



Penerapan Algoritma Naïve Bayes dan Aplikasi RadipMiner Untuk Penentuan Status Gizi Balita

Jon Idrison Molina¹

¹Teknik Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Tribuana Kalabahi

jhonmolina8788@gmail.com

Abstract

Determining the nutritional status of toddlers is an important activity carried out by medical personnel at the health center as technical implementers. Because with the results of determining the nutritional status, intervention efforts to overcome the condition of toddlers who are in a status other than good nutrition can be overcome as early as possible. However, medical personnel often have difficulty in determining the nutritional status of toddlers due to limited human resources and equipment, so it is important to have a method used to determine the nutritional status of children under five years old quickly and accurately. For this reason, in this study we used the naïve Bayes algorithm in calculating the classification of determining the nutritional status of toddlers. The results of the study showed that the naïve Bayes algorithm can increase the level of data accuracy in determining the nutritional status of toddlers with a data accuracy value of 70.33%.

Keywords: *Implementation, Naive Bayes, Status, Nutrition, Toddlers*

Abstrak

Penentuan status gizi balita merupakan suatu kegiatan yang penting dilakukan oleh tenaga medis di puskesmas sebagai pelaksana teknis. Sebagaimana yang terjadi pada dua Puskesmas di Kabupaten Alor yaitu Puskesmas Mademang dan Puskesmas Lantoka sebagai lokasi penelitian. Karena dengan adanya hasil penentuan status gizi ini maka upaya intervensi untuk mengatasi kondisi balita yang ada pada status selain gizi baik dapat diatasi sedini mungkin. Namun tenaga medis sering mengalami kesulitan dalam penentuan status gizi balita karena keterbatasan sumber daya manusia maupun peralatan sehingga penting adanya sebuah metode yang digunakan dalam penentuan status gizi anak dibawah lima tahun secara cepat dan akurat. Untuk itu dalam penelitian ini kami menggunakan algoritma naïve bayes dalam perhitungan klasifikasi penentuan status gizi balita. Hasil penelitian menunjukkan bahwa dengan algoritma naïve bayes dapat meningkatkan tingkat keakurasian data dalam penentuan status gizi balita dengan nilai akurasi datanya sebesar 70.33 %.

Kata kunci : *Penerapan, Naive Bayes, Status, Gizi, Balita*

1. Pendahuluan

Balita adalah anak yang berada pada usia lima tahun kebawah, usia ini merupakan usia emas anak atau dikenal dengan istilah *golden age*. Keberhasilan pertumbuhan dan perkembangan anak akan ditentukan pada masa-masa ini[1].

Sesuai data Badan Pusat Statistik (BPS) Provinsi NTT yang dirilis pada tanggal 20 Mei 2024 di websitenya, tergambar data balita dengan status gizi buruk 3 tahun terakhir, pada tahun 2021 sebanyak 232 anak, pada tahun 2022 sebanyak 310 anak, dan pada tahun 2023 sebanyak 275 anak. Untuk status gizi kurang pada

tahun 2021 sebanyak 1.096 anak, pada tahun 2022 sebanyak 1.514 anak, dan pada tahun 2023 sebanyak 1.575 anak. Anak dengan status gizi baik pada 2 tahun terakhir pada tahun 2022 sebanyak 15.702 anak, dan pada tahun 2023 sebanyak 15.406 anak. Sedangkan anak dengan status gizi lebih pada 2 tahun terakhir antara lain pada tahun 2022 sebanyak 113 anak, dan pada tahun 2023 sebanyak 95 anak. Dari data ini tergambar kalau data status gizi kurang dan gizi buruk masih tinggi. Untuk itu salah satu langkah yang cukup signifikan dalam meningkatkan kualitas perkembangan anak adalah dengan memperbaiki status gizi. Perbaikan status gizi sangat penting, terutama bagi balita yang



Lisensi

Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

sedang berada dalam periode emas pertumbuhan. Untuk mencegah masalah gizi pada balita, diperlukan penilaian status gizi secara rutin guna mendeteksi kondisinya sejak dini. Penilaian status gizi balita bertujuan untuk mengetahui kondisi gizinya. Untuk memperoleh informasi tersebut, diperlukan evaluasi status gizi secara berkala. Dengan adanya penentuan status gizi anak dibawah lima tahun yang dilakukan secara tepat dan akurat memungkinkan deteksi dini terhadap gangguan pertumbuhan dan perkembangan, sehingga dapat segera dilakukan tindakan yang diperlukan. Penilaian status gizi bayi biasanya dilakukan dengan mengacu pada tabel antropometri yang diterbitkan oleh Kementerian Kesehatan Republik Indonesia.

