

Analisis Keadaan Stunting pada Kelompok Balita di Kecamatan Tukdana dengan Pendekatan *Decision Trees*

Asep Budiyanto¹, Dodi Solihudin², Ryan Hamonangan³, Cep Lukman Rohmat⁴,
Ade Rizki Rinaldi⁵

^{1,2,3}Jurusan Teknik Informatika, STMIK IKMI Cirebon

^{4,5}Jurusan Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Email: ¹emailasep07@gmail.com, ²dodisolihudin@ikmi.ac.id, ³ryanhamonangan@ikmi.ac.id,
⁴ceplukman@ikmi.ac.id, ⁵aderizki@ikmi.ac.id

Abstrak

Dampak *stunting* pada bayi merupakan *parameter* penting untuk menilai kesehatan dan kesejahteraan anak dalam suatu daerah. *Stunting*, seringkali dipicu oleh faktor demografis dan kesehatan, memiliki implikasi serius terhadap pertumbuhan fisik dan kognitif anak. Penelitian ini bertujuan untuk memahami dampak faktor-faktor demografis dan kesehatan terhadap *stunting* anak-anak di Kecamatan Tukdana, Kabupaten Indramayu. Melalui analisis data, faktor-faktor seperti usia ibu, akses air bersih, fasilitas sanitasi, serta status berat dan panjang badan bayi diidentifikasi sebagai kontributor signifikan terhadap *stunting*. Metode *Decision Trees* digunakan untuk mengidentifikasi faktor-faktor yang berperan dalam *stunting* pada bayi, dengan tingkat akurasi sebesar 95.43%. Implikasi dari penelitian ini meliputi perencanaan intervensi yang lebih efektif untuk menangani *stunting*, baik di Kecamatan Tukdana maupun di wilayah-wilayah serupa di Indonesia. Meskipun mayoritas bayi di Kecamatan Tukdana memiliki status gizi yang baik, tetap diperlukan pemantauan dan upaya pencegahan lanjutan untuk memastikan kesejahteraan gizi yang optimal bagi mereka. Kesimpulannya, penelitian ini menyoroti pentingnya identifikasi faktor-faktor penyebab *stunting* pada bayi di Kecamatan Tukdana, sebagai dasar bagi perencanaan intervensi yang lebih efektif.

Kata kunci: *Stunting, Decision Trees, Analisis Data Kesehatan, Faktor Determinan, Gizi, Kecamatan Tukdana, Indonesia*

Abstract

The impact of stunting on babies is an important parameter for assessing the health and welfare of children in an area. Stunting, often triggered by demographic and health factors, has serious implications for children's physical and cognitive growth. This research aims to understand the impact of demographic and health factors on stunting in children in Tukdana District, Indramayu Regency. Through data analysis, factors such as maternal age, access to clean water, sanitation facilities, and baby weight and length status were identified as significant contributors to stunting. The Decision Trees method was used to identify factors that play a role in stunting in babies, with an accuracy rate of 95.43%. The implications of this research include planning more effective interventions to deal with stunting, both in Tukdana District and in similar areas in Indonesia. Even though the majority of babies in Tukdana District have good nutritional status, further monitoring and prevention efforts are still needed to ensure optimal nutritional well-being for them. In conclusion, this research highlights the importance of identifying factors that cause stunting in infants in Tukdana District, as a basis for planning more effective interventions.

Keywords: *Stunting, Decision Trees, Health Data Analysis, Determinant Factors, Nutrition, Tukdana District, Indonesia*

PENDAHULUAN

Perkembangan pesat di bidang Informatika memiliki dampak signifikan dalam pengelolaan dan analisis data kesehatan di Kabupaten Indramayu, Jawa Barat (Solehuddin dkk., 2022). Penelitian ini menerapkan algoritma C4.5 dan metode *Decision Trees* untuk menganalisis faktor-faktor yang memengaruhi

status *stunting* pada bayi di Kecamatan Tukdana, Kabupaten Indramayu pada tahun 2023, dengan fokus pada permasalahan *stunting* (Alawiah dkk., 2021). Tantangan di bidang Informatika, termasuk keamanan dan privasi data yang lebih ketat, serta kekurangan penelitian lokal, menjadi dasar penelitian ini (Ainurrohmah dan Wiyanti, 2023). Metode penelitian kuantitatif dengan

pendekatan analisis data CRISP-DM dan penerapan algoritma C4.5 digunakan untuk memberikan wawasan mendalam tentang faktor-faktor yang memengaruhi *stunting*, mendukung perencanaan program intervensi, dan berkontribusi pada pengembangan teknologi informasi di kesehatan masyarakat (Muttaqin dkk., 2021).

Penelitian ini berbeda dari penelitian-penelitian sebelumnya dengan penekanan pada penerapan *tuning parameter* pada algoritma *Decision Trees* untuk mencari nilai akurasi yang optimal. Pendekatan kuantitatif dengan pendekatan analisis data CRISP-DM juga digunakan untuk memberikan wawasan yang lebih mendalam tentang faktor-faktor yang memengaruhi *stunting*, mendukung perencanaan program intervensi yang lebih efektif, dan berkontribusi pada pengembangan teknologi informasi di bidang kesehatan masyarakat (Lestari & Amalia, 2023).

