

Penerapan *Algoritma C4.5* Untuk Prediksi Anak *Stunting* Di Kota Pagar Alam

Revaldo Xsanal Hakim¹, Ferry Putrawansyah², Riduan Syahri³

^{1,2,3} Program Studi Teknik Informatika, Institut Teknologi Pagar Alam, Pagar Alam, Indonesia
 Email: ¹revaldo22xsanal@gmail.com, ²feypuawansyah@gmail.com, ³syahririduan@gmail.com

Abstrak– Di Pagar Alam, *Prediksi* dan pengukuran tingkat *Stunting* masih mengandalkan analisis sekunder. Kader Posyandu melibatkan diri dalam mengukur kondisi balita, dan hasilnya diserahkan kepada ahli untuk menilai apakah balita tersebut mengalami *Stunting* atau tidak. Tujuan dari penelitian ini adalah mengaplikasikan *Algoritma C4.5* untuk melakukan *Prediksi* terkait kasus *Stunting* pada anak. Dari permasalahan yang ada diatas, maka metode yang dapat menyelesaikan permasalahan ini yaitu *Algoritma C4.5* yang termasuk dalam Pohon Keputusan pada *Data Mining*. Proses *Data Mining* peneliti menggunakan salah satu metode CRISP-DM dan pengujian *Data Mining* menggunakan *Confusion Matrix* serta pengujian *Sistem* menggunakan *Black Box Testing*. Hasil dalam penelitian ini berupa sebuah *Sistem* yang menerapkan aturan dari Pohon keputusan. *Sistem Prediksi* status gizi Anak yang dirancang penulis layak karena dapat mengkategorikan status gizi balita. balita secara otomatis berdasarkan *Zscore* yang ditetapkan dan hanya terdapat selisih 11,8% dari pengujian *Prediksi* dataset yang sama menggunakan *Rapid Miner*. *Sistem* yang penulis rancang dapat lebih cepat dan efektif dalam mem*Prediksi* status gizi Anak. Berdasarkan Berdasarkan data hasil uji, dapat disimpulkan bahwa akurasi *Algoritma C4.5* untuk mem*Prediksi* anak *Stunting* yaitu 88,20% tergolong baik. Sedangkan pengujian *Sistem* menggunakan *Black Box Testing* memperoleh total 4.35 masuk ke kategori sangat layak. Tujuan *Sistem* ini adalah membantu tenaga kesehatan dalam membuat keputusan terkait *Prediksi* status gizi balita. *Sistem Prediksi* ini bermanfaat untuk mengidentifikasi balita berisiko gizi buruk sehingga tindakan pencegahan dapat dilakukan dengan lebih efektif.

Kata kunci : *Data Mining; C4.5 Algorithm; Prediction; Stunting; Crisp Dm*

Abstract– In Pagar Alam, the prediction and measurement of *Stunting* rates still rely on secondary analysis. Posyandu cadres involve themselves in measuring the condition of toddlers, and the result are handed over to experts to assess whether the toddlers are stunted or not. The purpose of this study is to apply the *C4.5* algorithm to predict *Stunting* cases in children. From the problems above, the method that can solve this problem is the *C4.5* Algorithm which is included in the Decision Tree in *Data Mining*. The *Data Mining* process researchers use one of the CRISP-DM methods and *Data Mining* testing using the *Confusion Matrix* and system testing using *Black Box Testing*. The result in this study is in the form of a system that applies the rules of the Decision Tree. The Child Nutritional Status Prediction system designed by the author is feasible because it can categorize the nutritional status of toddlers. toddlers were automatically based on a set *Zscore* and there was only an 11.8% difference from testing the same dataset prediction using *Rapid Miner*. The system that the authors designed can be faster and more effective in predicting children's nutritional status. Based on the test data, it can be concluded that the accuracy of the *C4.5* algorithm to predict *Stunting* children, which is 88.20%, is good. While testing the system using *Black Box Testing* obtained a total of 4.35 into the very decent category. The purpose of this system is to assist health workers in making decisions related to the prediction of nutritional status of toddlers. This prediction system is useful for identifying toddlers at risk of malnutrition so that preventive measures can be taken more effectively.

Keywords: *Data Mining; C4.5 Algorithm; Prediction; Stunting; Crisp Dm*

1. PENDAHULUAN

Perkembangan pesat dalam teknologi informasi membuka pintu lebar-lebar untuk kemudahan akses informasi melalui internet, memungkinkan akses di berbagai lokasi dan waktu. Internet menjadi sumber beragam informasi yang mencakup segala aspek kehidupan, termasuk Kesehatan, yang memegang peranan krusial dalam kehidupan [1]. saat ini komputer dapat mempercepat proses pengolahan data seperti halnya untuk mem*Prediksi* suatu kejadian hanya dengan memasukkan beberapa *Variabel* yang telah ditentukan. Seorang *stoner* hanya tinggal memasukkan beberapa data, sehingga dari data tersebut *Sistem* dapat langsung mem*Prediksi* suatu kejadian. Dengan memanfaatkan *Sistem* komputer yang ada dan menggunakan jaringan komputer yang terintegrasi, maka akan didapatkan suatu *Sistem Prediksi* yang cepat, tepat dan akurat. Dalam proses kegiatannya salah satu permasalahan yang memerlukan *Sistem Prediksi* adalah untuk membantu Memprediksi anak yang kekurangan gizi di kota Pagar Alam. Hingga saat ini, Indonesia masih dihadapkan pada berbagai masalah kesehatan, termasuk isu gizi balita sebagai salah satu contohnya [2].

