

BAB II

LANDASAN TEORI

2.1. Konsep Dasar Status Gizi Balita

Status gizi balita adalah salah satu indikator utama yang digunakan untuk mengevaluasi kondisi kesehatan dan pertumbuhan anak usia dini. Balita usia 1 sampai 5 merupakan kelompok umur yang mudah terkena masalah gizi [S. B. Bulkisah][3]. World Health Organization (WHO) mendefinisikan status gizi sebagai kondisi kesehatan yang dinilai berdasarkan parameter antropometri, yaitu usia, berat badan, dan tinggi badan. Data tersebut dibandingkan dengan standar pertumbuhan yang diterbitkan oleh WHO untuk menentukan kategori status gizi, seperti gizi buruk, gizi kurang, gizi baik, dan gizi lebih[6].

Menurut Kementerian Kesehatan Republik Indonesia (Kemenkes), status gizi yang buruk pada balita dapat menimbulkan berbagai masalah kesehatan jangka panjang. Masalah tersebut mencakup keterlambatan perkembangan kognitif, penurunan daya tahan tubuh, dan peningkatan risiko penyakit kronis di masa dewasa[7]. Sistem penilaian status gizi yang akurat sangat diperlukan untuk mendukung pengambilan keputusan berbasis bukti dalam intervensi kesehatan.

Berdasarkan standar WHO, status gizi balita diklasifikasikan ke dalam enam kategori:

1. **Gizi Baik:** Kondisi di mana kebutuhan nutrisi terpenuhi secara optimal sehingga mendukung pertumbuhan dan perkembangan balita.

2. **Gizi Kurang:** Kondisi di mana terjadi defisit nutrisi yang dapat menyebabkan hambatan dalam pertumbuhan.
3. **Gizi Buruk:** Kekurangan nutrisi kronis yang menyebabkan gangguan serius pada tumbuh kembang balita.
4. **Gizi Lebih:** Kondisi di mana asupan nutrisi berlebih menyebabkan berat badan di atas normal tetapi belum mencapai obesitas.
5. **Beresiko Gizi Lebih:** Kondisi di mana berat badan balita mulai menunjukkan kecenderungan menuju gizi lebih, namun belum mencapai kategori tersebut.
6. **Obesitas:** Akibat dari asupan nutrisi yang berlebihan sehingga berisiko menyebabkan kelebihan berat badan dan komplikasi kesehatan lainnya.

2.1.1 Indikator Status Gizi

Indikator antropometri merupakan alat utama dalam menilai status gizi balita. Pengukuran seperti usia, berat badan, dan tinggi badan, digunakan untuk menghitung Z-Score, yaitu ukuran statistik yang menunjukkan seberapa jauh nilai pengukuran individu dari rata-rata populasi sehat[R. Rusdiarti]. Sebagai contoh, balita dengan Z-Score di bawah -2 standar deviasi (SD) dikategorikan memiliki gizi kurang, sedangkan di bawah -3 SD menunjukkan kondisi stunting.

Tabel 2.1 Indikator Antropometri untuk Penilaian Status Gizi Balita

Kategori Status Gizi	Z-Score
Gizi Buruk (surverely wasted)	< -3 SD
Gizi kurang (wasted)	-3 SD sd < -2SD
Gizi baik (normal)	-2 SD sd +1SD
Berisiko gizi lebih (possible risk of overweight)	> + 1 SD sd +2 SD
Gizi lebih (overweight)	> + 2 SD sd + 3SD
Obesitas (obese)	> + 3 SD

Sumber : Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020[8]

2.1.2 Pentingnya Pemantauan Status Gizi

Pemantauan status gizi balita memiliki peran penting dalam mencegah dan mengatasi masalah gizi sejak dini. Pemantauan pertumbuhan dan perkembangan anak adalah bagian integral dari peran bidan dalam upaya mencegah stunting[9]. Pemantauan ini melibatkan pengukuran berat badan, dan tinggi badan, secara rutin untuk memastikan bahwa pertumbuhan anak sesuai dengan kurva pertumbuhan yang telah ditetapkan oleh World Health Organization (WHO). Kelebihan berat badan dan obesitas pada masa anak-anak juga meningkatkan risiko terhadap obesitas, penyakit tidak menular (PTM), kematian dini dan disabilitas di masa dewasa[10].