TINJAUAN PUSTAKA

2.1 Klasifikasi Algoritma C4.5

Algoritma C4.5, yang dikembangkan oleh Ross Quinlan, merupakan algoritma klasifikasi *data mining* yang populer, menggunakan struktur pohon keputusan serupa dengan *Decision Trees* (Wahono dan Riana, 2020). C4.5 memiliki fleksibilitas dalam menangani beragam jenis atribut dan dapat diterapkan pada masalah klasifikasi maupun regresi. Keunggulannya terletak pada kemampuannya menghadapi data kompleks dan menghasilkan model yang mudah diinterpretasi. Penerapan C4.5 dalam penelitian ini bertujuan menganalisis faktor-faktor yang memengaruhi status *stunting* pada kelompok baduta di Kecamatan Tukdana, dengan harapan mengungkap pola-pola signifikan dalam data kesehatan masyarakat dan memberikan kontribusi pada pemahaman permasalahan *stunting* di wilayah tersebut.

Persamaan (1), Rumus untuk mengestimasi nilai *Entropy* adalah 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 = proporsi S_i terhadap S

Persamaan (2), Rumus untuk mengestimasi nilai *Gain* 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 = fitur

n = jumlah partisi atribut A

$|S_i|$ = proporsi S_i terhadap S

$|S|$ = jumlah kasus dalam S

2.2 Decision Trees

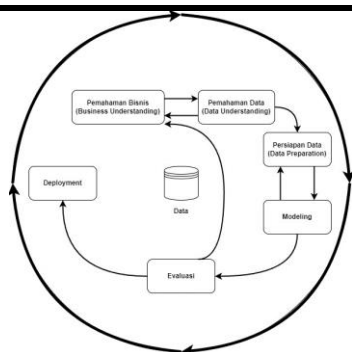
Decision Trees, atau pohon keputusan, adalah metode *data mining* yang menggunakan struktur pohon untuk menggambarkan serangkaian keputusan dan konsekuensinya, memfasilitasi analisis klasifikasi dan regresi (Hana, 2020). Setiap simpul dalam pohon mencerminkan keputusan berdasarkan atribut tertentu, memberikan representasi visual yang jelas terkait proses pengambilan keputusan. Dalam penelitian ini, penggunaan *Decision Trees* diharapkan membantu mengidentifikasi faktor-faktor yang memengaruhi status *stunting* pada kelompok baduta di Kecamatan Tukdana, memberikan wawasan berharga dalam konteks kesehatan masyarakat (Surya dkk., 2021).

2.3 CRISP-DM (Cross-Industry Standard Process for Data Mining)

CRISP-DM, sebuah kerangka kerja standar dalam *data mining*, terdiri dari enam tahap utama: pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran (Abidin dkk., 2023). Dalam penelitian ini, penggunaan CRISP-DM diharapkan memberikan pedoman yang terperinci dalam pengembangan model analisis data *stunting* pada kelompok baduta di Kecamatan Tukdana.

METODE PENELITIAN

Dalam konteks kerangka kerja penelitian CRISP-DM (*Cross-Industry Standard Process for Data Mining*), peneliti menggunakan berbagai strategi analisis data yang melibatkan:



Gambar 1. Rangkaian Proses CRISP-DM
(Wayan Wardani dkk., 2022)

CRISP-DM menampilkan kerangka metodologi terintegrasi dengan cermat dalam penelitian ini, memberikan panduan untuk serangkaian strategi analisis data yang mencakup tahap-tahap esensial seperti pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, hingga implementasi hasil analisis (Lestari dan Amalia, 2023). Pendekatan ini dirancang untuk memberikan dasar metodologis yang kuat dan terstruktur dalam menjalankan proses analisis data secara optimal dan terarah (Suhanda dkk., 2020).

3.1 Pemahaman Bisnis

Pada fase "Pemahaman Bisnis," tujuan utama penelitian adalah menganalisis status *stunting* pada kelompok baduta di Kecamatan Tukdana. Fokusnya adalah memberikan sumbangan signifikan pada pemahaman kesehatan masyarakat di wilayah tersebut, dengan penekanan pada faktor-faktor yang memengaruhi status *stunting*. Penelitian ini dapat menjadi dasar untuk pembentukan kebijakan pencegahan dan perbaikan status gizi baduta.

3.2 Pemahaman Data

Dalam tahap "Pemahaman Data," penelitian ini menyelidiki dataset bayi dari sumber data resmi <https://elsimil.bkkbn.go.id/>. Tujuannya adalah memahami secara menyeluruh sumber data tersebut dan fokus pada struktur serta karakteristik data. Analisis mendalam terhadap data ini menjadi kunci dalam merancang metode yang sesuai untuk hasil yang relevan dalam analisis status *stunting* pada kelompok bayi di Kecamatan Tukdana.