Jika data gizi bayi dalam jumlah yang banyak maka penentuan status gizi menggunakan tabel antropometri akan membutuhkan waktu yang lebih lama. Agar permasalahan ini teratasi maka diperlukan sebuah sistem yang lebih cepat dan akurat dalam penentuan status gizi anak dengan menggunakan metode klasifikasi. Metode klasifikasi dalam penelitian ini dapat digunakan untuk menentukan status gizi balita di masa mendatang, karena metode ini memanfaatkan data training dalam pengambilan keputusan serta mampu

menyesuaikan berbagai parameter penentuan status gizi, sehingga menghasilkan hasil yang lebih akurat.[2].

2. Metode Penelitian

Metode yang digunakan dalam penelitian ini dapat diuraikan di bawah ini :

A. Variabel Penelitian

Terdapat lima variabel yang digunakan meliputi :

- a. Jenis Kelamin
 Jenis kelamin adalah atribut mengenai informasi tentang jenis kelamin balita
- b. Umur
 Umur adalah atribut yang berisi informasi tentang umur balita
- c. Tinggi Badan
 Tinggi badan adalah atribut yang berisi informasi tentang tinggi badan balita
- d. Berat Badan
 Berat badan adalah atribut yang berisi informasi tentang berat badan balita
- e. Remark
 Remark adalah atribut yang berisi nilai akhir penentuan status gizi balita.

Dalam proses pengelompokan ini, digunakan lima atribut, dan parameter dari masing-masing atribut dapat dilihat pada Tabel 1.

Tabel 1. Atribut dan nilai

No	Atribut	Nilai	Nilai Baru	No	Atribut	Nilai	Nilai Baru
1	Jenis Kelamin	Laki-laki	L	4	Tinggi Badan	< 44.2	A
		Perempuan	P			44.3-48.0	B
2	Umur (Bulan)	0-2	A			48.1-52.0	C
		3-5	B			52.1-56.0	D
		7-9	C			56.1-60.0	E
		10-12	D			60.1-64.0	F
		13-15	E			64.1-68.0	G
		16-18	F			68.1-72.0	H
		19-21	G			72.1-76.0	I
		22-24	H			76.1-80.0	J
		25-27	I	80.1-84.0	K		
		28-30	J	84.1-90.0	L		
31-60	K	90.1-94.0	M				
3	Berat Badan	<2.1	A	5	Remark	94.1-98.0	N
		2.1-4.0	B			98.1-102.0	O
		4.1-6.0	C			102.1-106.0	P
		6.1-8.0	D			106.1-110.0	Q
		8.1-10.0	E			110.1-114.0	R
		10.1-12.0	F			114.1-120.0	S
		12.1-14.0	G			120.1-124.0	T
		14.1-16.0	H			> 124.0	U
		16.1-18.0	I			Gizi Buruk	GBu
		18.1-20.0	J			Gizi Kurang	GKu

	20.1-22.0	K		Gizi Baik	GBa
	22.1-24.0	L		Gizi Lebih	GLe
	24.1-26.0	M			
	26.1-28.0	N			
	28.1-30.0	O			
	> 30.0	P			

B. Metode Pengumpulan data

a. Studi Pustaka

Dalam penelitian ini data diperoleh melalui studi literatur, artikel penelitian, serta berbagai sumber lain yang relevan.[3].

b. Data sekunder

Selain studi pustaka dalam penelitian ini dilakukan juga melalui pengamatan langsung untuk mengambil data gizi balita dari dua Puskesmas melalui Dinas Kesehatan dalam bentuk fotokopi.

C. Teknik Analisis Data

Proses penemuan pengetahuan dalam basis data bersifat interaktif dan berulang, melibatkan berbagai langkah

serta keputusan yang dibuat oleh pengguna. Proses ini memberikan wawasan praktis dan menekankan sifat interaktifnya, yang terdiri dari sembilan tahap[4].

a. Data Gizi Balita

Menentukan dan menyusun kumpulan data untuk penelitian perlu dilakukan, termasuk mengidentifikasi data yang ada, mengumpulkan data lain yang dibutuhkan dalam penelitian, serta menggabungkan seluruh data ke dalam salah satu bagian yang mencakup atribut yang akan digunakan dalam proses penemuan pengetahuan.