Seperti yang diungkapkan oleh [3], gizi adalah proses di mana *Orgaisme* menggunakan makanan dengan cara normal melalui proses pencernaan, penyerapan, transportasi, penyimpanan, dan pengeluaran zat-zat yang tidak diperlukan untuk mendukung kelangsungan hidup, pertumbuhan, fungsi normal organ tubuh, dan penghasilan energi. Kelompok usia anak balita disebut sebagai "usia emas" karena pada periode ini, anak sedang

membangun dasar untuk pertumbuhan tubuhnya di masa depan. Pada *FAse* ini, anak sangat rentan terhadap masalah gizi karena membutuhkan asupan gizi yang memadai untuk mendukung proses tumbuh kembangnya [4]. Kekurangan gizi kronis dapat mengakibatkan kondisi *Stunting*.

Dalam pandangan [5], gizi buruk adalah kondisi yang terjadi ketika seseorang mengalami kekurangan asupan nutrisi atau jumlah nutrisi yang dikonsumsi berada di bawah standar. *Stunting*, yang umumnya terjadi pada balita, menjadi isu kesehatan masyarakat di Indonesia, berpotensi menurunkan produktivitas dan kualitas sumber daya manusia di masa depan. *Stunting* dapat meningkatkan risiko penyakit, kematian, gangguan perkembangan motorik otak, dan mengurangi tingkat produktivitas [6]. Meskipun *prevalensi* penyakit menular di Indonesia mencapai 24,4, hal ini dianggap sebagai masalah karena melebihi batas standar yang ditetapkan oleh Organisasi Kesehatan Dunia (WHO), di mana tingkat prevalensi penyakit menular dalam satu negara seharusnya tidak melebihi 20 [7]. Menurut [8], implementasi mengacu pada tindakan yang dilakukan oleh individu atau kelompok untuk mencapai tujuan yang telah ditetapkan dan disepakati dalam suatu keputusan. Sementara menurut [9], penerapan (implementasi) merupakan perluasan aktivitas yang menyesuaikan proses interaksi antara tujuan dan tindakan guna mencapainya, melibatkan jaringan pelaksana dan birokrasi yang efektif. Menurut [10], CRISP-DM adalah standar proses penambangan data yang dikembangkan oleh tiga perusahaan pembuat perangkat lunak penambangan data, yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), dan NCR. Standar ini kemudian mengalami pengembangan melalui serangkaian lokakarya antara tahun 1997 hingga 1999. Sebuah penelitian serupa berjudul "*Prediksi Status Gizi Balita Menggunakan Algoritma C4.5*" oleh [6] menggunakan *Algoritma C4.5* untuk mengklasifikasikan dan mem*Prediksi* status gizi balita di kecamatan Bringin. Pengujian dilakukan pada *Sistem* yang dibangun oleh penulis, serta melibatkan pengujian homemade dan menggunakan *Rapid Miner* sebagai pembanding.

Penelitian "*Simulasi Penerapan Metode Pohon Keputusan (C4.5) dalam Penentuan Status Gizi Balita*" oleh (bertujuan untuk mengetahui status gizi balita. Dalam penelitian ini, melibatkan 123 data balita yang lahir antara tahun 2015 hingga 2019. Hasil penelitian menunjukkan bahwa *Algoritma C4.5* dapat diaplikasikan untuk menentukan status gizi balita, dengan berat badan (C1) dan tinggi badan (C2) sebagai faktor penentu. Dalam penelitian "*Komparasi Algoritma C4.5 dengan Algoritma Naive Bayes Classifier untuk Klasifikasi Status Gizi Balita*" [11], terdapat lima atribut yang memengaruhi status gizi balita, termasuk jenis kelamin, umur, BB/U, TB/U, dan BB/TB. Penelitian ini menggunakan *Algoritma C4.5* dan *Algoritma Naive Bayes Classifier* untuk mengklasifikasikan status gizi balita.

Observasi dan wawancara di Dinas Kesehatan kota Pagar Alam menunjukkan bahwa *Prediksi Stunting* masih menggunakan metode Analisis Data Sekunder. Deteksi *Stunting* pada balita dapat dilakukan secara dini melalui pemantauan rutin kurva pertumbuhan anak di Posyandu terdekat, yang melibatkan pengukuran kondisi balita oleh kader Posyandu dan evaluasi hasil oleh ahli. Meskipun pendekatan ini memberikan wawasan tentang prevalensi *Stunting* dan faktor risiko, analisis data sekunder memiliki kelemahan seperti keterbatasan data, ketidakkonsistenan definisi dan metode, serta ketidakmungkinan memasukkan *Variabel* tertentu yang dapat mengakibatkan *Prediksi* tidak akurat dan tidak . Oleh karena itu, solusi yang diusulkan adalah menggunakan teknik *Data Mining*. Menurut [3], *Algoritma C4.5* merupakan salah satu *Algoritma Decision Tree* yang dikembangkan oleh J.Ross Quinlan. Sementara menurut [4], *Algoritma C4.5* juga dikenal sebagai *Decision Tree*. Pohon Keputusan adalah metode klasifikasi yang melibatkan konstruksi pohon keputusan dengan simpul keputusan terhubung dari simpul akar ke simpul daun akhir.

Data Mining, atau Penambangan Data, adalah proses ekstraksi informasi, pengidentifikasian pola, dan relasi dalam suatu himpunan data untuk menyelesaikan masalah tertentu. Tujuan utama *Data Mining* adalah mengenali relasi dan pola yang tidak terlihat langsung oleh analis manusia, memanfaatkan teknik komputasi dan statistik untuk menganalisis data dengan tujuan menghasilkan wawasan yang bermanfaat. Salah satu metode *Data Mining* untuk *Prediksi* adalah Metode Pohon Keputusan. Menurut [12], *Prediksi* adalah peramalan atau perkiraan, sementara menurut [13], ramalan atau soothsaying adalah ilmu yang mempelajari *Prediksi* kejadian di masa depan. Berdasarkan latar belakang ini, penelitian ini berjudul "*Penerapan Algoritma C4.5 untuk Prediksi Stunting di Kota Pagar Alam*." Harapannya, hasil penelitian ini dapat membantu Dinas Kesehatan dan masyarakat Kota Pagar Alam dalam mem*Prediksi* pemenuhan gizi anak-anak, sehingga kekurangan gizi dapat dicegah dan diminimalisir. Dengan mengurangi *Stunting* di Kota Pagar Alam, produktivitas dan kualitas masyarakat diharapkan meningkat, membuat sumber daya manusia di kota tersebut lebih kompetitif di masa depan. Tujuan penelitian ini adalah menerapkan *Algoritma C4.5* untuk mem *Prediksi Stunting* pada anak di Kota Pagar Alam.