3.3 Persiapan Data

Pada tahap "Persiapan Data," penelitian ini melibatkan serangkaian proses pengolahan

dataset untuk memenuhi kebutuhan analisis. Ini mencakup penanganan nilai-nilai kosong, normalisasi data, dan langkah-langkah *preprocessing* lainnya untuk memastikan integritas dan kualitas dataset. Fokus utama pada variabel-variabel terkait faktor-faktor yang berpengaruh terhadap status *stunting* pada bayi di Kecamatan Tukdana memastikan data memiliki pemahaman yang optimal terkait dengan fokus penelitian.

3.4 Pemodelan

Dalam tahap "Pemodelan," penelitian ini akan menerapkan algoritma C4.5 dengan tujuan mengidentifikasi pola dan faktor kunci yang terkait dengan status *stunting* pada kelompok baduta di Kecamatan Tukdana. Penerapan algoritma ini diharapkan mampu menghasilkan model yang tidak hanya dapat mengklasifikasikan status *stunting* secara akurat, tetapi juga memberikan wawasan mendalam mengenai tantangan kesehatan masyarakat di wilayah tersebut. Model ini diharapkan dapat memberikan kontribusi berharga untuk merancang intervensi dan kebijakan yang lebih efektif dalam upaya pencegahan dan perbaikan status gizi baduta di Kecamatan Tukdana.

3.5 Evaluasi

Dalam tahap "Evaluasi," penelitian ini akan melakukan penilaian terhadap kinerja model dalam mengklasifikasikan status *stunting* pada kelompok baduta di Kecamatan Tukdana. Evaluasi tidak hanya akan mempertimbangkan tingkat akurasi model, melainkan juga akan mengevaluasi sejauh mana model dapat memberikan informasi yang bermanfaat untuk pengambilan keputusan terkait permasalahan kesehatan masyarakat. Evaluasi secara menyeluruh ini diharapkan dapat memberikan gambaran yang akurat mengenai keandalan dan relevansi model dalam konteks penelitian ini.

3.6 Implementasi

Dalam tahap "Implementasi," penelitian ini akan merancang rencana efektif untuk menyampaikan hasil analisis dengan tujuan mendukung pengambilan keputusan terkait status gizi baduta di Kecamatan Tukdana. Selain itu, perhatian khusus akan diberikan untuk memastikan bahwa hasil penelitian dapat diaplikasikan secara praktis dan memberikan dampak positif pada upaya pencegahan dan perbaikan status *stunting* di wilayah tersebut. Rencana ini bertujuan untuk memastikan bahwa

informasi yang dihasilkan dari penelitian ini dapat memberikan kontribusi nyata dalam pemecahan masalah kesehatan masyarakat.

HASIL DAN PEMBAHASAN

4.1 Pemahaman Bisnis

BKKBN, lembaga yang berkomitmen pada kesejahteraan masyarakat melalui kebijakan keluarga berencana, memiliki kepentingan khusus dalam menerapkan algoritma C4.5 untuk menganalisis status *stunting* pada kelompok baduta. Dengan pemahaman lebih mendalam tentang faktor-faktor yang memengaruhi status *stunting*, BKKBN berharap dapat merancang intervensi yang lebih efektif dan tepat sasaran. Fokus utama penelitian ini adalah meningkatkan kesehatan dan kesejahteraan kelompok baduta di wilayah tertentu, dengan tujuan mengidentifikasi faktor-faktor kritis yang memengaruhi status *stunting* melalui penerapan algoritma C4.5.

4.2 Pemahaman Data

Dataset untuk analisis status *stunting* pada kelompok baduta diperoleh dari situs resmi BKKBN, mencakup periode Januari–Oktober 2023, dengan 43 atribut dan 1.768 entri data. Atribut mencakup informasi seperti NIK, Nama Ibu, Usia Ibu, No. HP, Provinsi, Kabupaten/Kota, Kecamatan, Alamat, Pemberian ASI Eksklusif, Akses Air Minum, dan lainnya.

4.3 Persiapan Data

4.3.1 Pembersihan Data

Pembersihan data dilakukan untuk menjaga integritas data dengan mengidentifikasi dan menangani nilai yang hilang atau *noise* selama proses input. Peneliti menggunakan rumus Excel (*counta*) untuk menghitung jumlah nilai pada setiap atribut dan memverifikasinya agar sejajar dengan jumlah rekam dalam dataset. Identifikasi nilai *noise* memerlukan pemahaman mendalam terhadap setiap atribut, dengan pertimbangan bahwa nilai yang tidak sesuai dengan konteks atribut lainnya dianggap sebagai *noise*. Saat menghadapi nilai yang hilang atau *noise* pada beberapa atribut, peneliti memutuskan untuk menghapus baris data yang mengandung nilai tersebut. Keputusan ini diambil karena jumlah data yang hilang relatif kecil, sehingga penghapusan tersebut dianggap tidak signifikan secara substansial terhadap

analisis keseluruhan, sambil meningkatkan kualitas dan keakuratan data yang digunakan dalam analisis.

