

Implementation of C4.5 Algorithm in Classification of Toddler Nutritional Status at Ujan Mas Health Center, Kepahiang Regency

Penerapan Algoritma C4.5 Pada Klasifikasi Status Gizi Balita Di Puskesmas Ujan Mas Kabupaten Kepahiang

Tia Ajeng Putri*, Murlena, Putri Milanda Binamus

Program Studi Ilmu Komputer, Fakultas Teknik, Universitas Pat Petulai, Bengkulu, Indonesia

*E-mail korespondensi: ¹tiaajengputi@icloud.com, ²ana.murlena@gmail.com,
³putrimilanda.bainamus@gmail.com

Abstract

Data Mining and the C4.5 Algorithm method, so that the calculation of the nutritional status of toddlers used to determine nutritional status can be easily calculated and the C4.5 Algorithm is suitable for class training and formation methods. decision tree or calculation tree. From the problem above, the author is interested in raising a research topic, namely: "Application of the C4.5 Algorithm in the Classification of Nutritional Status of Toddlers at the Ujan Mas Health Center, Kepahiang Regency". This research method is the basis for this research using data mining techniques with the C4.5 algorithm. The second step is to calculate the total data consisting of (48) and the total data based on the information Yes 24 and No 24. Then group and calculate all the total attributes stored in the "Amount" column and the total "Attribute" based on the nutritional status of the toddler which is stored in the "Yes" and "No" columns. Based on the rapidminer application, it gets an accuracy value of 60.87%

Keywords: Data Mining, Nutritional Status of Toddlers, C4.5 Algorithm, Classification

Abstrak

Data Mining dan metode Algoritma C4.5, dengan begitu penghitungan status gizi balita yang dipakai untuk penentuan status gizi dapat mudah di perhitungkan dan Algoritma C4.5 cocok untuk metode cllass trainig dan pembentukan decision tree atau pohon perhitungan. Dari masalah diatas penulis tertarik untuk mengangkat suatu topik penelitian yaitu : "Penerapan Algoritma C4.5 Pada Klasifikasi Status Gizi Balita di Puskesmas Ujan Mas Kabupaten Kepahiang". Metode penelitian ini menjadi dasar acuan dalam penelitian ini menggunakan teknik data mining dengan algoritma C4.5 Langkah ke dua adalah menghitung total data yang terdiri atas (48) dan total data berdasarkan Keterangan Ya 24 dan Tidak 24. Kemudian mengelompokkan dan menghitung semua total atribut disimpan dalam kolom "Jumlah" dan total "Atribut" berdasarkan dari status gizi balita yang disimpan dalam kolom "Ya" dan "Tidak". Berdasarkan Dari aplikasi rapidminer mendapatkan nilai akurasi 60,87 %

Kata Kunci : Data Mining , Status Gizi Balita , Algoritma C4.5, Klasifikasi

PENDAHULUAN

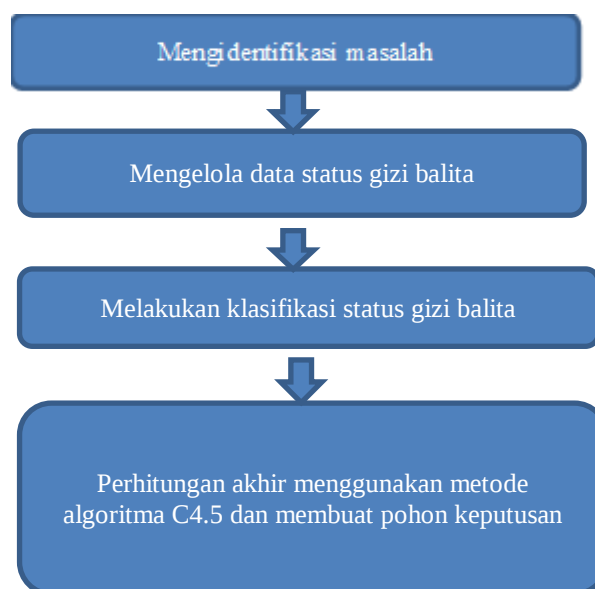
Pertumbuhan dan perkembangan anak menjadi salah satu hal yang sangat penting bagi para orang tua dalam memperhatikan masalah gizi pada anak. Pola makan juga mempengaruhi keadaan gizi balita. Gizi merupakan asupan terpenting untuk tumbuh kembang anak, karena gizi yang baik membantu anak untuk dapat berkembang secara normal. Gizi balita sangat penting untuk mengetahui kondisi balita sejak dini, agar penanggulangan gizi pada balita dapat teratasi.