Tabel 2. Data Status Gizi

No.	Nama Balita	J/K	Tanggal Lahir	Umur (Bulan)	Nama Orang Tua		Tinggi Badan(cm)	Berat Badan(Kg)	Nilai Z-Score	Ket.
					Ayah	Ibu				
1	Mika Kago	L	7/12/2019	15	Mesak Kagi	Eti Karfani	77.5	9.2	-1 SD	Gizi Baik
2	Laila Tahira	P	27/5/2019	22	Ahmad Tahira	Elsa Djami	83.0	9.4	-1.84	Gizi Baik
3	Rusti Bain	L	12/4/2019	23	Tomi Bain	Leni Bain	87.0	12.1	0	Gizi Baik
4	Ester Tangko	P	27/4/2019	23	Yanto Tangko	Merry Bana	91.5	16.9	3.42	Gizi Lebih
5	Sila Arangbain	P	9/5/2019	22	Ferdinan Arangbain	Wati Sailana	80.5	8.4	-2.69	Gizi Kurang
6	Aldrian Fadako	L	16/4/2019	23	Yosep Fadako	Nela Langfan	82.5	9.4	-2.07	Gizi Kurang
7	Maxsen Gaimau	L	3/7/2019	20	Samrits Gaimau	Alen Fankari	84.5	10.6	-0.66	Gizi Baik
8	Yabes Timung	L	28/10/2019	15	Yulius Timung	Dina Malmau	78.5	9.5	-0.81	Gizi Baik
9	Eben Letidena	L	9/8/2019	17	Lewi Letidena	Loisa Mohing	76.0	8.4	-2.18	Gizi Kurang
10	Daniel Kranding	L	20/4/2019	23	Jonisius Kranding	Bety Weni	80.5	9.1	-2.30	Gizi Kurang
11	Asri Djaha	P	21/10/2019	17	Eben Djaha	Nelci Banoet	80.5	13.8	2.30	Gizi Lebih
12	Marthen Laubila	L	3/5/2019	22	Yermias Laubila	Selfi Tangadang	87.0	15.6	2.64	Gizi Lebih
13	Oliv Laubesing	P	27/11/2019	16	Nikson Laubesing	Linda Tangel	74.5	7.4	-2.81	Gizi Kurang
14	Joel Lande	L	29/08/2019	19	Jensen Lande	Amalia Adang	80.5	10.5	-0.54	Gizi Baik
15	Sifra Pandu	P	14/09/2019	17	Andrias Lambila	Martha Pandu	80.0	9.6	-1	Gizi Baik
16	Aries Banaweng	L	2/6/2019	21	Lukas Banaweng	Fero Malaimola	79.0	11.2	-0.25	Gizi Baik
17	Jeck Puling	L	16/11/2019	16	Elisa Puling	Mimi Nanggula	75.0	8.1	-2.18	Gizi Kurang

18	Vivi Blegur	P	13/9/2019	15	Roki Blegur	Yunita Liubana	80.0	9.6	-0.63	Gizi Baik
19	Evender Djahadael	L	6/10/2019	15	Aristarkus Djahadael	Magdalena Blegur	80.5	13.7	2.83	Gizi Lebih
20	Betriks Hingkoil	P	2/6/2019	21	Simon Hingkoil	Deliana Maibila	80.0	8.9	-2.16	Gizi Kurang
21	Hermanto Banik	L	27/4/2019	23	Deni Banik	Mincelina Maure	80.5	11.5	-0.38	Gizi Baik
22	Michaela Soares	P	14/7/2019	20	Anderias Soares	Feronika Tangpen	81.5	10.7	-0.5	Gizi Baik
23	Jonkrius Fanpula	L	27/5/2019	22	Herianus Fanpula	Ester Balol	78.5	11.1	-0.53	Gizi Baik
24	Merlin Koilmo	P	13/6/2019	21	Yakob Koilmo	Natalia Tangpen	87.5	15.6	2.92	Gizi Lebih
25	Adrian Laulang	L	26/2/2020	25	Selfister Laulang	Helena Aloma	84.5	9.7	-1.92	Gizi Baik

b. Data Perubahan

Pada bagian ini, data dirancang dan diperluas secara optimal untuk mendukung proses data mining. Selain itu, data yang telah dipilih akan ditransformasikan agar sesuai dengan kebutuhan analisis. Proses ini bersifat inovatif dan juga bergantung pada bentuk informasi yang ingin dieksplorasi dalam basis data[5].

Data awal yang digunakan perlu diubah ke dalam kategori yang sesuai dengan Tabel 1, sehingga terdapat data yang mengalami perubahan sebagaimana tercantum dalam Tabel 3.