4.3.2 Transformation Data

Pada tahap ini, peneliti akan memilih 5 atribut utama dari total 43 atribut dalam data baduta yang dianggap memiliki hubungan signifikan dengan faktor-faktor yang memengaruhi status *stunting*. Atribut yang terpilih, yaitu USIA IBU, AKSES AIR MINUM, TEMPAT BUANG AIR BESAR, BERAT BADAN BAYI, dan PANJANG BADAN BAYI, akan difokuskan dalam analisis untuk menentukan status *stunting* pada kelompok baduta.

Atribut STATUS BB/PB dan STATUS *STUNTING* akan menjadi atribut utama dalam proses analisis klasifikasi. Proses ini melibatkan pengelompokan berbagai kategori berdasarkan jangka waktu data dari Januari hingga Oktober 2023. Pendekatan ini dilakukan untuk menghasilkan informasi terkini dan relevan mengenai status gizi baduta.

Analisis akan melibatkan atribut STATUS USIA IBU, AKSES AIR MINUM, TEMPAT BUANG AIR BESAR, STATUS BB/U, dan STATUS PB/U, karena atribut tersebut memiliki peran signifikan dalam pemahaman status *stunting*, memberikan informasi penting tentang kondisi gizi dan pertumbuhan bayi.

4.4 Pemodelan

4.4.1 Menggunakan Algoritma C4.5

Dalam tahap ini, peneliti membentuk model klasifikasi menggunakan metode *decision tree* C4.5, sebuah pendekatan supervised learning yang membutuhkan data latih sebagai dasar pembelajaran. Penggunaan metode ini dijelaskan secara rinci dan diakui sebagai panduan dalam menentukan klasifikasi status *stunting* pada penelitian ini.

Tabel 1. Ilustrasi Penerapan Algoritma C4.5

	Jumlah Gizi Baik	Gizi Kurang
Lebih	Obesitas Gizi Buruk	Gizi Kurang
Info	Tidak Terdeteksi	<i>Entropy Gain Split</i>
Total	1768	1664
	13	65
	13	2
		0,418416824

STATUS USIA IBU

	0,418416824	1,138156404		
0,367626824				
Ideal	1153	1086	44	6
9	8	0	0	
Terlalu Muda	88	81	2	
2	0	1	2	0
Terlalu Tua	527	497	19	
3	4	4	0	0

AKSES AIR MINUM

	0,274235784	0,603863383		
0,454135475				
Beresiko261	216	25	5	
8	5	2	0,976674632	
Tidak Beresiko	1507	1448	40	
6	5	8	0	0

TEMPAT BUANG AIR BESAR

	0,274235784			
0,603863383	0,454135475			
Beresiko261	216	25	5	
8	5	2	0,976674632	
Tidak Beresiko	1507	1448	40	
6	5	8	0	0

STATUS BB/U

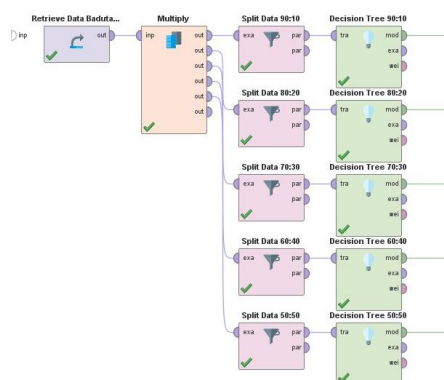
	0,330358964	0,439080244		
0,752388586				
Berat Kurang	31	21	5	
0	3	2	0	0
Berat Lebih	86	43	27	
7	5	2	2	
1,810305779				
Berat Normal	1645	1595	32	
4	5	9	0	0
Tidak Terdeteksi	6	5	1	
0	0	0	0	0

STATUS PB/U

	0,136434006	0,411599691		
0,331472566				
Normal	1653	1588	42	7
5	9	2	0,301600498	
Pendek	51	30	16	4
0	1	0	0	
Tinggi	64	46	7	0
8	3	0	0	

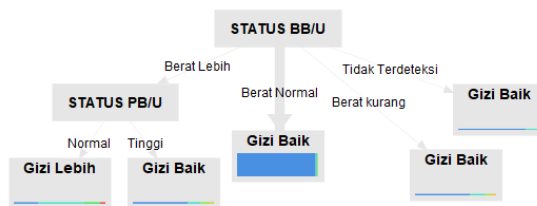
4.4.2 Pembentukan Model dengan Menggunakan RapidMiner

Peneliti akan menerapkan pembagian data latih dan uji, dengan variasi *rasio* pembagian seperti 90:10, 80:20, 70:30, 60:40, dan 50:50. Proses ini akan dilakukan secara simultan melalui operator *multiply* untuk menjalankan lima percobaan sekaligus. Pendekatan ini bertujuan menciptakan variasi dalam pembagian data, memastikan keragaman dalam analisis menggunakan metode *decision trees* terkait status *stunting* pada kelompok balita di Kecamatan Tukdana.

