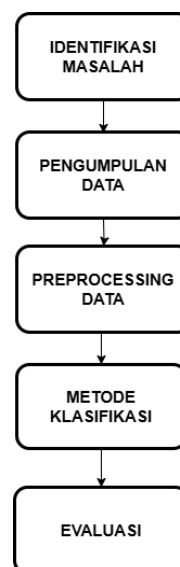
Salah satu algoritma Decision Tree yang banyak digunakan adalah C4.5, yang menghasilkan model dalam bentuk pohon keputusan (*tree structure*) yang mudah dipahami. Setiap cabang dalam pohon menggambarkan keputusan berdasarkan atribut, dan setiap daun mewakili hasil klasifikasi. Seperti yang dinyatakan oleh Azwanti dan Elisa algoritma C4.5 adalah model algoritma yang sangat populer di bidang penelitian. Kelebihan algoritma ini adalah memprediksi dengan menggunakan pohon keputusan, yang membuatnya lebih mudah dipahami [14].

3.2 Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data primer di mana dalam pengumpulan data bersumber dari pihak pertama yaitu Pos Kesehatan Desa Gombang Kabupaten Boyolali, dengan memperoleh sebanyak 275 data balita yang diambil pada bulan April-Mei 2025. Data primer ini diperoleh melalui metode sebagai berikut:

- 1) Wawancara dan Observasi: dilakukan dengan petugas Poskes dan Kader Posyandu guna memperoleh informasi mengenai pencatatan gizi, atribut yang digunakan, serta tantangan dalam klasifikasi manual.
- 2) Pengambilan data secara langsung di lokasi Poskes desa Gombang. Pengambilan data berupa *softfile* berformat xlsx yang berisi tentang informasi kesehatan balita. Atribut dalam data antara lain : nama, bb, tb, lila, lika, imb, status gizi.
- 3) Studi Pustaka : digunakan untuk memperkuat dasar teori penelitian.

3.3 Alur Penelitian



Gambar 1. Alur Penelitian (10 Point)

- 1) Identifikasi Masalah
Menentukan fokus penelitian, yaitu status gizi balita di desa Gombang Kabupaten Boyolali.
- 2) Pengumpulan Data
Melakukan wawancara, observasi, dan pengambilan dataset balita langsung bersumber dari Poskes Gombang dalam bentuk *softfile* (.xlsx).
- 3) *Pre-Processing* Data
Tahap ini melibatkan pembersihan data dan transformasi data sehingga siap untuk di analisis.
- 4) Metode Klasifikasi
Klasifikasi menggunakan RapidMiner untuk membangun model klasifikasi berdasarkan algoritma Decision Tree (C4.5), gunakan RapidMiner. Pemilihan atribut, pelatihan data (*training*), dan pengujian data (*testing*) adalah bagian dari proses ini.
- 5) Evaluasi
Mengukur kinerja model menggunakan *Confusion Matrix*, sensitivitas (*recall*), spesifisitas, dan interpretabilitas hasil klasifikasi. Visualisasi pohon keputusan dievaluasi berdasarkan kemudahan interpretasi oleh pengguna awam.

IV. HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Penelitian ini menggunakan dataset status gizi balita di Desa Gombang, Boyolali dengan jumlah 7 atribut dan 275 data balita yang bersumber langsung dari petugas Poskes Desa Gombang.

Tabel 1. Dataset Status Gizi Balita (10 Point)

No	Dataset Status Gizi Balita Desa Gombang April-Mei 2025					
	BB	TB (m)	Lila (cm)	Lika (cm)	IMB	Status Gizi
1	13,70	0,95	16	50	15,1800554	Normal
2	14,50	1,02	16	49	13,93694733	Normal
3	12,70	0,94	15.5	47	14,37301947	Normal
4	12,30	0,92	17	47	14,53213611	Normal
5	10,60	0,84	16	48	15,02267574	Normal
6	10,50	0,8	16	47	16,40625	Gizi Lebih
7	8,70	0,71	15	45	17,25848046	Gizi Lebih
8	7,40	0,71	13	44	14,67962706	Normal
9	12,80	0,88	16	49	16,52892562	Gizi Lebih
...	...	...	...	...	...	...