Program pemantauan seperti Posyandu di Indonesia menjadi wadah penting dalam mendeteksi dan menangani masalah gizi. Intervensi berbasis komunitas ini berkontribusi besar dalam menurunkan prevalensi malnutrisi melalui edukasi dan layanan kesehatan preventif.

2.2 Konsep Sistem Pengambilan Keputusan

Sistem pengambilan keputusan merupakan salah satu bagian dari sistem informasi berbasis computer[11]. Pendekatannya modern yang memanfaatkan informasi untuk mendukung proses pengambilan keputusan yang lebih tepat dan objektif. Pendekatan ini bertumpu pada integrasi data yang relevan dan representatif guna meminimalkan bias dalam proses pengambilan keputusan. Dalam bidang kesehatan, Perkembangan teknologi bertujuan untuk mengurangi permasalahan kesehatan pada masyarakat mengenai hal-hal klinis yang sangat perlu diperhatikan dan meningkatkan kualitas layanan dengan memberikan wawasan berbasis data kepada tenaga medis serta pemangku kebijakan[12]

Proses pengambilan keputusan umumnya memanfaatkan algoritma data mining untuk mengolah data dalam skala besar (big data) guna mengidentifikasi pola-pola yang relevan. Salah satu contohnya adalah penerapan algoritma C4.5 dalam klasifikasi data antropometri yang digunakan untuk menilai status gizi balita[13].

2.2.1 Keuntungan Sistem Pengambilan Keputusan Berbasis Data

Pendekatan berbasis data memiliki beberapa keuntungan utama:

1. Efisiensi

Sistem ini mampu memproses data dalam jumlah besar dengan lebih cepat dibandingkan metode manual, sehingga sangat cocok untuk implementasi di sektor yang memerlukan data dalam skala besar, seperti kesehatan masyarakat.

2. Akurasi Tinggi

Penggunaan algoritma canggih menjamin hasil yang lebih konsisten dan dapat diandalkan. Hal ini membantu tenaga kesehatan untuk membuat keputusan yang lebih tepat sasaran.

3. Kemudahan Interpretasi

Data yang dianalisis dapat divisualisasikan, misalnya dalam bentuk pohon keputusan atau diagram, untuk mempermudah interpretasi oleh pengguna akhir, seperti dokter atau tenaga kesehatan lainnya.

4. Fleksibilitas

Sistem ini dapat diterapkan di berbagai sektor, termasuk pendidikan, bisnis, dan kesehatan, yang menunjukkan bahwa pendekatan berbasis data bersifat adaptif.

2.2.2 Langkah-Langkah Sistem Pengambilan Keputusan

1. Identifikasi Masalah

Tahap awal dalam Sistem Pendukung Keputusan (SPK) adalah identifikasi masalah, yang mencakup pendefinisian tujuan keputusan, penentuan kriteria yang relevan, serta mempertimbangkan kendala atau batasan yang dapat memengaruhi proses pengambilan keputusan. Identifikasi masalah ini bertujuan untuk memastikan bahwa keputusan yang dihasilkan benar-benar sesuai dengan kebutuhan dan kondisi yang ada. Sebagai contoh, dalam bidang kesehatan, SPK dapat digunakan untuk menentukan status gizi balita berdasarkan indikator antropometri, seperti berat badan menurut usia (BB/U), tinggi badan menurut usia (TB/U), dan berat badan menurut tinggi badan (BB/TB).

2. Pengumpulan Data

Data yang dikumpulkan harus relevan dan representatif terhadap permasalahan yang dihadapi. Proses pengumpulan data dapat dilakukan melalui berbagai metode, seperti:

- a. Data primer (wawancara, kuesioner, observasi)
- b. Data sekunder (laporan kesehatan, jurnal penelitian, database)
- c. Data historis dan real-time.