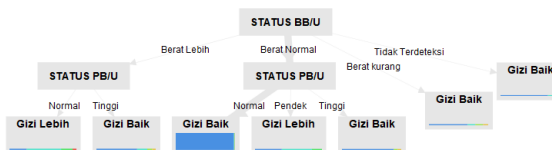


Gambar 2. Pembagian Data dalam Proses Pembentukan Model

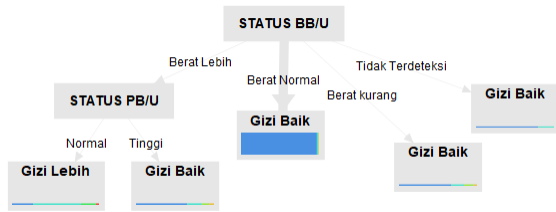
Dalam Gambar 2, terlihat langkah-langkah pemodelan menggunakan alat RapidMiner. Data latih dibagi menjadi lima tahap dengan perbandingan yang berbeda antara data latih dan data uji. Proses ini menghasilkan pohon keputusan yang mencerminkan analisis mengenai status *stunting* pada kelompok balita di Kecamatan Tukdana.



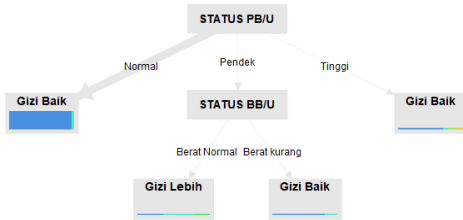
Gambar 3. Pohon Keputusan dengan 90% Data Latih



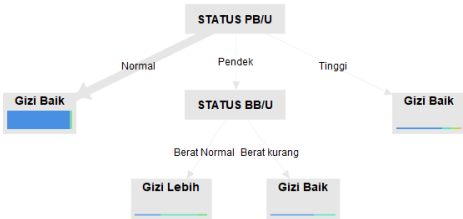
Gambar 4. Pohon Keputusan dengan 80% Data Latih



Gambar 5. Pohon Keputusan dengan 70% Data Latih



Gambar 6. Pohon Keputusan dengan 60% Data Latih



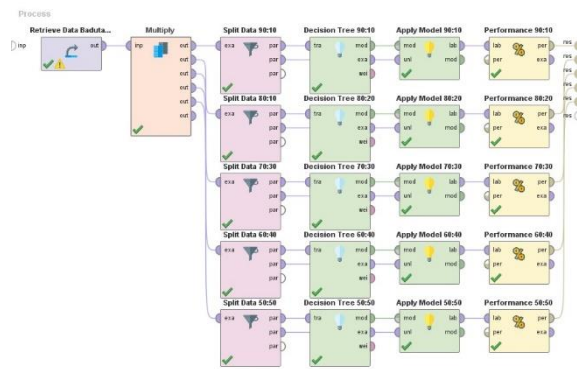
Gambar 7. Pohon Keputusan dengan 50% Data Latih

Dari Gambar 3 hingga Gambar 7, terlihat *output* proses pemodelan *decision tree* dengan *parameter criterion (gain_ratio)* dan *maximal_depth (3)*. Tahapan ini menyajikan visualisasi analisis status *stunting* pada kelompok balita di Kecamatan Tukdana menggunakan metode *decision tree* yang telah diatur secara optimal.

4.5 Evaluasi

Evaluasi dalam penelitian ini terdiri dari tiga langkah kunci. Pertama, penilaian terhadap model yang dibentuk selama proses pembuatan model. Kedua, validasi kinerja model menggunakan metode *cross-validation* untuk memastikan akurasi bebas dari bias potensial. Terakhir, fokus pada penyetelan *parameter* metode *decision tree* untuk mencari *set parameter* optimal yang menghasilkan kinerja model terbaik.

4.5.1 Pengaruh Rasio Pembagian Data Latih dan Data Uji terhadap Kinerja Model



Gambar 8. Data Latih dan Data Uji terhadap Kinerja Model *Decision Tree*

Tampilan pada gambar 8 Menyajikan Evaluasi Kinerja Model *Decision Tree* dengan Membandingkan Data Pelatihan dan Pengujian untuk Menilai Efektivitas dan Generalisasi Model.

accuracy: 94.29%

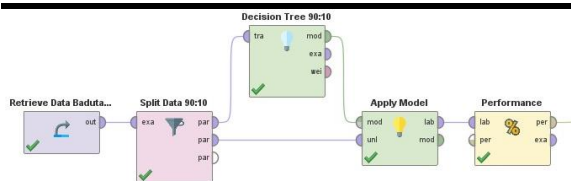
	true Gizi Baik	true Gizi Lebih	true Obesitas	true Gizi Buruk	true Gizi Kura...	true Tidak Ter...	class precision
pred. Gizi Baik	162	3	0	1	1	0	97.01%
pred. Gizi Leb...	4	3	0	0	0	0	37.50%
pred. Obesitas	0	0	0	0	0	0	0.00%
pred. Gizi Bur...	0	0	0	0	0	0	0.00%
pred. Gizi Kur...	0	0	0	0	0	0	0.00%
pred. Tidak T...	0	0	0	0	0	0	0.00%
class recall	97.59%	50.00%	0.00%	0.00%	0.00%	0.00%	

Gambar 9. Evaluasi Kinerja Model Terbaik Dengan Rasio 90:10

Tampilan pada gambar 9 Menyajikan evaluasi kinerja model terbaik dengan memanfaatkan pembagian data menjadi data pelatihan dan pengujian dalam proporsi 90:10. Hasilnya mencatatkan rata-rata akurasi sebesar 94.29%, dengan *presisi* mencapai 97.01%, *recall* sebesar 97.59%, dan skor *F1* mencapai 97.29%.