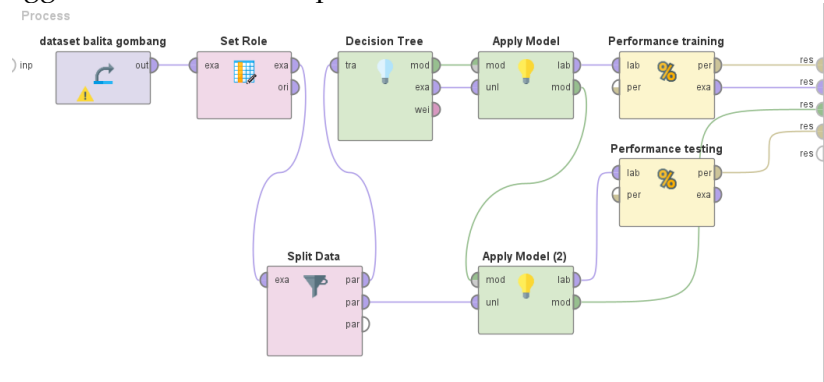
4.2 *Pre-Processing* Data

Pada penelitian ini dari dataset yang berjumlah 275, data dibagi menjadi dua dengan presentase 80% data *training* (pelatihan) didapatkan 220 data dan 20% data *testing* (pengujian) didapatkan 55 data untuk menguji kinerja pada model, memastikan bahwa model dapat membuat prediksi yang akurat. Atribut yang digunakan adalah nama, tanggal ukur, bb, tb, lila, lika, imb, dan status gizi sebagai label klasifikasi.

Proses klasifikasi dilakukan melalui pembentukan model pohon keputusan yang membagi data balita ke dalam tiga kelas: normal, gizi kurang, dan gizi buruk

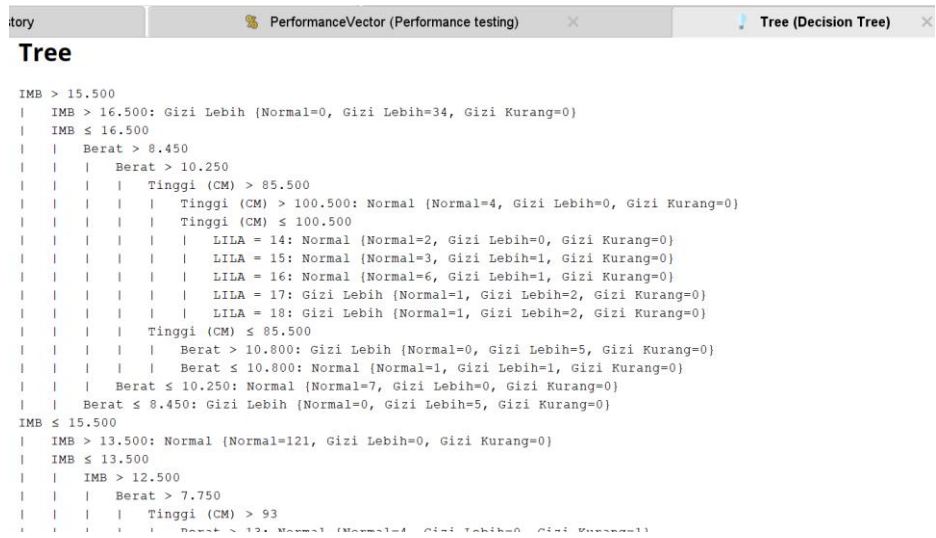
4.3 Algoritma Decision Tree

Model klasifikasi status gizi di bangun menggunakan metode Algoritma Decisio Tree dengan menggunakan software RapidMiner 2024.0.0.