3. Pemrosesan Data (Preprocessing Data)

Sebelum data digunakan dalam model SPK, perlu dilakukan beberapa tahap pemrosesan, seperti:

- a. Pembersihan Data: Menghilangkan data yang tidak valid atau memiliki nilai yang hilang
- b. Transformasi Data: Mengubah data ke dalam format yang dapat digunakan oleh model

4. Pemilihan Metode atau Model SPK

Berdasarkan jenis keputusan yang diambil, metode yang digunakan dalam SPK dapat bervariasi. Beberapa metode yang umum digunakan meliputi:

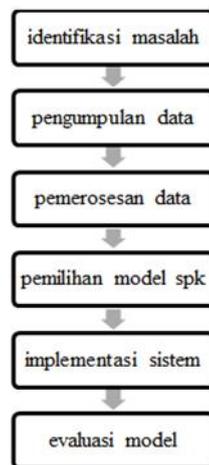
- a. Metode Berbasis Model: C4.5, Naïve Bayes, K-Nearest Neighbor (KNN)
- b. Metode Berbasis Multikriteria: Analytic Hierarchy Process (AHP), Simple Additive Weighting (SAW), Weighted Product (WP)
- c. Metode Berbasis Optimasi: Fuzzy Logic, Genetic Algorithm.

5. Implementasi Sistem

Sistem pengambilan keputusan berbasis algoritma C4.5 diimplementasikan menggunakan perangkat lunak seperti RapidMiner. Hasil klasifikasi divisualisasikan dalam bentuk pohon keputusan yang intuitif dan mudah diinterpretasi oleh pengguna.

6. Evaluasi Model

Evaluasi model menggunakan teknik confusion matrix. Confusion matrix sering dipakai dalam penelitian untuk menilai hasil serta mengevaluasi efektivitas suatu metode klasifikasi. Metrik kinerja seperti akurasi, presisi, recall, dan F1-score digunakan untuk memberikan gambaran menyeluruh mengenai performa sistem.



Gambar 2.1 Langkah-Langkah Sistem Pengambilan Keputusan

Keterangan : Gambar tersebut menunjukkan tahapan sistem pengambilan keputusan .

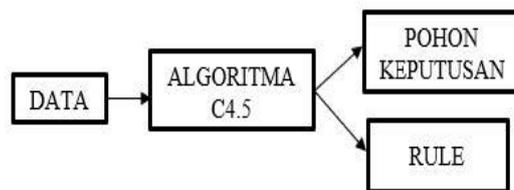
2.3 Algoritma C4.5

2.3.1 Dasar Teori Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma induksi pohon keputusan yang dikembangkan oleh J.Ross Quinlan[14]. Algoritma C4.5 merupakan kelompok algoritma decision tree. Algoritma C4.5 ini memiliki mekanisme

pruning untuk mengurangi kompleksitas pohon, yang dapat mengurangi risiko *overfitting* dan meningkatkan akurasi prediksi pada data baru. Algoritma C4.5 dilakukan dengan merancang atau membentuk pohon keputusan dengan memilih atribut pada variabel data sebagai akar (*node*), membuat cabang untuk setiap nilai, membagi kasus dalam cabang, dan mengulangi proses sampai semua cabang memiliki kelas yang sama[15]

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. tenaga kesehatan dan pengambil kebijakan. Selain itu, algoritma ini memiliki mekanisme pruning untuk mengurangi kompleksitas pohon, yang dapat mengurangi risiko *overfitting* dan meningkatkan akurasi prediksi pada data baru[16][17].



Gambar 2.2 Alur Pohon Keputusan Algoritma C4.5

Keterangan : Diagram tersebut menggambarkan alur proses di mana data diproses menggunakan algoritma C4.5 untuk menghasilkan pohon keputusan dan aturan (aturan) yang digunakan dalam sistem pengambilan keputusan.

2.3.2 Karakteristik Algoritma C4.5

Beberapa karakteristik utama dari algoritma C4.5 yang membedakannya dari algoritma sebelumnya adalah:

- a. Mampu menangani atribut kontinu dan diskrit

C4.5 dapat bekerja dengan data numerik dan kategorikal, dengan melakukan klasifikasi berdasarkan ambang batas optimal untuk atribut numerik.

- b. Mengirimkan data yang hilang

Jika terdapat nilai yang hilang, algoritma tetap dapat bekerja dengan menghitung probabilitas berdasarkan data yang tersedia.

- c. Melakukan pemangkasan untuk mengurangi kompleksitas pohon

Pruning adalah teknik untuk memangkas cabang pohon yang tidak signifikan guna menghindari overfitting terhadap data pelatihan.