4.5.2 Penilaian Kinerja Menggunakan Pendekatan *Split Data*

Penilaian model pada platform RapidMiner menggunakan operator "*apply model*" untuk menerapkan model pada data uji. Langkah berikutnya melibatkan operator "*Split Data*" untuk mengukur dan mendapatkan akurasi dari model yang telah dilatih.

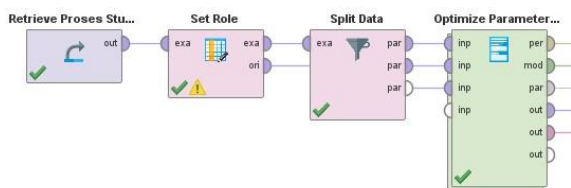


Gambar 10. Hasil Evaluasi Performa Model Melalui Proses *Split Data*

Tampilan pada gambar 10 Melalui proses Pembagian Data dengan rasio 90:10 antara data pelatihan dan pengujian, model terbaik menunjukkan kinerja yang sangat baik. Hasilnya mencatat akurasi rata-rata sebesar 94.29%, dengan tingkat presisi mencapai 97.01%, tingkat recall mencapai 97.59%, dan nilai F1-score mencapai 97.29%.

4.5.3 Penyetelan Parameter Pada *Decision Trees*

Langkah selanjutnya yang akan diambil oleh peneliti adalah melakukan optimasi atau penyetelan pada *parameter Decision Tree* dengan tujuan mencari konfigurasi *parameter* yang optimal dan suboptimal untuk melatih model. *Tuning parameter* dalam Data Mining merupakan serangkaian proses pemodelan yang dilakukan secara iteratif pada suatu metode dengan variasi *parameter* yang berbeda. Dalam upaya melakukan penyetelan *parameter*, peneliti akan menggunakan metode Pembagian Data sebagai dasar evaluasi model. Dalam konteks penyetelan algoritma *Decision Tree*, peneliti akan menggunakan akurasi sebagai pedoman utama untuk memilih *set parameter* terbaik.



Gambar 11. Model Proses *Tuning Parameter*

Tampilan pada gambar 11 *Tuning parameter* pada platform RapidMiner akan difokuskan pada penyetelan *parameter criterion*, *maximal depth*, dan *minimal leaf size*. Batasan yang telah ditetapkan untuk masing-masing *parameter* adalah sebagai berikut: kriteria *Criterion* akan divariasikan dengan opsi *Gain ratio*, *Information Gain*, dan *Gini Index*.

Parameter Maximal Depth akan dibatasi dalam rentang nilai antara 1 hingga 10, sementara ukuran *Minimal Leaf* akan dibatasi dalam rentang nilai antara 1 hingga 20. Oleh karena itu, proses *tuning parameter* ini akan melibatkan eksplorasi berbagai kombinasi nilai *parameter* dengan tujuan untuk mengoptimalkan kinerja algoritma *Decision Tree* di platform RapidMiner.

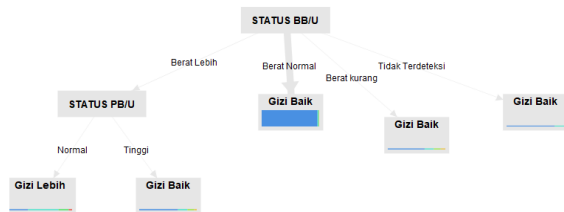
Tabel 2. Hasil Penyetelan *Parameter* dalam RapidMiner

<i>Criterion</i>	<i>Maximal Depth</i>	<i>Minimal Leaf Size</i>	<i>Accuracy</i>	<i>Rank</i>
<i>Gain Ratio</i>	3	5	95,43 %	1
<i>Information Gain</i>	3	5	95,43 %	2
<i>Gini Index</i>	3	5	95,43 %	3
<i>Gain Ratio</i>	5	5	95,43 %	4
<i>Information Gain</i>	5	5	95,43 %	5
.....				
<i>Gini Index</i>	10	1	94,29 %	60

Dari hasil penyetelan *parameter* (Tabel 2), didapatkan tingkat akurasi tertinggi pada konfigurasi tertentu. *Criterion Gain Ratio*, *maximal depth 3*, dan *minimal leaf size 5* memberikan hasil optimal, mencapai tingkat akurasi 95.43%.

4.6 Implementasi

Dalam tahap akhir implementasi model, model yang dioptimalkan melalui penyetelan *parameter* di platform RapidMiner dipilih untuk digunakan dalam situasi praktis. Proses ini melibatkan penyesuaian model, integrasi ke dalam sistem yang dibutuhkan, pengujian, dan pemantauan performa untuk memastikan konsistensi prediksi. *Deployment* dirancang untuk mendukung pengambilan keputusan yang lebih baik dengan menggunakan model *decision tree*.