Tabel 3. Perubahan Data Status Gizi Balita

Jenis Kelamin	Umur	Berat Badan	Tinggi Badan	Nilai Z-Score	Remark
L	F	E	J	-1 SD	GBa
P	H	E	K	-1.84	GBa
L	H	F	L	0	GBa
P	H	I	M	3.42	GLe
P	H	E	K	-2.69	GKu
L	H	E	K	-2.07	GKu
L	G	F	L	-0.66	GBa
L	F	E	J	-0.81	GBa
L	F	E	I	-2.18	GKu
L	H	E	K	-2.30	GKu
P	F	G	K	2.30	GLe
L	H	H	L	2.64	GLe
P	F	D	I	-2.81	GKe
L	G	F	K	-0.54	GBa
P	F	E	J	-1	GBa
L	G	F	J	-0.25	GBa
L	F	E	I	-2.18	GKu
P	E	E	J	-0.63	GBa
L	E	G	K	2.83	GLe
P	G	E	J	-2.16	GKu

L	H	F	K	-0.38	GBa
P	G	F	K	-0.5	GBa
L	G	F	J	-0.53	GBa
P	H	H	L	2.92	GLe
L	I	E	L	-1.92	GBa

Data yang telah dipersiapkan untuk pengelompokan data dibagi menjadi dua bagian, yaitu data latihan dan data pengujian, dengan proporsi 80% dan 20%. Pembagian ini dilakukan menggunakan teknik *systematic random sampling*. Teknik ini dilakukan dengan pengacakan hanya sekali untuk menentukan elemen pertama dari sampel yang akan diambil. Elemen berikutnya ditentukan berdasarkan interval sampel yang telah ditetapkan.

Interval sampel, yang juga dikenal sebagai *sampling ratio*, dihitung dengan membagi ukuran populasi dengan jumlah sampel yang diinginkan (N/n). Berikut adalah contoh perhitungan untuk pengambilan data testing :

$$\begin{aligned} \text{Jumlah data pengujian} &= 20\% \times 206 = 41 \\ \text{Jumlah populasi (N)} &= 206 \\ \text{Jumlah sampel (n)} &= 41 \\ \text{Interval sampling (k)} &= N/n = 206/41 = 5 \\ \text{Unsur pertama yang diambil untuk data testing (s)} &= 1 \\ \text{Unsur kedua} &= s + k \\ \text{Unsur ketiga} &= s + 2k \\ \text{Unsur keempat} &= s + 3k, \text{ dan seterusnya} \\ &\text{hingga unsur ke-n.} \end{aligned}$$

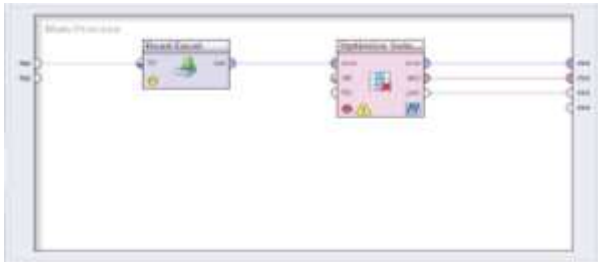
Berdasarkan hasil perhitungan di atas, diperoleh 41 data gizi balita sebagai data testing, sedangkan sisanya, yaitu 165 data gizi balita, digunakan sebagai data training.

c. Eksperimen dan Pengujian Model

Pengujian model dilakukan dengan menggunakan empat variabel, di mana 10% data pelatihan dipilih secara acak sebagai data uji. Proses ini dilakukan sebanyak 10 kali, kemudian hasil pengujian model,

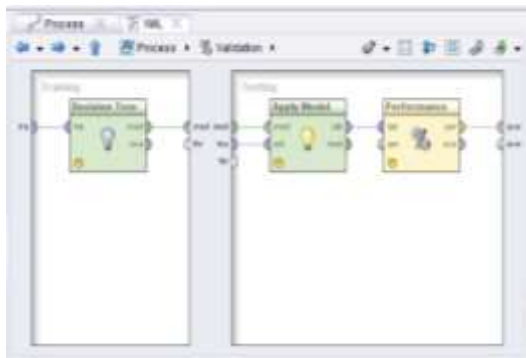
termasuk accuracy, precision, dan recall, dihitung rata-ratanya. Pengujian ini dilakukan menggunakan RapidMiner dengan building block yang berfungsi untuk melakukan prediksi.