- d. Dapat menghasilkan aturan keputusan (if-then)

Setelah membangun pohon keputusan, algoritma dapat mengubah sekumpulan aturan klasifikasi yang lebih mudah dipahami.

2.3.3 Algoritma Dalam Berbagai Bidang

Algoritma C4.5 telah diterapkan di berbagai domain, menunjukkan fleksibilitas dan efektivitasnya. Beberapa penelitian terdahulu yang relevan mencakup:

1. Prediksi Loyalitas Pelanggan

Penelitian oleh [18] menggunakan algoritma C4.5 untuk memprediksi

loyalitas pelanggan di PT. Pegadaian. Hasil penelitian menunjukkan tingkat akurasi sebesar 97,5%, yang menegaskan efektivitas algoritma dalam mengidentifikasi pola perilaku pelanggan. Keberhasilan ini memungkinkan perusahaan untuk menyusun strategi pemasaran yang lebih terarah.

2. Prediksi Indeks Pembangunan Manusia

Penelitian oleh [19] menerapkan algoritma C4.5 untuk menganalisis dan memprediksi Indeks Pembangunan Manusia di Indonesia. Penelitian ini menunjukkan bagaimana algoritma mampu mengelola data sosial-ekonomi yang kompleks serta menghasilkan model yang dapat diinterpretasikan secara jelas.

3. Klasifikasi Kesiapan Siswa untuk Ujian Nasional

[20] Menggabungkan algoritma C4.5 dengan pendekatan pengoptimalan particle swarm optimization (PSO) untuk meningkatkan akurasi klasifikasi kesiapan siswa menghadapi ujian nasional. Penelitian ini membantu mengidentifikasi siswa yang memerlukan dukungan tambahan, sehingga mendukung proses pembelajaran yang lebih efektif.

4. Aplikasi di Bidang Kesehatan

Penelitian [21] menunjukkan kontribusi signifikan algoritma C4.5 dalam aplikasi perawatan kesehatan, seperti diagnosis penyakit dan analisis data medis. Pohon keputusan yang dihasilkan memberikan wawasan yang dapat ditindaklanjuti oleh tenaga medis untuk pengambilan keputusan yang lebih baik.

2.3.4 Proses Kerja Algoritma C4.5

Proses kerja algoritma C4.5 terdiri atas langkah-langkah berikut:

1. Menghitung Entropy Dataset Awal

Entropy digunakan untuk mengukur tingkat ketidakteraturan atau ketidakpastian dalam dataset.

2. Menghitung Information Gain Setiap Atribut

Information gain mengukur pengurangan entropy setelah data dibagi berdasarkan atribut tertentu.

3. Memilih Atribut dengan Information Gain Tertinggi

Atribut dengan nilai information gain tertinggi dipilih sebagai simpul akar pohon.

4. Mengulangi Proses hingga Semua Data Terklasifikasi

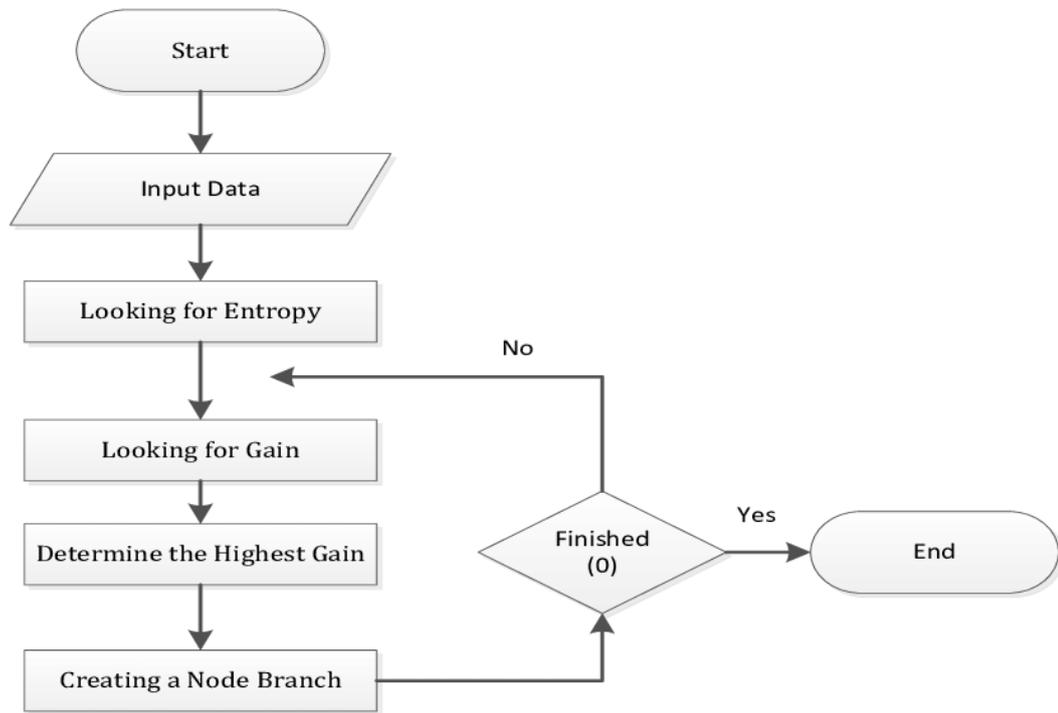
Langkah ini dilakukan secara rekursif hingga semua data terklasifikasi atau kriteria penghentian, seperti jumlah data minimum pada node, terpenuhi.