Gambar 12. Menampilkan Hasil Akhir Dari Pohon Keputusan Yang Dikembangkan

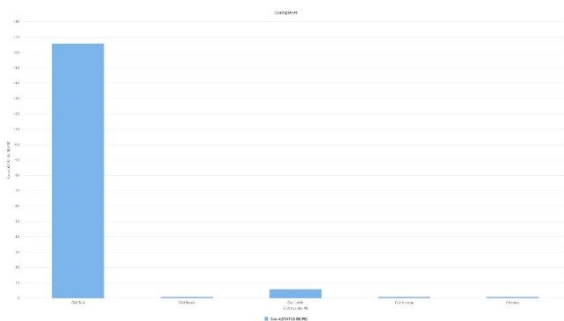
Tampilan pada gambar 12, terlihat hasil akhir dari *Decision Tree* setelah melalui penyetulan *parameter* dan evaluasi dengan metode *split data*. Model ini mencerminkan struktur pohon keputusan yang telah dioptimalkan untuk menganalisis data *stunting* pada kelompok baduta di Kecamatan Tukdana, dengan tujuan meningkatkan akurasi prediksi dan memperjelas interpretabilitas model dalam konteks penelitian.

accuracy: 95.43%

	true Gizi Baik	true Gizi Lebih	true Obesitas	true Gizi Buruk	true Gizi Kura.	true Tidak Ter...	class precision
pred. Gizi Baik	165	4	0	1	1	0	96.49%
pred. Gizi Leb...	1	2	1	0	0	0	50.00%
pred. Obesitas	0	0	0	0	0	0	0.00%
pred. Gizi Bur...	0	0	0	0	0	0	0.00%
pred. Gizi Kura...	0	0	0	0	0	0	0.00%
pred. Tidak T...	0	0	0	0	0	0	0.00%
class recall	99.40%	33.33%	0.00%	0.00%	0.00%	0.00%	

Gambar 13. Hasil Akhir Dengan Nilai Akurasi Terbaik Pada Split Data

Tampilan pada gambar 13 Hasil akhir dari proses Pembagian Data pada model *Decision Tree* menunjukkan tingkat akurasi rata-rata sebesar 95.43%, dengan tingkat ketepatan sebesar 96.49%, tingkat pemulihan sebesar 99.40%, dan nilai *F1* mencapai 97.90%.



Gambar 14. Hasil Akhir Dengan Menggunakan Grafik

Tampilan pada gambar 14 Menampilkan hasil visualisasi grafik yang mengilustrasikan

perbandingan posisi puncak dari kategori "gizi baik" dibandingkan dengan kategori lainnya. Grafik ini menyajikan perbandingan relatif antara status gizi bayi dalam kategori tersebut, memberikan gambaran yang jelas tentang distribusi dan perbandingan proporsi antar kategori gizi.

SIMPULAN

Dalam menerapkan algoritma C4.5 untuk menganalisis status *Stunting* pada kelompok balita di Kecamatan Tukdana, langkah awal melibatkan identifikasi atribut yang krusial. Evaluasi berbagai nilai, termasuk *Entropy*, *Gain*, *split info*, dan *Gain info*, menunjukkan bahwa STATUS BB/U memiliki peran yang sangat penting, dengan *Gain info* mencapai 0,752388586. Informasi dari STATUS BB/U memiliki kontribusi yang signifikan dalam pemahaman tentang status *Stunting* pada kelompok balita. Dengan nilai *Entropy* sebesar 0,418416824, pemilihan STATUS BB/U menjadi langkah kunci dalam pembentukan model *Decision Tree*, dengan harapan dapat meningkatkan efektivitas model dalam menganalisis status *Stunting*.

Evaluasi pemodelan *Decision Trees* untuk status *Stunting* di Kecamatan Tukdana menemukan bahwa atribut utama seperti STATUS USIA IBU, AKSES AIR MINUM, TEMPAT BUANG AIR BESAR, STATUS BB/U, dan STATUS PB/U memiliki pengaruh signifikan. STATUS BB/PB dan STATUS STUNTING terbukti sebagai atribut paling berpengaruh, dengan akurasi, *presisi*, *recall*, dan *F1-score* yang sangat baik. *Tuning parameter*, khususnya pada *maximal depth* (1-10) dan *minimal leaf size* (1-20), memberikan fleksibilitas untuk mengoptimalkan model. Mayoritas bayi di Kecamatan Tukdana menunjukkan status gizi yang baik, namun proporsi kecil menunjukkan kondisi gizi yang tidak optimal. Meskipun demikian, prevalensi *stunting* dalam sampel tersebut relatif rendah. Hasil analisis ini memberikan pemahaman yang lebih dalam tentang kesehatan anak di wilayah tersebut, menekankan perlunya pemantauan lanjutan untuk memastikan pemeliharaan kesehatan optimal.