Jumlah balita yang semakin banyak dan perhitungan status gizi balita masih dilakukan secara manual yang membutuhkan waktu yang sangat lama. Dari data status gizi balita tersebut Puskesmas Ujan Mas membutuhkan pengolahan data status gizi balita untuk menemukan informasi baru yang berguna sebagai pembantu pengambilan keputusan, dalam hal ini yaitu perhitungan status gizi balita sehingga petugas Posyandu Puskesmas Ujan Mas dapat mengetahui status gizi balita tersebut.

Berdasarkan hal tersebut, Puskesmas Ujan Mas membutuhkan suatu alat bantu dalam pengolahan data untuk perhitungan status gizi balita yang dapat memudahkan petugas dalam menghitung status gizi pada balita. Aplikasi yang nantinya diterapkan di Puskesmas Ujan Mas ini merupakan aplikasi penentuan status gizi pada balita, dapat dikategorikan menjadi 5 bagian status gizi yaitu gizi normal, gizi kurang, beresiko gizi lebih dan obesitas. Jadi perlu adanya pengklasifikasian status gizi balita yang menentukan gizi balita.

METODE PENELITIAN

Pada penelitian ini membentuk kerangka pikir sehingga menjadi gambaran bagaimana peneliti menerapkan langkah-langkah pembuktian sampai memperoleh hasil yang menjadi keputusan. Adapun kerangka pikir penelitian ini yaitu:



Gambar 1 Kerangka Pemikiran

Keterangan :

1. Kelola data adalah proses pengumpulan, penyimpanan, pengolahan dan menganalisis data.
2. Melakukan klasifikasi adalah mengelompokkan status gizi pada balita di Puskesmas Ujan Mas.
3. Perhitungan akhir menggunakan Algoritma C4.5 dan membuat pohon keputusan adalah tahap untuk mengetahui apakah Algoritma C4.5 yang digunakan sudah sesuai untuk digunakan pada penelitian sehingga bisa dapat menentukan status gizi balita di Puskesmas Ujan Mas.

HASIL DAN PEMBAHASAN

Menghitung total data yang terdiri atas (48) dan total data berdasarkan Keterangan Ya =24 dan Tidak = 24. Kemudian mengelompokkan dan menghitung semua total atribut disimpan dalam kolom "Jumlah" dan total "Atribut" berdasarkan dari banyaknya status gizi balita yang terdapat dalam kolom "Ya" dan "Tidak". seperti tabel data set dibawah ini:

Tabel 1 data set

NO	JK	BB	TB	JUMLAH	YA/TIDAK
1	L	8	65	4	Tidak
2	L	11	81	8	Tidak
3	L	12	82	6	Tidak
4	L	13	88	10	Tidak
5	L	16	95	2	Tidak
6	L	17	100	1	Tidak
7	L	10	75	10	Tidak
8	L	14	90	3	Tidak
9	L	15	92	1	Tidak
10	L	6	58	1	Tidak
11	L	7	64	3	Tidak
12	L	9	72	7	Tidak
13	L	22	100	1	Tidak
14	L	4	55	2	Ya
15	L	5	59	3	Ya
16	L	6	66	8	Ya
17	L	7	69	11	Ya
18	L	8	77	21	Ya
19	L	9	78	22	Ya
20	L	10	85	26	Ya
21	L	11	88	18	Ya
22	L	12	95	12	Ya
23	L	13	95	11	Ya
24	L	14	98	8	Ya
25	P	6	58	4	Tidak
26	P	7	64	8	Tidak
27	P	11	81	10	Tidak
28	P	12	83	3	Tidak
29	P	10	76	4	Tidak
30	P	13	88	4	Tidak
31	P	15	94	2	Tidak
32	P	4	52	1	Tidak
33	P	8	68	2	Tidak
34	P	9	70	3	Tidak
35	P	14	86	1	Tidak
36	P	3	52	1	Ya
37	P	4	57	4	Ya
38	P	5	62	6	Ya
39	P	6	66	16	Ya
41	P	7	69	9	Ya
42	P	8	78	21	Ya
43	P	9	84	21	Ya
44	P	10	87	17	Ya
45	P	11	94	14	Ya
46	P	12	97	9	Ya
47	P	13	95	5	Ya
48	P	14	97	6	Ya