5. Pohon Keputusan: Melalui pembagian data menjadi subset berdasarkan atribut yang memiliki Information Gain tertinggi.

6. Pembuatan pohon keputusan dan Rule.

7. Melakukan Pruning

Setelah pohon selesai dibangun, dilakukan pruning untuk menghapus cabang yang tidak signifikan, sehingga mengurangi risiko overfitting.



Gambar 2.3 Flowchart Proses Kerja Algoritma C4.5

Keterangan: Diagram ini menggambarkan langkah-langkah utama dalam penerapan algoritma C4.5, mulai dari preprocessing data hingga evaluasi dan optimasi model.

2.3.5 Formula Matematika Algoritma C4.5

1. Entropy

Entropy digunakan untuk mengukur tingkat ketidakaturan atau ketidakpastian dalam dataset. Nilai entropy dihitung menggunakan formula berikut:

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

Keterangan:

S = Himpunan kasus

n = Jumlah partisi S

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

2. Information Gain Setiap Atribut

Information gain mengukur pengurangan entropy setelah data dibagi berdasarkan atribut tertentu. Formula untuk information gain adalah:

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

Keterangan:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

$|S_i|$: Jumlah kasus pada partisi ke- i

$|S|$: Jumlah kasus pada S

2.3.6 Kelebihan dan Kekurangan Algoritma C4.5

Tabel 2.2 Perbandingan Keunggulan dan Kekurangan Algoritma C4.5

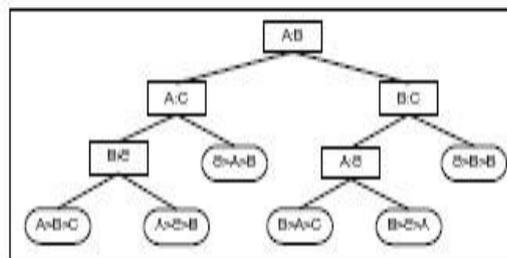
Aspek	Keunggulan	Kekurangan
Penanganan Data	Mampu menangani atribut numerik, kategorikal, dan nilai	Memerlukan waktu pemrosesan yang lama pada dataset besar
Interpretabilitas	Hilang Pohon keputusan	Rentan terhadap atribut
Kemampuan Generalisasi	mudah dipahami Menggunakan pruning untuk mencegah overfitting	Irrelevan Membutuhkan seleksi atribut jika banyak atribut yang tidak relevan

2.4 Pohon Keputusan (*Decision Tree*)

Pohon keputusan adalah suatu gambaran pemodelan dari suatu persoalan yang terdiri dari serangkaian keputusan yang mengarah kesolusi [22]. Tiap simpul dalam menyatakan keputusan dan daun menyatakan solusi. Proses pembangunan pohon keputusan pada C4.5 dimulai dengan menghitung gain informasi atau gain rasio setiap atribut yang ada pada data. Atribut dengan gain informasi atau gain rasio tertinggi akan dipilih sebagai atribut terbaik untuk membagi data. Proses ini dilakukan secara berulang-ulang hingga terbentuk pohon keputusan yang optimal.