2. METODE PENELITIAN

2.1 Alur Penelitian

1. Objek Penelitian

Penelitian ini dilakukan pada Dinas Kesehatan kota Pagar Alam yang ber alamat di Jl. Ais Nasution No. 1, Kota Pagar Alam, Sumatera Selatan. Penelitian dilakukan dari bulan Oktober sampai dengan Desember 2023, Dilanjutkan pada Desember 2023 sampai dengan Febuari 2023.

2. Prosedur Pengumpulan Data

a. Observasi

Observasi merupakan teknik pengumpulan data dan informasi yang dilakukan dengan cara mengamati langsung/ *survey Sistem* pada Dinas Kesehatan kota Pagar Alam.

b. Wawancara

Wawancara merupakan Teknik pengumpulan data yang dilakukan dengan cara mengajukan pertanyaan-pertanyaan langsung kepada Narasumber yang bersangkutan dengan subjek penelitian.

c. Studi Pustaka

Studi Pustaka merupakan teknik pengumpulan data yang dilakukan dengan cara mengumpulkan buku, *Ebook* dan jurnal untuk melakukan tinjauan pada sumber tersebut serta referensi relevan yang berhubungan dengan penelitian.

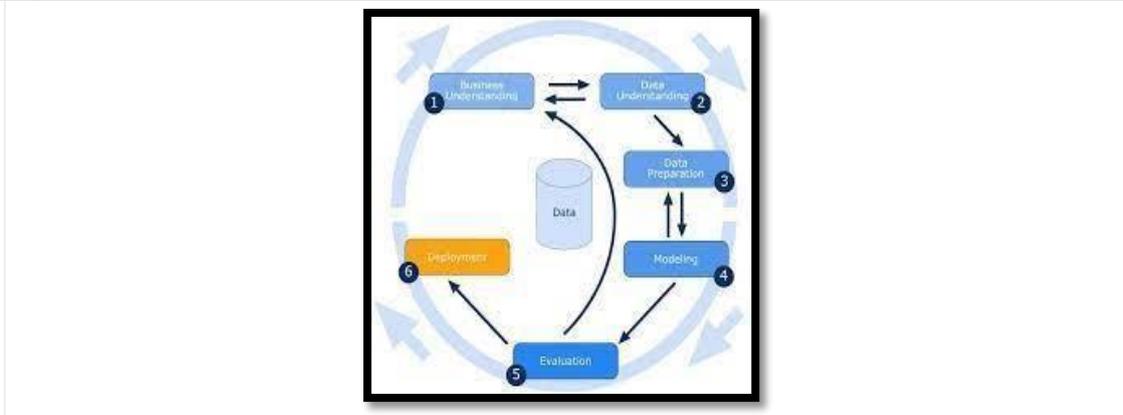
d. Dokumentasi

Dokumentasi merupakan teknik pengumpulan data yang dilakukan dengan cara memperoleh data dan menganalisa data yang berhubungan dengan subjek penelitian.

3. Metode Pengembangan Sistem

Pendekatan Pengembangan *Sistem* Cross Industry Standard Process for *Data Mining* (CRISPDM) adalah suatu proses standar untuk *Data Mining* yang dirancang oleh tiga inisiator utama dalam industri *Data Mining*, yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), dan NCR. Proses standar ini kemudian mengalami perkembangan melalui serangkaian lokakarya yang berlangsung antara tahun 1997 dan 1999 [10]. Gambaran tahap proses pada pengembangan *Sistem* ini dapat dilihat pada ilustrasi [14].

Penelitian mengadopsi metode *Cross Industry Standard Process* untuk *Data Mining* (CRISPDM) menurut [15]. Tahap-tahap dalam metode ini melibatkan:



Gambar 1. Metode CRIPSDM

- 1) **Pengertian Bisnis (*Business Understanding*)** : Dinas Kesehatan kota Pagar Alam menunjukkan bahwa *Prediksi Stunting* masih menggunakan metode Analisis Data Sekunder. Deteksi *Stunting* pada balita dapat dilakukan secara dini melalui pemantauan rutin kurva pertumbuhan anak di Posyandu terdekat, yang melibatkan pengukuran kondisi balita oleh kader Posyandu dan evaluasi hasil oleh ahli. Meskipun pendekatan ini memberikan wawasan tentang prevalensi *Stunting* dan faktor risiko, analisis data sekunder memiliki kelemahan seperti keterbatasan data, ketidakonsistenan definisi dan metode, serta ketidakmungkinan memasukkan *Variabel* tertentu yang dapat mengakibatkan *Prediksi* tidak akurat dan tidak representatif.
- 2) **Pemahaman Data (*Data Understanding*)** : Pada tahap pengumpulan data, Data diambil dari Dinas Kesehatan kota Pagar Alam, Data yang diambil berupa data balita dari tahun 2018 sampai tahun 2023, data yang digunakan dalam bentuk file *Microsoft Excel* dengan jumlah data 75210 data. Data tersebut

berisikan data balita untuk atribut yang diperoleh berupa No, NIK, Nama, JK, Usia Saat Ukur, Tanggal Pengukuran, Berat, Tinggi.