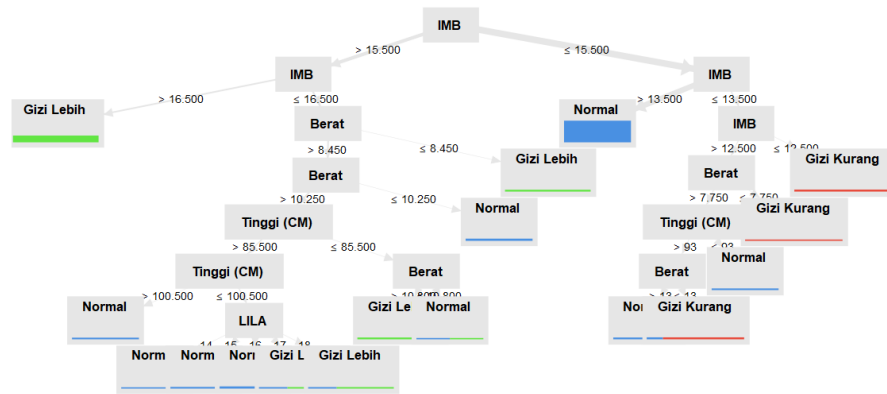


Gambar 2. Proses Klasifikasi Decision Tree RapidMiner

Gambar 2 menunjukkan alur kerja proses RapidMiner ini. Langkah pertama adalah memasukkan data balita Gombang. Kemudian, pada "set role" memilih atribut "status gizi" sebagai target (label) yang akan diprediksi. Data dibagi menjadi dua bagian: data latihan (*training*) dan data uji (*testing*). Model Decision Tree digunakan untuk memprediksi status gizi pada kedua bagian data setelah menggunakan data *training*. Untuk mengetahui seberapa baik model mengklasifikasikan status gizi balita secara akurat dan konsisten, metrik seperti akurasi, recall, dan precision digunakan untuk mengevaluasi kinerjanya. Melalui proses tersebut akan di dapatkan hasil seperti Gambar 3 dan Gambar 4.



Gambar 3. Description Decision Tree



Gambar 4. Hasil Pohon Keputusan

Hasil pohon Keputusan Gambar 3 dan 4, menunjukkan bahwa faktor yang paling berpengaruh dalam klasifikasi status gizi balita adalah Indeks Massa Balita (IMB), diikuti oleh berat badan, tinggi badan, dan lingkaran lengan atas (LILA). Balita dengan IMB di atas 16,5 langsung dikategorikan sebagai "Gizi Lebih", sedangkan balita dengan IMB antara 15,5 dan 16,5 diklasifikasikan berdasarkan kombinasi berat badan, tinggi badan, dan LILA. Misalnya, jika IMB kurang dari 16,5 dan berat badan lebih dari 10,25 kg, balita dikategorikan sebagai "Normal", sedangkan balita dengan IMB di bawah 16,5 diklasifikasikan sebagai "Normal" atau "Gizi Lebih".

4.4 Evaluasi *Confusion Matrix*

Tabel 2. Hasil Akurasi Data Training

accuracy: 96.82%

	true Normal	true Gizi Lebih	true Gizi Kurang	class precision
pred. Normal	152	3	1	97.44%
pred. Gizi Lebih	2	48	0	96.00%
pred. Gizi Kurang	1	0	13	92.86%
class recall	98.06%	94.12%	92.86%	

Tabel 3. Hasil Akurasi Data Testing

accuracy: 98.18%

	true Normal	true Gizi Lebih	true Gizi Kurang	class precision
pred. Normal	38	0	0	100.00%
pred. Gizi Lebih	0	13	0	100.00%
pred. Gizi Kurang	1	0	3	75.00%
class recall	97.44%	100.00%	100.00%	

Hasil dari evaluasi dengan *confusion matrix* dilakukan dua kali yaitu evaluasi hasil klasifikasi terhadap data *training* dan data *testing*. Berdasarkan Tabel 2 dan 3, hasil evaluasi data *training*, akurasi sebesar 96.82% dengan klasifikasi eror sebesar 3.18% dan evaluasi data *testing* akurasi sebesar 98.18% dengan klasifikasi eror sebesar 1.82%, hal ini menunjukkan bahwa model mampu mengenali pola status gizi dengan baik. Hal ini mengindikasikan bahwa atribut-atribut lokal yang digunakan memang relevan dan informatif dalam klasifikasi. Model ini memiliki *recall* yang tinggi untuk kategori gizi buruk, yang sangat penting untuk upaya deteksi dini dan pencegahan stunting.

V. KESIMPULAN DAN SARAN

Studi ini menemukan bahwa metode "Decision Tree" yang digunakan oleh platform RapidMiner dapat mengklasifikasikan status gizi balita di Desa Gombang dengan akurasi sebesar 96,82%. Model mengidentifikasi kategori "Normal", "Gizi Kurang", dan "Gizi Lebih" dengan nilai recall dan ketepatan yang seimbang di atas 92%. Indeks massa balita (IMB), berat badan, tinggi badan, dan LILA adalah atribut yang paling membantu dalam proses klasifikasi. Model ini memiliki keunggulan besar karena dapat membuat pohon keputusan yang mudah dipahami, yang sangat relevan bagi tenaga kesehatan di tingkat desa untuk membantu menemukan masalah gizi dan membuat keputusan yang cepat dan akurat.