Pohon keputusan memiliki *node* pohon yang merepresentasikan atribut yang telah diuji dan setiap cabangnya merupakan suatu pembagian hasil uji serta *node* daun (*leaf*) merepresentasikan kelompok kelas tertentu (Han & Kamber,

2006). Level *node* teratas dari sebuah pohon keputusan adalah *node* akar (*root*) yang biasanya berupa atribut yang memiliki pengaruh paling besar pada suatu kelas tertentu. Konsep dasar dari pohon keputusan adalah mengubah data menjadi model pohon keputusan, kemudian mengubah model pohon menjadi *rule* dan menyederhanakan *rule*. Data dalam pohon keputusan.



Gambar 2.4 Contoh Pohon Keputusan

Keterangan : Gambar ini menunjukkan struktur pohon keputusan yang dibentuk berdasarkan atribut dan nilai tertentu.

Pohon keputusan pada gambar 2.4 diatas dibaca dari atas ke bawah. Simpul paling atas pada pohon ini adalah simpul akar. Simpul yang ditandai dengan tanda kotak disimpul tersebut dinamakan simpul keputusan. Cabang-cabang yang mengarah kekanan dan kekiri dari cabang keputusam mempresentasikan kumpulan dari alternatif keputusan yang bisa diambil. Hanya satu keputusan yang dapat diambil dalam suatu waktu.

2.5 Pengolahan Aplikasi Rapidminer

2.5.1 RapidMiner

RapidMiner adalah perangkat lunak berbasis data mining yang dikembangkan oleh *Dr. Markus Hofmann* dari *Institute of Technology Blanchardstown* dan *Ralf Klinkenberg* dari *rapid-i.com*. Sebagai perangkat lunak open-source, RapidMiner dirancang menggunakan bahasa pemrograman Java dan dilisensikan di bawah GNU Public License, sehingga dapat dijalankan pada berbagai sistem operasi. Perangkat lunak ini dilengkapi dengan antarmuka grafis (Graphical User Interface/GUI) yang intuitif, memungkinkan pengguna untuk melakukan analisis data secara efisien tanpa memerlukan kemampuan pemrograman yang mendalam[23].

RapidMiner juga digunakan sebagai solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi[24]

RapidMiner memiliki kemampuan untuk menangani preprocessing data, analisis visualisasi, penerapan algoritma, hingga evaluasi model dengan akurasi yang tinggi. Fitur-fitur tersebut menjadikannya alat yang fleksibel dan andal

dalam berbagai jenis penelitian, termasuk yang melibatkan data besar dan kompleks. RapidMiner juga menyediakan beragam metode, termasuk *klasifikasi*, *klustering*, dan *asosiatif*. Selain itu, jika terdapat algoritma atau model yang belum tersedia, pengguna dapat menambahkan modul tambahan, karena sifatnya yang open source memungkinkan perangkat lunak ini untuk dikembangkan lebih lanjut oleh siapa saja.

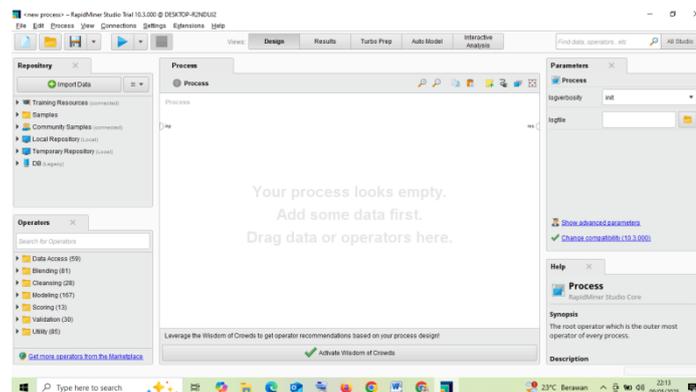
2.5.2 Keunggulan Rapidminer

RapidMiner memiliki sejumlah keunggulan utama yang menjadikannya salah satu perangkat lunak data mining yang paling banyak digunakan, yaitu:

1. Antarmuka Pengguna yang Intuitif : Antarmuka pengguna RapidMiner dirancang untuk mempermudah proses analisis data. Tidak seperti perangkat lunak berbasis pemrograman, RapidMiner menyediakan drag-and-drop interface yang memungkinkan pengguna membangun alur analisis tanpa harus menulis kode program. [25] Menyebutkan bahwa fitur antarmuka grafis ini sangat membantu dalam proses analisis prediksi harga menggunakan regresi linier, memberikan efisiensi waktu serta keakuratan dalam penelitian yang dilakukan.
2. Dukungan Algoritma yang Beragam : RapidMiner menyediakan berbagai algoritma terutama Algoritma C4.5 . Kemampuan ini memungkinkan fleksibilitas dalam memilih algoritma yang paling sesuai untuk berbagai jenis data dan analisis. Salah satu fitur unggulan RapidMiner adalah kemampuannya dalam tahap preprocessing, termasuk normalisasi data, deteksi outlier, serta pembersihan data. Penelitian oleh [26] menunjukkan

bahwa preprocessing data menggunakan RapidMiner meningkatkan akurasi hasil clustering dalam mengelompokkan data penyebaran HIV di wilayah Jawa Barat .

3. Validasi Model yang Akurat : Teknik validasi model yang didukung oleh RapidMiner, seperti K-Fold Cross-Validation dan visualisasi Confusion Matrix, memberikan keandalan tinggi dalam mengevaluasi performa model. Validasi ini juga membantu mengidentifikasi overfitting, sehingga model dapat bekerja secara optimal pada data baru.
4. Kemampuan Visualisasi yang Mendalam : RapidMiner menyediakan alat visualisasi data untuk menganalisis pola yang kompleks. [27] mencatat bahwa fitur ini meningkatkan pemahaman tentang pola distribusi data yang rumit, seperti dalam penelitian analisis penjualan.



Gambar 2.5 Aplikasi Rapid Miner

Keterangan : Gambar tersebut menunjukkan tampilan aplikasi RapidMiner Studio, sebuah perangkat lunak berbasis GUI yang digunakan untuk

pemrosesan data, analisis prediktif, dan pembuatan model machine learning tanpa perlu menulis kode secara manual.

2.6 Perhitungan Akurasi

Perhitungan akurasi dilakukan untuk mengevaluasi sejauh mana tingkat ketepatan sistem dalam mengklasifikasikan data pada studi kasus yang menjadi fokus penelitian ini. Akurasi mencerminkan perbandingan antara jumlah prediksi yang tepat dengan total keseluruhan data uji yang digunakan. Nilai akurasi diperoleh dengan membandingkan jumlah data uji yang diprediksi benar sesuai kelas sebenarnya terhadap jumlah total data uji, kemudian dikalikan dengan 100 persen. Persamaan perhitungan akurasi dapat dituliskan sebagai berikut:

$$Akurasi = \frac{\text{jumlah data uji yang diklasifikasikan benar}}{\text{jumlah data yang diuji}} \times 100\%$$

Jumlah data uji yang diklasifikasikan dengan benar mengacu pada data yang hasil prediksinya sesuai dengan label atau kelas aslinya. Sementara itu, jumlah total data uji mencakup seluruh data yang digunakan dalam proses pengujian model klasifikasi.

Untuk mendukung perhitungan akurasi dan mengevaluasi performa model secara lebih rinci, digunakan Confusion Matrix. Confusion Matrix merupakan sebuah tabel yang menggambarkan kinerja algoritma klasifikasi dengan membandingkan antara hasil prediksi dan nilai aktual dari kelas target.