Tabel 1. Potongan Data Belum Siap Diolah

No	NIK	Nama	JK	Usia Saat Ukur	Berat	Tinggi
1.	1672030509 170001	ADE BIMANTARA	L	4 Tahun - 11 Bulan - 17 Hari	15	107
2.	1672041610 170001	AHMAD REZA	L	4 Tahun - 9 Bulan - 24 Hari	18,9	115
3.	1672041310 170001	ALFARIZKI	L	4 Tahun - 9 Bulan - 19 Hari	18,8	117
4.	1672051908 180001	RAHMAT PUTRA PRAMANAT A	L	4 Tahun - 11 Bulan - 15 Hari	15,4	104
5.	1672042008 160001	LEO	L	4 Tahun - 10 Bulan - 17 Hari	16	108
6.	1672046409 100001	ALZAHRA	P	3 Tahun - 10 Bulan - 16 Hari	19,4	120
7.	1672045808 160002	AGUSTI	P	4 Tahun - 11 Bulan - 13 Hari	20	118
8.	1672034105 180001	FAUZIAH	P	4 Tahun - 10 Bulan - 15 Hari	15	107
9.	1672032412 170002	ABID AQILA	L	4 Tahun - 7 Bulan - 26 Hari	15	104
10.	1672030708 180001	ANWAR MAULANA	L	4 Tahun - 10 Bulan - 13 Hari	15	104
11.
752	1673305908	ELVIA	P	0 Tahun - 0 Bulan	3,2	49
1.	235235	ADILA		- 0 Hari		

3) **Persiapan Data (Data Preparation) :**

Pada sesi ini dicoba persiapan data dengan memilih *Field* yang akan dipakai kemudian ditransformasikan. Prosesnya sebagai berikut :

1. Pembersihan Data
Tahap pembersihan data dilakukan pembersihan terhadap record data yang kosong ataupun tidak lengkap.
2. Pemilihan *Field*
Untuk pemilihan *Field* sendiri akan dipilih 5 *Field* yaitu Nama, Jenis Kelamin, Umur, BB, TB.
3. Transformasi Data
Untuk transformasi data sendiri agar dapat dipakai didalam *Rapid Miner* studio maka ditambahkan Atribut *Zscore* dan Status Agar Rstudio Dapat membuat aturan berdasarkan *Algoritma C4.5* yang nantinya aturan tersebut akan diimplementasikan kedalam *Sistem* yang akan dibuat nantinya. Potongan data transformasi dari data balita dilihat pada Tabel 2.

Tabel 2. Potongan Hasil Data Siap Diolah

No	NIK	Nama	JK	Usia Saat Ukur	Berat	Tinggi	Zscore	Status
1.	1672030509170001	ADE BIMANTARA	L	59	15	107	-1,9	Normal
2.	1672041610170001	AHMAD REZA	L	57	18,9	115	2,3	obesitas
3.	1672041310170001	ALFARIZKI	L	57	18,8	117	2,2	obesitas
4.	1672051908180001	RAHMAT PUTRA PRAMANATA	L	59	15,4	104	-1,5	Normal

5 .	1672042008160001	LEO	L	58	16	108	-0,8	Normal
6 .	1672046409100001	ALZAHRA	P	46	19,4	120	4,4	obesitas
7 .	1672045808160002	AGUSTI	P	59	20	118	3,1	obesitas
8 .	1672034105180001	FAUZIAH	P	58	15	107	-1,8	Normal
9 .	1672032412170002	ABID AQILA	L	55	10	104	-5,4	Stunting
10 .	1672030708180001	ANWAR MAULANA	L	58	15	104	-1,8	Normal
11
100	1672036402190001	MIKHAYLA AZAHRA PUTRI	P	53	14	99	-2,1	Stunting

4) Modelling Phase

Ditahap pemodelan penelitian menggunakan metode *Data Mining* dengan prosedur *Algoritma C4.5* yang dibantu dengan Rstudio. Proses ini menghasilkan sebuah pohon keputusan yang aturan-aturan dari pohon keputusan tersebut akan dapat diimplementasikan kedalam sebuah *Sistem Prediksi* anak *Stunting* di kota Pagar Alam.

- 5) **Evaluasi (Evaluation):** Model yang dihasilkan Pada tahap Evaluasi model yang dihasilkan sudah membentuk pohon keputusan dan sudah memiliki cabang berdasarkan *Entropy* dan *Gain* yang sudah dihitung. Model ini juga telah diuji menggunakan confusion matrix dan menghasilkan akurasi sebesar 90,00%, sehingga sangat cocok untuk diimplementasikan kedalam sebuah *Sistem Prediksi* anak *Stunting* di kota pagar Alam.

- 6) **Penyebaran (Deployment):** Pada tahap ini, pengetahuan dan data serta aturan yang diperoleh disusun untuk membangun sebuah *Sistem Prediksi* yang nantinya bertujuan untuk membantu Dinas Kesehatan kota Pagar Alam agar pelaksanaan dalam melayani masyarakat dapat lebih efektif dan efisien.

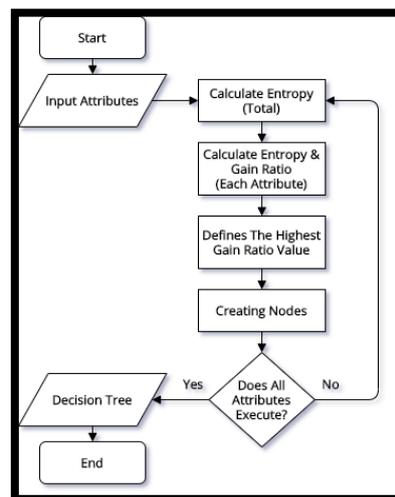
3. HASIL DAN PEMBAHASAN

3.1 Hasil

Hasil dalam penelitian ini berupa sebuah *Sistem* yang menerapkan aturan dari Pohon keputusan yang terbentuk melalui data yang telah di pelajari, dengan mengisi atribut inti yaitu umur, berat badan, tinggi badan, maka apakah seorang anak tersebut terkena stunting atau tidak dapat di *Prediksi*, apabila seorang anak tersebut terkena stunting maka dapat dilakukan tindakan pencegahan secara dini

3.2 Pembahasan

Flowchart proses *Algoritma C4.5* bisa dilihat di gambar 3 flowchat *Algoritma C4.5*.