Penelitian ini memerlukan eksperimen dan pengujian terhadap model yang diusulkan, yang dilakukan dengan memanfaatkan sebagian dari dataset yang tersedia. Selanjutnya, seluruh dataset diuji menggunakan metode yang telah dirancang dalam aplikasi RapidMiner 5. Berikut adalah model yang diimplementasikan dalam RapidMiner 5:



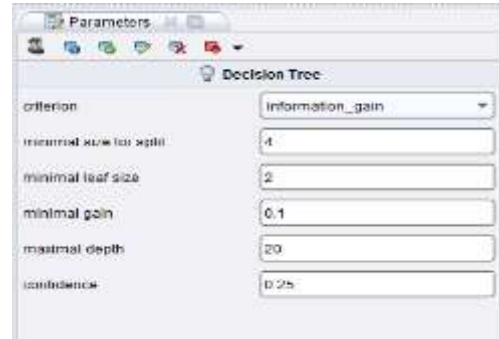
Gambar 1. Model yang diusulkan pada aplikasi RapidMiner 5

Gambar 1. Menunjukkan keterkaitan dengan model yang diusulkan, yaitu proses membaca dataset dari Excel. Selanjutnya, kedua proses *training* dihubungkan dalam tahap validasi untuk memperkirakan kinerja operator. Selanjutnya, hasil validasi pada tahap training diteruskan ke bagian *optimize selection example* sebagai input.



Gambar 2. Metode *cross validation* pada RapidMiner 5

Dataset diproses menggunakan algoritma naïve bayes pada Gambar 2 untuk mengevaluasi performanya. Parameter information gain diatur sesuai dengan yang ditampilkan pada gambar 3 di bawah ini :



Gambar 3. Parameter algoritma naïve bayes pada RapidMiner

Hasil yang didapatkan, dianalisis dan dikomparasi untuk menentukan metode yang paling akurat dalam model yang diusulkan. Perbandingan dilakukan berdasarkan atribut yang digunakan dalam model, nilai AUC, tingkat akurasi, serta performa model yang mencakup f-measure, precision, dan recall.

3. Hasil dan Pembahasan

A. Klasifikasi Algoritma Naïve Bayes

Penerapan algoritma ini dengan data training pada Tabel 3 diawali dengan menghitung probabilitas prior untuk menentukan nilai yang diterima dan tidak diterima berdasarkan data secara keseluruhan. Dalam data training yang terdiri dari 165 balita, terdapat 112 balita dengan gizi baik, 23 dengan gizi kurang, 22 dengan gizi lebih, dan 8 dengan gizi buruk. Hasil perhitungan prior probability dapat dilihat pada Tabel 4.[6].

Tabel 4. Perhitungan Probabilitas *Prior*

Atribut	Nilai Baru	Kasus	GBa	GKu	GBu	GLE	p(x cl)			
							GBa	GKu	GBu	GLE
Total		165	112	23	8	22	0.678787879	0.139393939	0.048484848	0.133333333
Jenis Kelamin	L	89	62	13	4	10	0.696629213	0.146067416	0.146067416	0.112359551
	P	76	50	10	4	12	0.657894737	0.131578947	0.052631579	0.157894737
Umur	A	2	0	0	0	2	0	0	0	1

	B	11	4	0	0	7	0.363636364	0	0	0.636363636
	C	14	9	3	1	1	0.642857143	0.214285714	0.071428571	0.071428571
	D	10	7	2	0	1	0.7	0.2	0	0.1
	E	22	15	4	2	1	0.681818182	0.181818182	0.090909091	0.045454545
	F	16	12	3	0	1	0.75	0.1875	0	0.0625
	G	12	6	3	1	2	0.5	0.25	0.083333333	0.166666667
	H	23	16	5	0	2	0.695652174	0.217391304	0	0.086956522
	I	10	8	0	1	1	0.8	0	0.1	0.1
	J	7	5	1	0	1	0.714285714	0.142857143	0	0.142857143
	K	38	30	2	3	3	0.789473684	0.052631579	0.078947368	0.078947368
Berat Badan	A	0	0	0	0	0	0	0	0	0
	B	0	0	0	0	0	0	0	0	0
	C	2	0	0	2	0	0	0	1	0
	D	21	5	13	2	1	0.238095238	0.619047619	0.095238095	0.047619048
	E	50	36	8	1	5	0.72	0.16	0.02	0.1
	F	41	33	2	1	5	0.804878049	0.048780488	0.024390244	0.12195122
	G	17	15	0	0	2	0.882352941	0	0	0.117647059
	H	17	14	0	0	3	0.823529412	0	0	0.176470588
	I	10	7	0	0	3	0.7	0	0	0.3
	J	4	1	0	1	2	0.25	0	0.25	0.5
	K	1	1	0	0	0	1	0	0	0
	L	0	0	0	0	0	0	0	0	0
	M	2	0	0	1	1	0	0	0.5	0.5
	N	0	0	0	0	0	0	0	0	0
	O	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	
Tinggi Badan	A	0	0	0	0	0	0	0	0	0
	B	0	0	0	0	0	0	0	0	0
	C	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0
	E	1	0	0	1	0	0	0	1	0
	F	3	1	1	0	1	0.333333333	0.333333333	0	0.333333333
	G	11	5	2	1	3	0.454545455	0.181818182	0.090909091	0.272727273
	H	16	12	2	1	1	0.75	0.125	0.0625	0.0625
	I	20	7	6	1	6	0.35	0.3	0.05	0.3
	J	24	17	6	1	0	0.708333333	0.25	0.041666667	0
	K	19	13	3	1	2	0.684210526	0.157894737	0.052631579	0.105263158
	L	28	24	2	0	2	0.857142857	0.071428571	0	0.071428571
	M	12	8	0	2	2	0.666666667	0	0.166666667	0.166666667
	N	9	7	1	0	1	0.777777778	0.111111111	0	0.111111111
	O	9	8	0	0	1	0.888888889	0	0	0.111111111
P	9	7	0	0	2	0.777777778	0	0	0.222222222	
Q	3	3	0	0	0	1	0	0	0	