2.6 Penelitian Terdahulu

Tabel 2.3 Penelitian Terdahulu

No	Nama Penelitian	Judul Penelitian	Tahun	Hasil Penelitian
1	Revaldo[28]	Penerapan Algoritma C4.5 Untuk Prediksi Anak Stunting Di Kota Pagar Alam	2024	Sistem berbasis prediksi C4.5 yang dikembangkan memiliki akurasi 88,20% dan diuji menggunakan metode Black Box Testing. Sistem dapat mengidentifikasi anak berisiko stunting dengan lebih cepat.
2	Mahfuz[29]	Penerapan Algoritma C4.5 Dalam Mengklasifikasi Status Gizi Balita Pada Posyandu Desa Dames Damai Kabupaten Lombok Timur	2022	Menggunakan algoritma C4.5, penelitian mencatat akurasi 97%, menunjukkan efisiensi dalam klasifikasi data gizi balita berdasarkan antropometri.
3	Abdulqader[21]	Penerapan Pohon Keputusan untuk Diagnostik Kesehatan	2024	C4.5 menguraikan data medis kompleks untuk mendiagnosis pasien
4	Nuraisana[30]	Implementasi Algoritma C4.5 dalam Mengklasifikasi Status Gizi Balita pada Posyandu Desa Sekip Lubuk Pakam	2024	Penelitian ini berhasil mengklasifikasikan status gizi balita menggunakan algoritma C4.5 dengan tingkat akurasi yang baik.

No	Nama Penelitian	Judul Penelitian	Tahun	Hasil Penelitian
5	W.Wahyudi[5]	Optimasi Klasifikasi Status Gizi Balita Berdasarkan Indeks Antropometri Menggunakan Algoritma C4.5 Adaboost Classification	2019	Akurasi prediksi mencapai 90,23%
6	Abdurrahman Naufal [31]	Klasifikasi Keluarga Berisiko Stunting Menggunakan Algoritma C4.5	2024	Menggunakan Algoritma C4.5 untuk mengklasifikasikan keluarga berisiko stunting dengan akurasi rata-rata 98,4%.
7	Agustiena Merdekawati [32]	Sistem Pendukung Keputusan Penerimaan Beasiswa Menggunakan Algoritma C4.5	2018	Algoritma C4.5 diterapkan untuk seleksi penerimaan beasiswa dengan akurasi 69,73%.
8	Abiyyu Ramadhan [33]	Sistem Pendukung Keputusan Evaluasi Problematika Pembelajaran Daring dengan Algoritma C4.5	2022	Algoritma C4.5 digunakan untuk menentukan metode pembelajaran daring terbaik berdasarkan data guru dan siswa.
9	Windy Julianto[34]	Algoritma C4.5 untuk Penilaian Kinerja Karyawan	2014	C4.5 digunakan untuk evaluasi kinerja karyawan berdasarkan berbagai parameter dengan akurasi tinggi.

No	Nama Penelitian	Judul Penelitian	Tahun	Hasil Penelitian
10	Saefudin & Sri Lestari[35]	Sistem Pendukung Keputusan untuk Penilaian Kinerja Karyawan PT. Mulya Spindo Mills Menggunakan Algoritma C4.5	2015	Menggunakan algoritma C4.5 untuk menentukan evaluasi kinerja karyawan secara objektif.
11	Liliana Swastina [36]	Penerapan Algoritma C4.5 untuk Penentuan Jurusan Mahasiswa	2013	Algoritma C4.5 diterapkan dalam pemilihan jurusan mahasiswa dengan akurasi 93,31%.
12	Aghniya Nadira[37]	Implementasi Data Mining dalam Mengidentifikasi Faktor Pasien yang Berpotensi Mengalami Obesitas Menggunakan Algoritma C4.5	2024	Algoritma C4.5 diterapkan untuk mengidentifikasi faktor risiko obesitas dengan akurasi 98,8%.
13	Muhamad Iqbal[38]	Sistem Pendukung Keputusan Menentukan Hasil Budidaya Udang Vaname dengan Metode Algoritma C4.5	2020	Algoritma C4.5 diterapkan dalam klasifikasi faktor yang memengaruhi hasil budidaya udang dengan akurasi 76,47%.
14	Marini [39]	Penerapan Data Mining Klasifikasi Gizi Bayi dengan Algoritma Decision Tree C4.5	2022	Algoritma C4.5 diterapkan untuk klasifikasi status gizi bayi dan balita di Posyandu Gunung Maligas. Faktor yang paling berpengaruh terhadap gizi buruk dianalisis menggunakan data mining.

No	Nama Penelitian	Judul Penelitian	Tahun	Hasil Penelitian
15	Tukino [15]	Penerapan Metode Algoritma C4.5 dalam Penilaian Kelayakan Pemberian Kredit Kepada Mitra Usaha	2023	C4.5 diterapkan dalam sistem klasifikasi kelayakan kredit untuk mitra usaha PT Arita Prima Sukses dan meningkatkan akurasi keputusan.