Gambar 2. Flowchart *Algoritma C4.5*

Berikut ini adalah tahapan flowchart *Algoritma C4.5* :

Buat atribut yang akan dihitung;

1. Menghitung *Entropy* dan *Gain* dari setiap atribut;

a. Pilih Atribut akar

Untuk memilih atribut sebagai akar, didasarkan pada nilai *Gain* tertinggi dari atribut atribut yang ada. Dengan menggunakan persamaan diatas, maka akan didapatkan *Entropy* dan *Gain* yang digunakan sebagai akar dalam membuat pohon keputusan.

1) *Entropy* Total

$$\begin{aligned} \text{Entropy Total} &= \left(-\frac{51}{100} * \log_2\left(\frac{51}{100}\right)\right) + \left(-\frac{41}{100} * \log_2\left(\frac{41}{100}\right)\right) + \left(-\frac{8}{100} * \log_2\left(\frac{8}{100}\right)\right) \\ &= 1,314323 \end{aligned}$$

2) *Entropy* Atribut Umur

Untuk atribut Umur terdiri dari *Fase* nilai yaitu umur ≤ 23 Bulan, umur ≥ 24 Bulan, nilai *Entropy* nya adalah

a) umur ≤ 23 Bulan

(Jumlah kasus = 0, Hasil Lebih=0 , Hasil Baik = 0, Hasil Kurang =0

$$\begin{aligned} \text{Entropy Atribut umur} &= \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) + \left(-\frac{0}{0} * \log_2\left(\frac{0}{0}\right)\right) \\ &= 0 \end{aligned}$$

b) umur ≥ 24 Bulan

(Jumlah Kasus = 100, Hasil Lebih=51 , Hasil Baik = 41, Hasil Kurang = 8

$$\begin{aligned} \text{Entropy Atribut umur} &= \left(-\frac{51}{100} * \log_2\left(\frac{51}{100}\right)\right) + \left(-\frac{41}{100} * \log_2\left(\frac{41}{100}\right)\right) + \left(-\frac{8}{100} * \log_2\left(\frac{8}{100}\right)\right) \\ &= 1,314323 \end{aligned}$$

3) *Entropy* Jenis Kelamin

Untuk atribut jenis kelamin terdiri dari 2 nilai yaitu : jenis kelamin laki laki, dan jenis kelamin perempuan, dimana nilai *Entropy* masing masing adalah :

a) Jenis kelamin Laki-Laki

(Jumlah Kasus = 47, Hasil Lebih = 26, Hasil Baik = 18, Hasil Kurang = 3)

$$\begin{aligned} \text{Kelamin Laki-Laki} &= \left(-\frac{26}{47} * \log_2\left(\frac{26}{47}\right)\right) + \left(-\frac{18}{47} * \log_2\left(\frac{18}{47}\right)\right) + \left(-\frac{3}{47} * \log_2\left(\frac{3}{47}\right)\right) \\ &= -1,256185 \end{aligned}$$

b) Jenis Kelamin Perempuan

(Jumlah Kasus = 53, Hasil Lebih = 25, Hasil Baik = 23, Hasil Kurang = 3

$$\begin{aligned} \text{Kelamin Perempuan} &= \left(-\frac{25}{53} * \log_2\left(\frac{25}{53}\right)\right) + \left(-\frac{23}{53} * \log_2\left(\frac{23}{53}\right)\right) + \left(-\frac{3}{53} * \log_2\left(\frac{3}{53}\right)\right) \\ &= 1,355317 \end{aligned}$$

4) *Entropy* atribut Berat Badan

Untuk atribut berat badan terdiri dari 2 nilai yaitu: berat badan dengan berat kurang dari 10Kg, dan berat badan dengan berat lebih dari 10 Kg, dimana nilai *Entropy* masing masing adalah sebagai berikut :

a) Berat Ringan

(Jumlah Kasus = 8, Hasil Lebih =8 , Hasil Baik = 0, Hasil Kurang = 0

$$\begin{aligned} \text{Entropy Berat Badan} &= \left(-\frac{8}{8} * \log_2\left(\frac{8}{8}\right)\right) + \left(-\frac{0}{8} * \log_2\left(\frac{0}{8}\right)\right) + \left(-\frac{0}{8} * \log_2\left(\frac{0}{8}\right)\right) \\ &= -0,301029 \end{aligned}$$

b) Berat Sedang

(Jumlah Kasus = 41, Hasil Lebih = 0, Hasil Baik = 41, Hasil Kurang = 0)

$$\begin{aligned} \text{Entropy Berat} &= \left(-\frac{0}{41} * \log_2\left(\frac{0}{41}\right)\right) + \left(-\frac{41}{41} * \log_2\left(\frac{41}{41}\right)\right) + \left(-\frac{0}{41} * \log_2\left(\frac{0}{41}\right)\right) \\ \text{Badan} & \\ &= -0,301029 \end{aligned}$$

c) Massa Tubuh yang Berlebihan

(Jumlah Kejadian = 51, Output Lebih = 0, Output Baik = 0, Output Kurang = 51)

$$\begin{aligned} \text{Entropy Ber} &= \left(-\frac{0}{51} * \log_2\left(\frac{0}{51}\right)\right) + \left(-\frac{0}{51} * \log_2\left(\frac{0}{51}\right)\right) + \left(-\frac{51}{51} * \log_2\left(\frac{51}{51}\right)\right) \\ \text{Badan} & \\ &= -0,301029 \end{aligned}$$