	R	0	0	0	0	0	0	0	0	0
	S	1	0	0	0	1	0	0	0	1
	T	0	0	0	0	0	0	0	0	0
	U	0	0	0	0	0	0	0	0	0

Untuk menentukan apakah suatu kasus baru termasuk dalam kategori gizi baik, gizi kurang, gizi buruk, atau gizi lebih, dilakukan perhitungan probabilitas posterior berdasarkan probabilitas prior yang telah dihitung sebelumnya pada Tabel 4. Perhitungan probabilitas posterior ini digunakan untuk mengklasifikasikan data testing, sebagaimana ditampilkan pada Tabel 5. Sebagai contoh, data testing X dengan nilai yang tercantum pada kolom 2 Tabel 5 akan diklasifikasikan berdasarkan hasil perhitungan probabilitas posterior, yang ditampilkan pada kolom 3, 4, 5, dan 6 Tabel 5.

Tabel 5. Perhitungan untuk Menentukan Klasifikasi Data testing X

Data X	p(X Ci)			
	GBa	GKu	GBu	GLe
Jenis Kelamin	0.696629213	0.146067416	0.146067416	0.112359551
Umur	0.789473684	0.052631579	0.078947368	0.078947368
Berat Badan	0.882352941	0	0	0.117647059
Tinggi Badan	0.888888889	0	0	0.111111111

Dari tabel tersebut terdapat beberapa langkah untuk menghitung, yaitu :

- a. $P(X|C_i) = P(X)|\text{remark} = \text{GBa}$
 $= 0.696629213 \times 0.789473684 \times 0.882352941 \times 0.888888889$
 $= 0.431349358$
- b. $P(X|C_i) = P(X)|\text{remark} = \text{GKu}$
 $= 0.146067416 \times 0.052631579 \times 0 \times 0$
 $= 0$
- c. $P(X|C_i) = P(X)|\text{remark} = \text{GBu}$
 $= 0.146067416 \times 0.078947368 \times 0 \times 0$
 $= 0$
- d. $P(X|C_i) = P(X)|\text{remark} = \text{GLe}$
 $= 0.112359551 \times 0.078947368 \times 0.117647059 \times 0.111111111$
 $= 0.000115954$
- e. $P(X|C_i) P(C_i) = P(X| \text{remark} = \text{GBa})$
 $P(\text{remark} = \text{GBa})$
 $= 0.431349358 \times 0.678787879$
 $= 0.292794716$
- f. $P(X|C_i) P(C_i) = P(X| \text{remark} = \text{GBa})$
 $P(\text{remark} = \text{GKu})$
 $= 0 \times 0.139393939$
 $= 0$
- g. $P(X|C_i) P(C_i) = P(X| \text{remark} = \text{GBa})$
 $P(\text{remark} = \text{GBu})$
 $= 0 \times 0.048484848$

$$= 0$$

$$\text{h. } P(X|C_i) P(C_i) = P(X| \text{remark} = \text{GBa})$$

$$P(\text{remark} = \text{GLe})$$

$$= 0.000115954 \times 0.133333333$$

$$= 1.546060000$$

Berdasarkan perhitungan di atas, diperoleh nilai $P(X|C_i)$ dan $P(PX|C_i) P(C_i)$ yang lebih tinggi untuk $\text{remark} = \text{GBa}$, sehingga dapat disimpulkan bahwa data testing tersebut termasuk dalam klasifikasi GBa.