Dari analisis faktor-faktor yang memengaruhi status *stunting* di Kecamatan Tukdana, atribut utama, seperti STATUS USIA IBU, AKSES AIR MINUM, TEMPAT BUANG AIR BESAR, STATUS BB/U, dan STATUS

PB/U, memiliki peran penting. Algoritma C4.5 memberikan panduan untuk merancang program intervensi, tidak hanya untuk mencegah *stunting* di Kecamatan Tukdana, tetapi juga untuk meningkatkan efektivitas program kesehatan masyarakat secara lebih luas.

Dalam penelitian selanjutnya, dapat dieksplorasi lebih lanjut mengenai faktor-faktor lain yang mungkin berkontribusi terhadap status *stunting* pada kelompok balita di Kecamatan Tukdana. Selain itu, mempertimbangkan penggunaan metode-metode *machine learning* atau teknik analisis data yang lebih canggih dapat membantu dalam meningkatkan akurasi dan keakuratan prediksi model. Selain itu, penting untuk melakukan survei yang lebih luas dan mendalam serta mengumpulkan data yang lebih lengkap untuk memperoleh pemahaman yang lebih komprehensif tentang kondisi kesehatan balita di wilayah tersebut. Hal ini dapat membantu dalam merancang program intervensi yang lebih efektif dan berkelanjutan untuk mencegah *stunting* dan meningkatkan kesehatan anak-anak secara keseluruhan.

DAFTAR PUSTAKA

- Abidin, Z., Nurhana, E., Permata, & Ulum, F. (2023). Analisis Perbandingan Algoritma Decision Tree C4.5 Dan C5.0 Pada Data Karyawan Berpotensi Promosi Jabatan. *ZA. Pagar Alam*, 17(2), 567–582. <https://doi.org/https://doi.org/10.33365/jti.v17i2.2702>
- Ainurrohmah, A., & Wiyanti, D. T. (2023). Analisis Performa Algoritma Decision Tree, Naive Bayes, K-Nearest Neighbor untuk Klasifikasi Zona Daerah Risiko Covid-19 di Indonesia. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 10(1), 115–122. <https://doi.org/10.25126/jtiik.20231015935>
- Alawiah, R. H., Saifullah, & Damanik, I. S. (2021). Analisis Kepuasan Konsumen Terhadap Pelayanan Bengkel Menggunakan Metode Algoritma C4.5. *Januari*, 2(1), 31–38. <https://doi.org/https://doi.org/10.30645/kesatria.v2i1.55>
- Hana, F. M. (2020). Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4.5. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, 4(1), 32–39. <https://doi.org/10.47970/siskom-kb.v4i1.173>
- Lestari, S., & Amalia, R. A. (2023). Penerapan Algoritma C. 45 Pada Klasifikasi Status Gizi Balita di Posyandu Desa Sukalilah Cibatu Kabupaten Garut Jawa Barat. *Jurnal Sains Dan Teknologi*, 5(1), 177–182. <https://doi.org/https://doi.org/10.55338/saintek.v5i1.1375>
- Muttaqin, A. R., Wibawa, A., & Nabila, K. (2021). Inovasi Digital untuk Masyarakat yang Lebih Cerdas 5.0: Analisis Tren Teknologi Informasi dan Prospek Masa Depan. *Jurnal Inovasi Teknologi Dan Edukasi Teknik*, 1(12), 880–886. <https://doi.org/10.17977/um068v1i122021p880-886>
- Solehuddin, M., Syafei, W. A., & Gernowo, R. (2022). Metode Decision Tree untuk Meningkatkan Kualitas Rencana Pelaksanaan Pembelajaran dengan Algoritma C4.5. *Jurnal Penelitian Dan Pengembangan Pendidikan*, 6(3), 510–519. <https://doi.org/10.23887/jppp.v6i3.52840>
- Suhanda, Y., Kurniati, I., & Norma, S. (2020). Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik. *Jurnal Teknologi Informatika Dan Komputer*, 6(2), 12–20. <https://doi.org/10.37012/jtik.v6i2.299>
- Surya, K., Kusuma, H., Atros, A., Padri, A. R., Nurdiawan, O., Faqih, A., & Anwar, S. (2021). Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan Online Menggunakan K-Means dan Decision Tree. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 323–329. <https://doi.org/10.30865/jurikom.v8i6.3680>
- Wahono, H., & Riana, D. (2020). Prediksi Calon Pendonor Darah Potensial Dengan Algoritma Naive Bayes, K-Nearest Neighbors dan Decision Tree C4.5. *JURIKOM (Jurnal Riset Komputer)*, 7(1), 7–14. <https://doi.org/10.30865/jurikom.v7i1.1953>
- Wayan Wardani, N., Gede Surya Cipta Nugraha, P., Hartono, E., Wayan Dharma Suryawan, I., Manik Dirgayusari, A., Wayan Darmadi, I., & Surya Mahendra, G. (2022). Penerapan Data Mining Untuk Klasifikasi Penjualan Barang Terlaris Menggunakan Metode Decision Tree C4.5. *Jurnal Teknologi Informasi Dan Komputer*, 8(3), 268–279. <https://jurnal.undhirabali.ac.id/index.php/jutik/article/view/2081>