5) Entropy Tinggi Badan

Untuk atribut Tinggi badan terdiri dari 2 nilai yaitu : Tinggi badan dengan tinggi kurang dari 87 cm dan tinggi badan lebih dari 88 cm dimana nilai Entropy masing masing adalah sebagai berikut :

a) Tinggi Rendah

(Jumlah Kasus = 7, Hasil Lebih = 6, Hasil Baik = 1, Hasil Kurang = 0)

$$\begin{aligned} \text{Entropy Tinggi} &= \left(-\frac{6}{7} * \log_2\left(\frac{6}{7}\right)\right) + \left(-\frac{1}{7} * \log_2\left(\frac{1}{7}\right)\right) + \left(-\frac{0}{7} * \log_2\left(\frac{0}{7}\right)\right) \\ \text{Rendah} & \\ &= -0,2277924173 \end{aligned}$$

b) Postur Tubuh Normal

(Jumlah Kejadian = 36, Output Lebih = 2, Output Baik = 24, Output Kurang = 10)

$$\begin{aligned} \text{Tinggi} &= \left(-\frac{2}{36} * \log_2\left(\frac{2}{36}\right)\right) + \left(-\frac{24}{36} * \log_2\left(\frac{24}{36}\right)\right) + \left(-\frac{10}{36} * \log_2\left(\frac{10}{36}\right)\right) \\ \text{Sedang} & \\ &= -0,1579478372 \end{aligned}$$

c) Elevasi (ketinggian)

(Jumlah Kejadian = 57, Output Lebih = 0, Output Baik = 16, Output Kurang = 41)

$$\begin{aligned} \text{Entropy Tinggi} &= \left(-\frac{0}{57} * \log_2\left(\frac{0}{57}\right)\right) + \left(-\frac{16}{57} * \log_2\left(\frac{16}{57}\right)\right) + \left(-\frac{41}{57} * \log_2\left(\frac{41}{57}\right)\right) \\ \text{(Tinggi)} & \\ &= -0,1794690987 \end{aligned}$$

6). Gain untuk atribut umur adalah :

Gain Umur

$$\begin{aligned} &= 1,314323 - \left(\left(\frac{0}{100} * 0\right) + \left(\frac{100}{100}(-1,314323)\right)\right) \\ &= -0 \end{aligned}$$

7). Hitung Gain atribut

Gain untuk atribut jenis Kelamin adalah :

Gain JK

$$\begin{aligned} &= 1,314323 - \left(\left(\frac{47}{100} * (-1,256185)\right) + \left(\frac{53}{100} * (-1,355317)\right)\right) \\ &= 0,00559804 \end{aligned}$$

8. Keuntungan atribut Berat Badan Keuntungan yang terkait dengan atribut berat badan adalah:

Gain BB

$$\begin{aligned} &= 1,314323 - \left(\left(\frac{8}{100} * -0,301029\right) + \left(\frac{41}{100} * -0,301029\right)\right) \\ &\quad + \left(\frac{51}{100} * -0,301029\right) \\ &= 1,151466311 \end{aligned}$$

8) Gain atribut Tinggi Badan

Gain untuk atribut Tinggi badan berikut ini :

$$\begin{aligned}
 & \text{Gain TB} \\
 & = 1,314323 \left(\left(\frac{7}{100} * -0,22779 \right) + \left(\frac{36}{100} * -0,1579 \right) + \left(\frac{57}{100} * -0,1794 \right) \right) \\
 & = 1,1392189231
 \end{aligned}$$

3. Melihat atribut yang memiliki *Gain* tertinggi

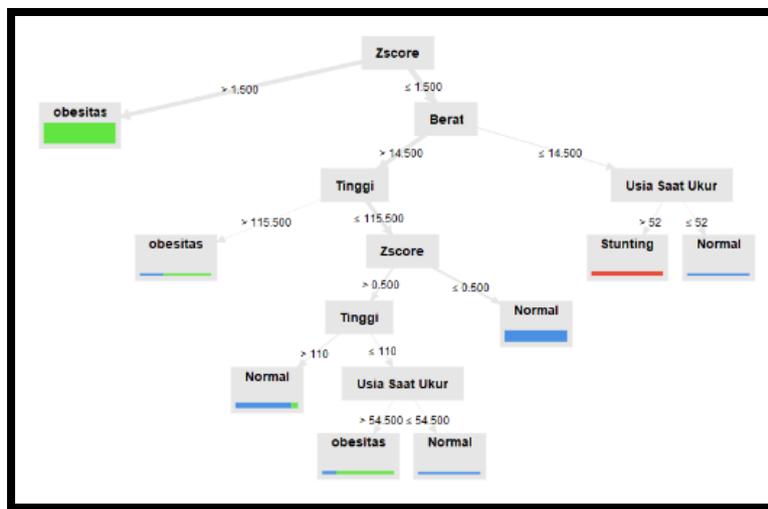
Gain tertinggi yaitu pada Atribut Berat Badan 1,151466311, *Gain* tertinggi yaitu pada Atribut tinggi badan 1,1392189231, *Gain tertinggi ketiga yaitu* pada Atribut Jenis Kelamin 0,00559804, dan *Gain* yang paling terendah pada Atribut *Umur* -0

4. Atribut yang memiliki *Gain* tertinggi dijadikan sebagai root (akar)

Atribut Berat Badan dipilih sebagai node akar karena memiliki *Gain* tertinggi.

5. Membagi data berdasarkan atribut dengan *Gain* tertinggi. Dengan BB sebagai Akar, dilanjutkan dengan Atribut TB, lalu Atribut Jenis Kelamin dan yang terakhir yaitu Atribut Umur

6. Membentuk pohon keputusan;

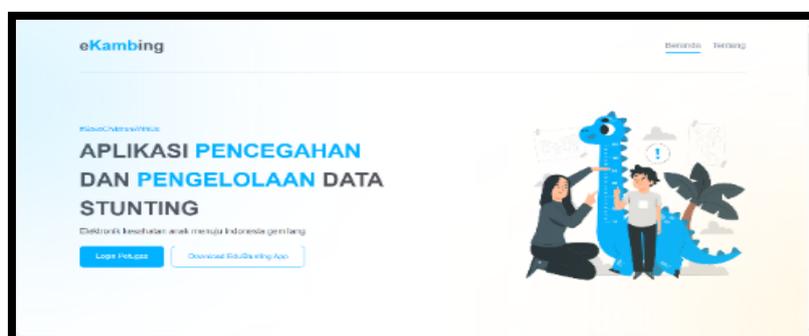


Gambar 3. Pohon Keputusan yang terbentuk

3.3 Implementasi

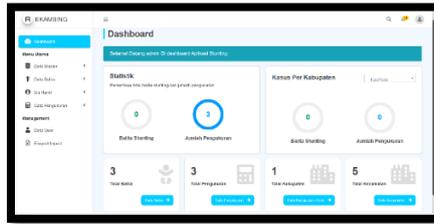
Selanjutnya adalah mengimplementasikan aturan – aturan yang terbentuk dari proses *Data Mining* menggunakan *Algoritma C4.5* berupa sebuah pohon keputusan ke dalam sebuah *Sistem* untuk mem*Prediksi* Anak terkena *Stunting* di kota Pagar Alam.