B. Evaluasi dan Validasi

Seperti pada gambar 1 sampai dengan gambar 3 untuk langkah penelitian pada algoritma Naïve Bayes, dilakukan uji validitas data dengan data training. Uji validitas akan dilakukan dengan *Confusion Matrix*.

C. Confusion Matrix Algoritma Naïve Bayes

Hasil perhitungan terhadap 4 atribut dengan 165 data menunjukkan bahwa 112 data tergolong dalam kategori gizi baik (GBa), 22 data sebagai gizi lebih (GLe), 23 data sebagai gizi kurang (GKu), dan 8 data sebagai gizi buruk (GBu).

Tabel 6. Tabel *Confusion Matrix* untuk algoritma Naïve Bayes

	true GBa	true GLe	true GKU	true GBu	class precision
pred GBa	99	11	14	3	77.95%
pred GLe	4	7	0	1	58.33%
pred GKU	9	2	9	3	38.13%
pred GBu	0	2	0	1	33.33%
class recall	88.38%	31.82%	38.13%	12.50%	

Performance Vector

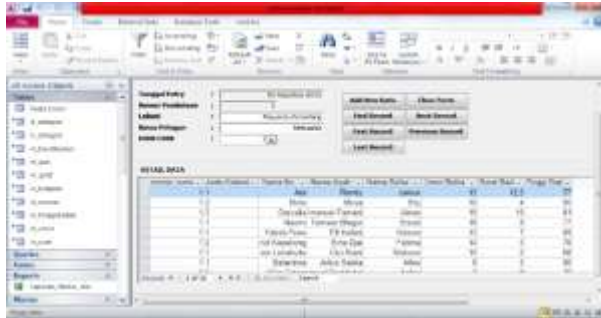
PerformanceVector:
 accuracy: 70.33% +/- 8.20% (mikro: 70.30%)
 ConfusionMatrix:
 True: GBa GLe GKU GBu
 GBa: 99 11 14 3
 GLe: 4 7 0 1
 GKU: 9 2 9 3
 GBu: 0 2 0 1
 kappa: 0.344 +/- 0.154 (mikro: 0.3316)
 ConfusionMatrix:
 True: GBa GLe GKU GBu
 GBa: 99 11 14 3
 GLe: 4 7 0 1
 GKU: 9 2 9 3
 GBu: 0 2 0 1

Gambar 4. Model *Confusion Matrix* untuk algoritma Naïve Bayes

Dari hasil validasi menggunakan aplikasi RapidMiner pada gambar di atas tergambar bahwa nilai *accuracy* dari algoritma naïve bayes 70.33%, nilai *class precision* 77,95%, dan nilai *class recall* 88.39%.

D. Database

Pembuatan database menggunakan Microsoft Access 2016. Dengan alat ini, database Access dapat dikelola, mulai dari tahap pembuatannya, tabel-tabel, dan otorisasi pengguna database.



Gambar 5. Kontruksi Database Dengan Microsoft Access

E. Tampilan Laporan Status Gizi Balita

Laporan Status Gizi Balita
 Nomor Sampel : 1
 Lokasi : Pasopati Kecamatan
 Nama Pengisi : Mutiawati

No	Kode	Keterangan
1	GBa	Gizi Baik
2	GKu	Gizi Kurang
3	GLe	Gizi Lebih
4	GBu	Gizi Buruk

No	Nama Balita	Umur				Status
		1	2	3	4	
1	Adi	11	7	10	10	Salah
2	Adi	11	7	10	10	Salah
3	Adi	11	7	10	10	Salah
4	Adi	11	7	10	10	Salah
5	Adi	11	7	10	10	Salah
6	Adi	11	7	10	10	Salah
7	Adi	11	7	10	10	Salah
8	Adi	11	7	10	10	Salah
9	Adi	11	7	10	10	Salah
10	Adi	11	7	10	10	Salah
11	Adi	11	7	10	10	Salah
12	Adi	11	7	10	10	Salah
13	Adi	11	7	10	10	Salah
14	Adi	11	7	10	10	Salah
15	Adi	11	7	10	10	Salah
16	Adi	11	7	10	10	Salah
17	Adi	11	7	10	10	Salah
18	Adi	11	7	10	10	Salah
19	Adi	11	7	10	10	Salah
20	Adi	11	7	10	10	Salah