Selanjutnya yaitu menghitung total data yang terdiri atas (48) dan total data berdasarkan Keterangan Ya =24 dan Tidak = 24. Kemudian mengelompokan dan menghitung semua total atribut

disimpan dalam kolom “Jumlah” dan total “Atribut” berdasarkan dari banyaknya status gizi balita yang terdapat dalam kolom “Ya” dan “Tidak”. Hasil pengelompokan dapat dilihat pada table dibawah ini:

Tabel 2 total data, atribut dan kelas

Keterangan		jumlah kasus (s)	tidak (s1)	ya (s2)
total		48	24	24
jenis kelamin				
	laki-laki	24	13	11
	Perempuan	24	11	13
berat badan				
	3-10 kg	26	11	15
	11-22 kg	22	13	9
tinggi badan				
	52-70 cm	17	8	9
	71-90 cm	18	11	7
	91-100 cm	13	5	8

Proses Pembentukan Decesion Tree Menggunakan Algoritma C4.5

Berikut ini merupakan langkah selanjutnya yaitu mencari *Entropy* total

Rumus *Entropy* Total

$$Entropy (total) : Entropy (s) = \sum -p * \log^2 p_i$$

$$Entropy Total: (-24/48 * \log^2 (24/48)) + (-24/48 * \log^2 (24/48)) = 1$$

Langkah selanjutnya yaitu : Mencari *Entropy* katagori (Jenis kelamin, Berat badan, Tinggi badan, jumlah, Ya, Tidak). dari masing-masing atribut dihitung berdasarkan dari jumlah kelas (keterangan Ya dan Tidak).

$$Jk \text{ laki-laki: } (-11/24 * \log^2 (11/24)) + (-13/24 * \log^2 (13/24)) = 0,994985$$

$$Jk \text{ perempuan: } (-13/24 * \log^2 (13/24)) + (-11/24 * \log^2 (11/24)) = 0,994985$$

$$BB \text{ 3-10 kg: } (-15/26 * \log^2 (15/26)) + (-11/26 * \log^2 (11/26)) = 0,982859$$

$$BB \text{ 11-22 kg: } (-9/22 * \log^2 (9/22)) + (-13/22 * \log^2 (13/22)) = 0,976021$$

$$TB \text{ 52-70 cm: } (-9/17 * \log^2 (9/17)) + (-8/17 * \log^2 (8/17)) = 0,884115122$$

$$TB \text{ 71-90 cm: } (-7/18 * \log^2 (7/18)) + (-11/18 * \log^2 (11/18)) = 0,964079$$

$$TB \text{ 91-100 cm: } (-8/13 * \log^2 (8/13)) + (-5/13 * \log^2 (5/13)) = 0,961237$$

Perhitungan Gain:

$$\text{Jenis kelamin : } 1 - (24/48 * 0,994985) + (24/48 * 0,994985) = 1$$

$$\text{Berat badan : } 1 - (26/48 * 0,982859) + (22/48 * 0,976021) = 0,914961$$

$$\text{Tinggi badan: } 1 - (17/48 * 0,997503) + (18/48 * 0,964079) + (13/48 * 0,961237) = 1,2685823$$

Tabel 3 nilai hasil perhitungan *node* 1.1

Keterangan		jumlah kasus (s)	tidak (s1)	ya (s2)	entropy	gain
total		48	24	24	1	
jenis kelamin						
	laki-laki	24	13	11	0,994985	1
	perempuan	24	11	13	0,994985	
berat badan						
	3-10 kg	26	11	15	0,982859	0,914961
	11-22 kg	22	13	9	0,976021	
tinggi badan						
	52-70 cm	17	8	9	0,997503	1,2685823
	71-90 cm	18	11	7	0,964079	
	91-100 cm	13	5	8	0,961237	