- 1) Halaman tampilan depan akan muncul sebagai tampilan awal saat *Admin* ingin melakukan Login ke dalam *Sistem*.



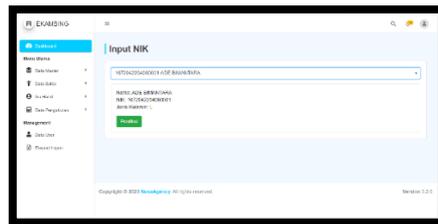
Gambar 4. Tampilan Depan

- 2) Halaman Menu *Dashboard* adalah menu utama yang memuat seluruh fitur di *Website* tersebut



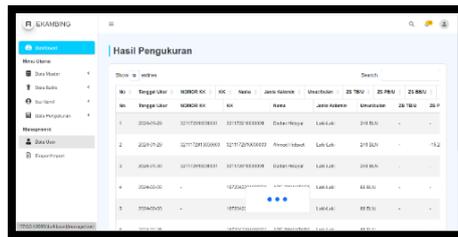
Gambar 5. Menu Dashboard

- 3) Halaman Menu Pengukuran merupakan menu yang digunakan untuk melakukan pengukuran *Stunting*. *Admin* dapat memasukkan data dan melihat hasil dari pengukuran.



Gambar 6. Halaman Pengukuran/Prediksi

- 4) Halaman Menu *Hasil Prediksi* merupakan menu yang berisikan semua data Pengukuran yang ada di dalam *Sistem*. *Admin* dapat melakukan edit, melakukan hapus dan melihat data



Gambar 7. Halaman Hasil Prediksi

5) Pengujian

a. Confusion Matrix

Pada proses *Data Mining* penelitian ini, pengujian dilakukan dengan menggunakan metode pengujian *Confusion Matrix* untuk menemukan dan memastikan bahwa penambangan data yang dihasilkan bebas dari kesalahan atau kegagalan dari keseluruhan.

Berdasarkan hasil pengujian, pengujian *Pred Normal* menghasilkan skor rata-rata 91,67%, pengujian *Pred Obesitas* mendapatkan Skor rata-rata 87,50%, kemudian pengujian *Pred Stunting* skor-rata-rata 100%. Berikut ini tabel pengujian *Confusion Matrix*.

accuracy: 90.00%				
	true Normal	true obesitas	true Stunting	class precision
pred. Normal	11	2	0	84.62%
pred. obesitas	1	14	0	93.33%
pred. Stunting	0	0	2	100.00%
class recall	91.67%	87.50%	100.00%	

Gambar 8. Halaman Hasil Prediksi

b. Blackbox Testing

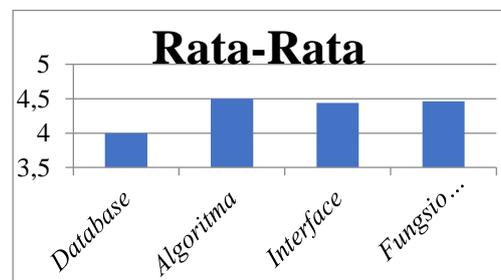
Pada *Sistem* ini, pengujian dilakukan dengan menggunakan metode pengujian *black box* untuk menemukan dan memastikan bahwa *Sistem* yang dihasilkan bebas dari kesalahan atau kegagalan dari keseluruhan *Sistem*.

Berdasarkan hasil kuisioner, pengujian *Database* menghasilkan skor rata-rata 4.0, pengujian *Algoritma* mendapatkan Skor rata-rata 4,5, kemudian pengujian *Interface* skor-rata-rata 4.4 dan pengujian *fungsionalitas* skor rata-rata 4.5. Berikut ini tabel ringkasan kuisioner melalui uji kotak hitam.

Tabel 3. Hasil Pengujian

Pengujian	Rata-Rata
<i>Database</i>	4.00
<i>Algoritma</i>	4.50
<i>Interface</i>	4.44
<i>Fungsionalitas</i>	4.46

Di bawah ini adalah bagan ringkasan kuisioner melalui uji kotak hitam.



Gambar 9. Halaman Hasil *Prediksi*

Di bawah ini adalah tabel Kriteria Kelayakan menurut [16]

Tabel 4. Tabel Kriteria kelayakan

No	Kategori	Tingkat Penilaian
1.	Sangat Layak	85 – 100%
2.	Layak	75 – 84%
3.	Cukup Layak	60 – 74%
4.	Tidak Layak	40 – 59%
5.	Sangat Tidak Layak	0 – 39%

Berdasarkan hasil grafik diatas menunjukkan bahwa Penerapan Metode *Algoritma C4.5* memiliki nilai skor rata-rata 4.00 dan untuk total skor mendapatkan total 4.35 sangat layak untuk diimplementasikan.