Gambar 6. Tampilan Laporan Status Gizi Balita

4. Kesimpulan

Hasil perhitungan menggunakan aplikasi RapidMiner ditemukan bahwa pada algoritma *Naïve Bayes* nilai akurasi datanya sebesar 70.33 %, perhitungan pada data training yang terdiri dari 165 record menunjukkan bahwa 99 data diklasifikasikan sebagai GBa, sementara 11 data diprediksi sebagai GBa tetapi sebenarnya GLe, 14 data diprediksi sebagai GBa tetapi sebenarnya GKU, dan 3 data diprediksi sebagai GBa tetapi sebenarnya GBu. Selain itu, 4 data diprediksi sebagai GLe tetapi sebenarnya GBa, 7 data diklasifikasikan dengan benar sebagai GLe, dan 1 data diprediksi sebagai GLe tetapi sebenarnya GBu. Selanjutnya, 9 data diprediksi sebagai GKU tetapi sebenarnya GBa, 2 data diprediksi sebagai GKU tetapi sebenarnya GLe, 9 data diklasifikasikan dengan benar sebagai GKU, serta 3 data diprediksi sebagai GKU tetapi sebenarnya GBu, 2 data diprediksi

GBu tetapi ternyata GLe, 1 data secara benar diklasifikasikan GBu.

Hasil validasi tergambar bahwa nilai *accuracy* dari algoritma naïve bayes 70.33%, nilai *class precision* 77,95%, dan nilai *class recall* 88.39%.

Berdasarkan pembahasan di atas, algoritma *Naïve Bayes* dapat dimanfaatkan untuk meningkatkan akurasi data dalam menentukan status gizi balita. Oleh karena itu, algoritma ini dapat diterapkan sebagai alat bantu bagi tenaga kesehatan untuk penentuan status gizi secara cepat dan akurat. Selain itu pada penelitian selanjutnya dapat menggunakan algoritma lain untuk membandingkan algoritma dalam penentuan status gizi balita agar mendapatkan hasil yang terbaik.

Daftar Rujukan

[1] S. D. Nugraha, R. R. M. Putri, and R. C. Wihandika, "Penerapan Fuzzy K-Nearest Neighbor (FK-NN) Dalam Menentukan Status Gizi Balita," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 925–932, 2017.

[2] A. Saleh, "Implementasi Metode Klasifikasi *Naïve Bayes* Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," vol. 2, no. 3, pp. 207–217, 2015.

[3] J. Bimbingan and K. Indonesia, "EVALUASI PROGRAM BIMBINGAN DAN KONSELING: SEBUAH STUDI PUSTAKA," vol. 4, no. September, pp. 39–42, 2019.

[4] F. Nasari, S. Darma, and S. Informasi, "PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU," pp. 6–8, 2015.

[5] U. Suriani, "Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma," vol. 3, no. 2, pp. 55–66, 2023.

[6] R. Rachman and R. N. Handayani, "Klasifikasi Algoritma *Naive Bayes* Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," vol. 8, no. 2, pp. 111–122, 2021.

[7] Ilham, W., Putra, N., Putri, T. E., Kurniawan, E., & Molina, J. I. (2025). ANALISIS SISTEM FORECASTING PADA PRODUKSI DAN PERMINTAAN TELUR IMPLEMENTASI METODE LEAST SQUARE. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1), 1392-1398.

[8] Molina, J. I., & Malese, L. P. (2024). SISTEM PENDUKUNG KEPUTUSAN PENENTUAN STATUS GIZI BALITA PADA DINAS KESEHATAN KABUPATEN ALOR MENGGUNAKAN ALGORITMA C4. 5. *Jurnal Teknologi Informasi dan Komunikasi*, 17(1), 1-13.

[9] Tanggraeni, A. I., & Sitokdana, M. N. (2022). Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma *Naïve Bayes*. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(2), 785-795.

[10] Prasetyo, S. D., Hilabi, S. S., & Nurapriani, F. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma *Naïve Bayes* dan KNN. *Jurnal KomtekInfo*, 1-7.

[11] Septianingrum, F., & Irawan, A. S. Y. (2021). Metode Seleksi Fitur Untuk Klasifikasi Sentimen Menggunakan Algoritma *Naive Bayes*: Sebuah Literature Review. *Jurnal Media Informatika Budidarma*, 5(3), 799.

[12] Lestari, S., & Saepudin, S. (2021, September). Analisis sentimen vaksin sinovac pada twitter menggunakan algoritma *Naive Bayes*. In *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, pp. 163-170).

[13] Widodo, Y. B., Anggraeini, S. A., & Sutabri, T. (2021). Perancangan Sistem Pakar Diagnosis Penyakit Diabetes Berbasis Web Menggunakan Algoritma *Naive Bayes*. *J. Teknol. Inform. dan Komput*, 7(1), 112-123.