Tinggi Badan

57-70 cm 71-90 cm 91-100 cm

Gambar 1 Pohon Keputusan Node 1.1

Entropy Total: $(-7/12 * \text{Log}^2 (7/12)) + (-5/12 * \text{Log}^2 (5/12))=0,979869$

JK : $(-3/7 * \text{Log}^2 (3/7)) + (-4/7 * \text{Log}^2 (4/7))=0,985228$

Jk perempuan: $(-4/5 * \text{Log}^2 (4/5)) + (-1/5 * \text{Log}^2 (1/5))=0,-2,06439$

BB 3-10 kg :0

BB 11-22 kg : $(-7/12 * \text{Log}^2 (7/12)) + (-5/12 * \text{Log}^2 (5/12))$
=0,979869

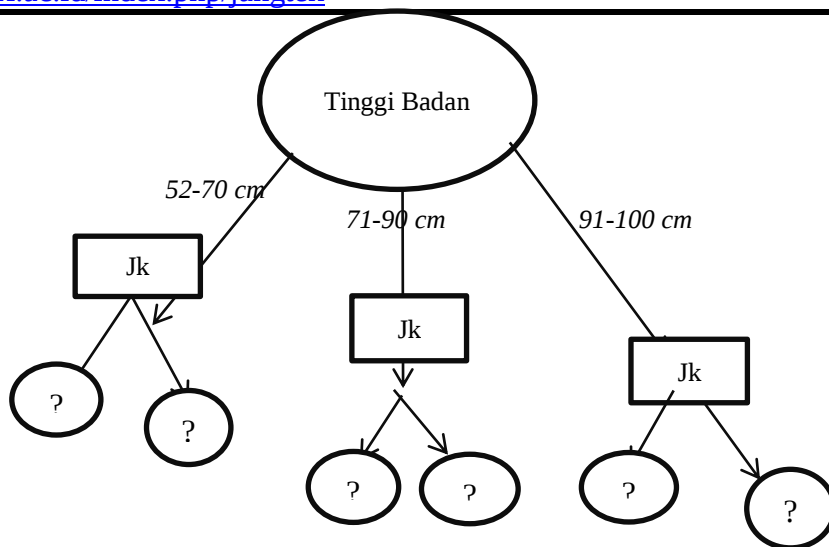
Gain:

JK : $0,979869 - (7/12 * 0,985228) + (5/12 * -2,06439)=-1,276033$

BB $0,979869 - (12/12 * 0,979869)= 0$

Tabel 4 Hasil Perhitungan Tinggi Badan 9-100 cm

Keterangan		jumlah	Tidak	Ya	entropy	gain
total		12	5	7	0,979869	
jenis kelamin						
	laki-laki	7	4	3	0,985228	-1,276033
	perempuan	5	1	4	-2,06439	
berat badan						
	3-10 kg	0	0	0	0	0
	11-22 kg	12	5	7	0,979869	