4. KESIMPULAN

Berdasarkan analisis perancangan dan pengujian yang penulis lakukan maka penulis dapat mengambil kesimpulan dari “Penerapan *Algoritma C4.5* Untuk *Prediksi* Anak *Stunting* Di Kota Pagar Alam” sebagai berikut: Penggunaan *Algoritma* sangat dianjurkan untuk mem*Prediksi* status gizi Anak dengan menggunakan metode *Algoritma C4.5* di kota Pagar Alam. karena dapat menghasilkan pohon keputusan dan sangat efisien dalam menangani atribut diskrit dan numerik, serta menghasilkan akurasi yang dapat diterima. *Sistem Prediksi* status gizi Anak yang dirancang penulis layak karena dapat mengkategorikan status gizi balita. balita secara otomatis berdasarkan *Zscore* yang ditetapkan dan hanya terdapat selisih 11,8% dari pengujian *Prediksi* dataset yang sama menggunakan *Rapid Miner*. *Sistem* yang penulis rancang dapat lebih cepat dan efektif dalam mengklasifikasikan status gizi Anak. Berdasarkan Berdasarkan data hasil uji *Sistem*, dapat disimpulkan bahwa akurasi *Algoritma C.45* untuk

memPrediksi anak *Stunting* yaitu 88,20% tergolong baik. Diharapkan bagi penelitian selanjutnya untuk menggunakan metode lain dalam menerapkan klasifikasi kasus status gizi Anak, selain itu juga diharapkan kedepannya dapat dikembangkan ke dalam versi aplikasi android/ ios.

UCAPAN TERIMAKASIH

Peneliti mengucapkan terimakasih kepada Dinas Kesehatan kota Pagar Alam yang sudah mengizinkan melakukan penelitian disana dan diizinkan dalam meminta data Balita, lalu dan tak lupa lagi terimakasih juga untuk Institut Teknologi Pagar Alam beserta dosen pembimbing yang memberikan dukungan untuk menyelesaikan tugas akhir dengan lancar.

REFERENCES

- [1] R. A. Anggraini and A. J. Mulyana, "Implementasi Metode Prototyping Pada Rancang *Sistem* Informasi Kesehatan Gizi Berbasis Mobile di Kota Tasikmalaya," vol. 9, no. 1, pp. 9–18, 2022.
- [2] A. Prasetio, "Simulasi Penerapan Metode Decision Tree (C4.5) Pada Penentuan Status Gizi Balita," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 4, no. 3, pp. 209–214, 2021, doi: 10.32672/jnkti.v4i3.2983.
- [3] M. Mahpuz, A. Muliawan Nur, and L. M. Samsu, "Penerapan Algoritma C4.5 Dalam Mengklasifikasi Status Gizi Balita Pada Posyandu Desa Dames Damai Kabupaten Lombok Timur," *Infotek J. Inform. dan Teknol.*, vol. 5, no. 1, pp. 72–81, 2022, doi: 10.29408/jit.v5i1.4414.
- [4] J. Ilmiah and W. Pendidikan, "3 1,2,3," vol. 8, no. July, pp. 116–125, 2022.
- [5] H. S. Tambunan, Z. A. Siregar, A. P. Windarto, and F. Rizki, "Penerapan Data Mining Klasifikasi Gizi Bayi Dengan Algoritma," vol. 1, no. 2, pp. 88–96, 2022.
- [6] S. Mujiyono and N. Syaiful, "*Prediksi* Status Gizi Balita Menggunakan C4 . 5 Algoritma," vol. 8984, no. 1, pp. 41–52, 2023.
- [7] F. Adzim *et al.*, "KLASIFIKASI STATUS STUNTING BALITA MENGGUNAKAN METODE C4 . 5 BERBASIS WEB," vol. 5, no. 3, pp. 515–525, 2023.
- [8] D. Memberikan, P. Prima, P. Wisatawan, D. I. Desa, and W. Pentingsari, "Jurnal Administrasi Bisnis Terapan (JABT)," vol. 2, no. 2, 2020, doi: 10.7454/jabt.v2i2.1071.
- [9] F. Revita, "Penerapan Penilaian Kinerja Karyawan Pada Politeknik Kutaraja," *J. Kreasi Rakyat*, vol. 1, no. 1, pp. 25–035, 2023, [Online]. Available: <https://jurnal.kreyatcenter.com/index.php/jkr/index>
- [10] N. S. Mianti, S. Hartati, and M. Arafat, "Membuat Website Uptd Puskesmas Batumarta Ii Menggunakan Php & Mysql," *J. Inform. dan Komput.*, vol. 14, no. 1, pp. 82–88, 2023, [Online]. Available: <https://garuda.ristekbrin.go.id/documents/detail/1064467>
- [11] A. Parkhan and M. Sugarindra, "... Kualitas Mekanis Kain Tenun Menggunakan Metode Vikor Optimal Design of Woven Fabric Mechanical Quality Using Vikor ...," *J. Disprotek*, vol. 13, no. 2, pp. 137–145, 2022, doi: 10.34001/jdpt.v12i2.
- [12] C. Algoritma, U. *Prediksi*, and B. Siswa, "BERDASARKAN TIPOLOGI HIPPOCRATES-GALENUS," vol. 11, no. 1, pp. 1–8, 2023.
- [13] A. Wahyudi and S. Mujilahwati, "Implementasi Metode Time Series untuk *Prediksi* dan Monitoring Pendapatan Masla Delivery Berbasis Website," vol. 8, no. 2, pp. 1–6, 2023.
- [14] R. Wirth and J. Hipp, "CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39," *Proc. Fourth Int. Conf. Pract. Appl. Knowl. Discov. Data Min.*, no. 24959, pp. 29–39, 2000, [Online]. Available: https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining
- [15] H. Asyraf and M. E. Prasetya, "Implementasi Metode CRISP DM dan Algoritma Decision Tree Untuk Strategi Produksi Kerajinan Tangan pada UMKM A," vol. 8, pp. 94–105, 2024, doi: 10.30865/mib.v8i1.7050.
- [16] D. M. Pithri, N. Safina, and R. Soraya, "Kelayakan Media Pembelajaran Berbasis Android Menggunakan Ispring Suite Pada Teks Prosedur Kelas XI SMAS Istiqlal Deli Tua," vol. 4, pp. 5034–5044, 2024.