Untuk memudahkan klasifikasi dan pelaporan status gizi secara otomatis, model ini dapat diintegrasikan ke dalam sistem informasi posyandu berbasis digital. Selain itu, pelatihan sederhana mengenai penggunaan RapidMiner kepada kader posyandu akan sangat membantu dalam penerapan hasil penelitian ini di lapangan secara berkelanjutan.

VI. DAFTAR PUSTAKA

- [1] H. Hafizan and A. N. Putri, "Penerapan Metode Klasifikasi Decision Tree Pada Status Gizi Balita Di Kabupaten Simalungun," 2020.
- [2] 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)," *Journal of Information Technology Ampera*, vol. 5, no. 1, pp. 2774–2121, 2024, doi: 10.51519/journalita.v5i1.538.
- [3] H. Hairani, L. Nurhayati, and M. Innuddin, "Web-Based Application for Toddler Nutrition Classification Using C4.5 Algorithm," *International Journal of Engineering and Computer Science Applications (IJECSA)*, vol. 1, no. 2, pp. 77–82, Sep. 2022, doi: 10.30812/ijecsa.v1i2.2387.
- [4] S. Amalia Alhamid, B. Tiara Carolin, R. Lubis, P. Studi Kebidanan, F. Ilmu Kesehatan, and U. Nasional Jakarta, "STUDI MENGENAI STATUS GIZI BALITA," 2021.
- [5] Y. Maria Hawa Keytimu and M. Yustina Yantiana Guru, *PEMANFAATAN PANGAN LOKAL DENGAN METODE PIS-PK DALAM MENINGKATKAN STATUS GIZI BALITA*. 2025. [Online]. Available: <https://isbn.perpusnas.go.id/bo-penerbit/penerbit/isbn/data/view->
- [6] A. Aufa, A. Putri, and S. A. Rahmah, "IMPLEMENTASI DATA MINING DENGAN ALGORITMA K-MEANS CLUSTERING UNTUK ANALISIS BISNIS PADA PERUSAHAAN ASURANSI," *Jurnal Teknologi Informasi*, vol. 5, no. 1, 2024, doi: 10.46576/djtechno.
- [7] Z. Setiawan *et al.*, *Buku Ajar Data Mining*. PT.Sonpedia Publishing Indonesia, 2023.
- [8] M. R. Nahjan, N. Heryana, and A. Voutama, "IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL," 2023.
- [9] L. Qadrini, A. Seppewali, and A. Aina, "DECISION TREE DAN ADABOOST PADA KLASIFIKASI PENERIMA PROGRAM BANTUAN SOSIAL," *Jurnal Inovasi Penelitian*, vol. 02, Dec. 2021.
- [10] C. I. Ifaru *et al.*, "Penentuan Sparepart Genset Paling Sering Digunakan Pada Operator Indosat Ooredoo Hutchison Dengan Algoritma C4.5 Determination of the most

- frequently used generator spare parts at Indosat Ooredoo Hutchison Operator with C4.5 Algorithm,” 2024.
- [11] P. Rahayu, J. W., and S. L., *Pengantar Sistem Pakar: Teori dan Implementasi*. PT. Green Pustaka Indonesia, 2025.
- [12] P. Kurnia Illahi, A. Rina Viana, N. Fitria, M. Permata, and M. Y. Pratama, “SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Application of Decision Tree Algorithm and Linear Regression for Breast Cancer Classification Penerapan Algoritma Decision Tree aan Regresi Linear untuk Klasifikasi Kanker Payudara.” [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [13] O. P. Moerdyanto, I. Kadek, and D. Nuryana, “Prediksi Kelulusan Tepat Waktu Menggunakan Pendekatan Pohon Keputusan Algoritma Decision Tree,” *Journal of Informatics and Computer Science*, vol. 05, 2023.
- [14] J. Ilmu, K. Ruru, A. S. Sitio, and F. A. Sianturi, “PENINGKATAN KINERJA DECISION TREE C4.5 MELALUI PENDEKATAN MEAN DAN MEDIAN PADA DATA NUMERIK,” vol. 2, pp. 24–32, 2025, doi: 10.69688/jikr.v2i1.17.