Gambar 2 Pohon Keputusan Node 1.2

Entropy Total: $(-2/4 * \log^2 (2/4)) + (-2/4 * \log^2 (2/4)) = 1$

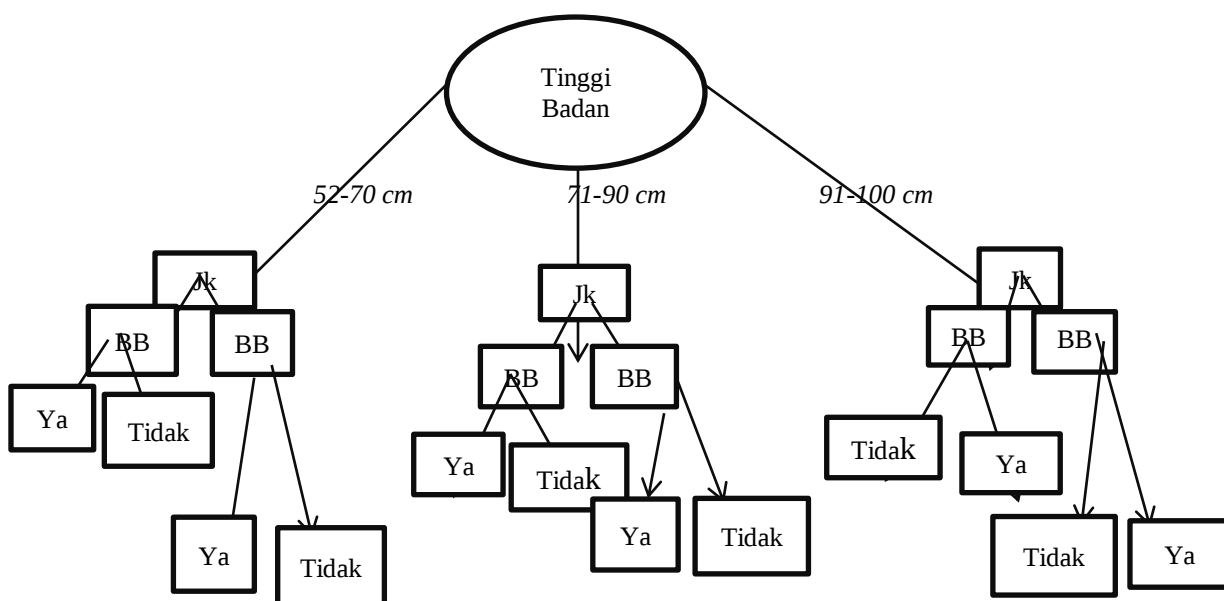
BB 11 kg = $(-2/4 * \log^2 (2/4)) + (-2/4 * \log^2 (2/4)) = 1$ BB 22 kg = 0

Gain :

BB 1 - $(2/4 * 1) + 0 = 0,5$

Tabel 5 hasil perhitungan berat badan 11- 22 kg

Keterangan		jumlah	Tidak	Ya	entropy	gain
total		4	2	2	1	
BB 11-22 Kg						
	BB 3-10 Kg	0	0	0	0	
	BB 11 Kg	4	2	2	1	0,5
	BB 22 Kg	0	0	0	0	



Gambar 3 Pohon Keputusan Node 1.3

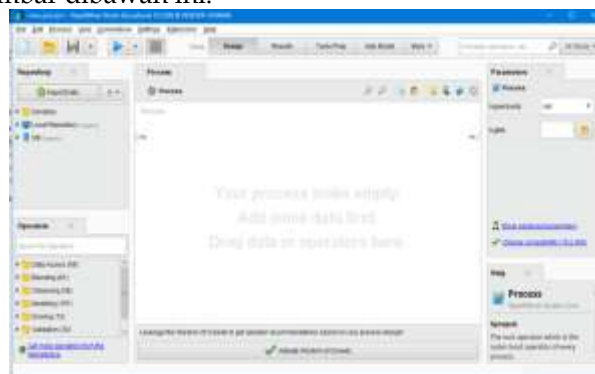
Tahapan Implementasi *Decision tree*

Berikut merupakan langkah-langkah menentukan pohon keputusan pada aplikasi rapidminer.

Tampilan Awal Aplikasi *Rapidminer*



Setelah beberapa tahapan sudah dilakukan pada rapidminer maka mendapatkan hasil pohon keputusan seperti gambar dibawah ini:



Serta Menampilkan hasil akurasi sebesar 60,87%



KESIMPULAN

Kesimpulan yang dapat diperoleh dari penelitian ini yaitu mengacu pada rumusan masalah, maka dapat disimpulkan: Penerapan Algoritma C4.5 untuk mengklasifikasi status gizi balita menggunakan 4 atribut yang digunakan untuk menentukan status gizi balita adalah jenis kelamin, berat badan, tinggi badan dan status gizi. Diolah menggunakan Algoritma C4.5 untuk mendapatkan hasil status gizi balita dengan proses sebagai berikut: Menentukan atribut yang digunakan, Menghitung jumlah kasus dari semua data pada setiap

atribut, Membuat data set, Menghitung nilai *entropy*, Menghitung nilai *gain*, Menentukan *tree* dari nilai *gain* tertinggi, Menentukan *decision tree*, Hasil akurasi yang diperoleh dari pengujian menggunakan aplikasi *rapidminer* sebesar 60,87% Algoritma C4.5 ini cocok digunakan untuk menghitung klasifikasi status gizi balita.

DAFTAR PUSTAKA

- Akhiri, S., & Windarto, T. (2019). Data mining: Konsep, teknik, dan aplikasinya. Yogyakarta: Pustaka Pelajar.
- Almatsier, S. (2005). Prinsip Dasar Ilmu Gizi. Jakarta: Gramedia Pustaka Utama.
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques. Elsevier.
- Hutasoit, A. S., Tarigan, P., & Siagian, E. R. (2018). Implementasi Data Mining Klasifikasi Status Gizi Balita Pada Posyandu Medan Timur Dengan Menggunakan Metode C 4.5. *Pelita Informatika: Informasi dan Informatika*, 7(2), 120-125.
- Irianto, Djoko Pekik. (2007). Panduan Gizi Lengkap Keluarga dan Olahragawan. Yogyakarta: Penerbit Andi.
- Islam, Hajar Izzatul, Muhamad Khandava Mulyadien, and Ultach Enri. "Penerapan Algoritma C4. 5 dalam Klasifikasi Status Gizi Balita." *Jurnal Ilmiah Wahana Pendidikan* 8.10 (2022): 116-125.
- Larose & larose, 2020. Aplikasi sistem pendukung keputusan menggunakan algoritma c4.5. Untuk penjurusan sma. Dalam jurnal simantec. Vol. 5, no. 2 juni 2016
- Lestari, S., & Amalia, R. A. (2023). Penerapan Algoritma C. 45 Pada Klasifikasi Status Gizi Balita di Posyandu Desa Sukalilah Cibatuk Kabupaten Garut Jawa Barat. *Jurnal Sains dan Teknologi*, 5(1), 177-182.
- Mahpuz, M., Nur, A. M., & Samsu, L. M. (2022). Penerapan Algoritma C4. 5 Dalam Mengklasifikasi Status Gizi Balita Pada Posyandu Desa Dames Damai Kabupaten Lombok Timur. *Infotek: Jurnal Informatika dan Teknologi*, 5(1), 72-81.
- Moertini, Y. (2007). Data Mining: Konsep dan Implementasi. Yogyakarta: Andi.
- Muaris, A. (2006). Ilmu Gizi Anak dan Remaja. Jakarta: Universitas Indonesia.
- Murlena, D. A., & Apriana, D. (2022). Penerapan Data Mining Untuk Memprediksi Ketersediaan Stok Produk HNI HPAI Menggunakan Algoritma C4. 5. *Arcitech: Journal of Computer Science and Artificial Intelligence*, 2(1), 19-32.
- Murlena, M., & Syahindra, W. (2020). Data Mining Pengolahan Penempatan Library Books Menggunakan Metode Association Rule Dengan Algoritma Apriori. *Jurnal INSTEK (Informatika Sains Dan Teknologi)*, 5(2), 199-208.
- Nuraisana, N., windarti Halawa, S., & Harun, M. (2024). Implementasi Algoritma C4. 5 Dalam Mengklasifikasi Status Gizi Balita Pada Posyandu Desa Sekip Lubuk Pakam. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Infor*
- Purnomo, J., Sukemi, S., Parwito, P., & Ermatita, E. (2022). Implementation of Fuzzy C-Means and Topsis in College Rankings. *Journal of Information Systems and Informatics*, 4(4), 1094-1111. <https://doi.org/10.51519/journalisi.v4i4.409>
- Rerung, S. (2018). Data Mining: Konsep dan Implementasi Algoritma. CV Pustaka Mandiri.
- Supariasa, D. A. (2002). Ilmu Gizi dan Diet. Gramedia Pustaka Utama.
- Sutomo, B. (2010). Menu Sehat Alami Untuk Balita & Batita. PT. Agromedia Pustaka.