

APLIKASI KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN METODE NAÏVE BAYES BERBASIS ANDROID

Eva Darnila✉, Maryana, Muhammad Azmi

Program Studi Teknik Informatika, Universitas Malikussaleh, Aceh Utara, Indonesia

Email: eva.darnila@unimal.ac.id

DOI: <https://doi.org/10.46880/jmika.Vol5No2.pp135-141>

ABSTRACT

Malnutrition is a condition that a person experience because the lack of nutrient intake or the amount of nutrients that exist is still below standard. nutritional problems are most commonly experienced by toddlers. Nutritional status in toddlers is a benchmark of the health condition of toddlers. The nutritional status of toddlers can currently be measured based on anthropometric index using WHO standards. Anthropometrics is an assessment of the nutritional status of toddlers determined through measurements of the human body In this study measurement of nutritional status using a technique in data mining namely Naïve Bayes Method that aims to facilitate health centers in measuring the nutritional status of toddlers quickly through android-based applications. Naïve Bayes method performs nutritional status classification of toddlers that will produce a classification very thin, thin, normal and fat. Based on data taken from banda sakti health center Lhokseumawe city which then carried out the classification process, obtained the results of nutritional status "skinny" with data testing male sex, weight 7.4 kg and height 64.6.

Keyword: *Data Mining, Nutrition Status of Toddlers, Classification, Naïve Bayes Methods.*

ABSTRAK

Gizi buruk merupakan suatu keadaan yang dialami seseorang dikarenakan kurangnya asupan nutrisi atau jumlah nutrisi yang ada masih dibawah standar. masalah gizi paling sering dialami oleh balita. Status gizi pada balita merupakan tolak ukur kondisi kesehatan balita. Status Gizi balita saat ini bisa diukur berdasarkan indeks antropometri yang merupakan standar baku WHO. Antropometri adalah penilaian status gizi balita yang ditentukan melalui pengukuran tubuh manusia. Pada penelitian ini pengukuran status gizi menggunakan suatu teknik dalam data mining yaitu Metode *Naïve Bayes* yang bertujuan untuk mempermudah puskesmas dalam melakukan pengukuran status gizi balita secara cepat melalui aplikasi berbasis android. Metode *Naïve Bayes* melakukan klasifikasi status gizi balita yang akan menghasilkan klasifikasi sangat kurus, kurus, normal dan gemuk. Berdasarkan data balita yang diambil dari puskesmas Banda Sakti Kota Lhokseumawe yang kemudian dilakukan proses klasifikasi, diperoleh hasil status gizi "kurus" dengan data pengujian jenis kelamin laki-laki, berat badan 7,4 kg dan tinggi badan 64,6.

Kata Kunci: *Data Mining, Status Gizi Balita, Klasifikasi, Metode Naïve Bayes.*

PENDAHULUAN

Gizi buruk adalah kondisi yang dialami seseorang dikarenakan kekurangan asupan nutrisi atau jumlah nutrisi yang dikonsumsi berada dibawah standar. Nutrisi yang dibutuhkan diantaranya karbohidrat, protein serta kalori. Salah satu permasalahan gizi yang paling utama dan paling sering dialami bayi usia di bawah lima tahun (balita) adalah kekurangan energi protein (Rahmawati, 2017).

Berdasarkan data Profil Kesehatan Aceh 2018, kasus gizi buruk Kabupaten Aceh Utara menempati urutan pertama dengan jumlah sebanyak 429 kasus, diikuti kabupaten Bireuen sebanyak 17 kasus. Dari data tersebut yang mendapatkan perawatan di Aceh Utara

hanya 5% dari jumlah kasus sedangkan kabupaten Bireuen dan kabupaten lainnya di Aceh mendapat perawatan 100% dari kasus yang ditemukan (Dinas Kesehatan Aceh, 2018).

Angka balita gizi buruk yang sangat tinggi, maka diperlukan pendeteksian awal bagi balita untuk mengurangi kasus gizi buruk, salah satu langkah yang dapat dilakukan ialah dengan menguji status gizi balita, yang bertujuan untuk mendapatkan informasi status balita gizi secara cepat agar mempermudah instansi kesehatan melakukan penanganan pada balita.

Untuk melakukan pengujian status gizi balita, dapat dilakukan dengan menggunakan sebuah teknik dalam data mining yaitu klasifikasi. Klasifikasi

melakukan analisis data pelatihan untuk dapat membentuk aturan klasifikasi yang biasa disebut sebagai proses belajar atau pelatihan. Selanjutnya, klasifikasi akan menganalisis data pengujian untuk menguji aturan klasifikasi yang telah diperoleh dari proses pelatihan (Putri, Sendari, & Widiyaningtyas, 2018).

Salah satu algoritma klasifikasi yang sering dipakai ialah algoritma naïve bayes. Algoritma Naïve Bayes adalah algoritma yang memanfaatkan teori kemungkinan, yakni memprediksi kemungkinan di masa depan berdasarkan pengalaman pada masa sebelumnya (Hariri & Pamungkas, 2016).

METODE

Metode Naïve Bayes

Naïve Bayes adalah sebuah metode pengelompokan statistik yang bisa dipakai untuk memprediksi probabilitas anggota suatu class. *Naïve Bayes* juga mempunyai akurasi dan kecepatan yang sangat kuat ketika diaplikasikan pada database dengan big data (Widiastuti, Santosa, & Supriyanto, 2014). Teorema Bayes memiliki bentuk umum sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \dots\dots\dots(1)$$

Keterangan:

- X* : Data dengan class yang belum diketahui
- H* : Hipotesis data *X* merupakan suatu class spesifik
- P(H|X)* : Probabilitas hipotesis *H* berdasarkan kondisi *X*
- P(H)* : Probabilitas hipotesis *H* (prior probability)
- P(X|H)* : Probabilitas *X* berdasarkan kondisi pada hipotesis *H*
- P(X)* : Probabilitas dari *X*

Jika atribut ke-*i* bersifat diskret, maka *P(H|X)* diestimasi sebagai frekuensi relatif dari sampel yang memiliki nilai *H* sebagai atribut ke *i* dalam kelas *X*. Namun, jika atribut ke-*i* bersifat kontinu, maka *P(H|X)* diestimasi dengan fungsi densitas *Gauss* (Putri et al., 2018).

$$P(X_i = x_i|Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i - \mu)^2}{2\sigma_{ij}^2}} \dots\dots\dots(2)$$

Keterangan:

- P* : Peluang
- X_i* : Atribut ke *i*
- x_i* : Nilai atribut ke *i*
- Y* : Kelas yang dicari

- y_j* : Sub kelas *Y* yang dicari
- μ* : Mean, menyatakan rata-rata seluruh atribut
- σ* : Standar deviasi

Persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat dibawah ini (Guntur, Santony, & Yuhandri, 2018):

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \dots\dots\dots(3)$$

Keterangan:

- σ* : Standar deviasi
- x_i* : Data ke *i*
- μ* : Rata-rata populasi
- n* : Jumlah data

Dan rumus yang digunakan untuk menghitung nilai rata-rata (mean) dapat dilihat sebagai berikut (Guntur et al., 2018):

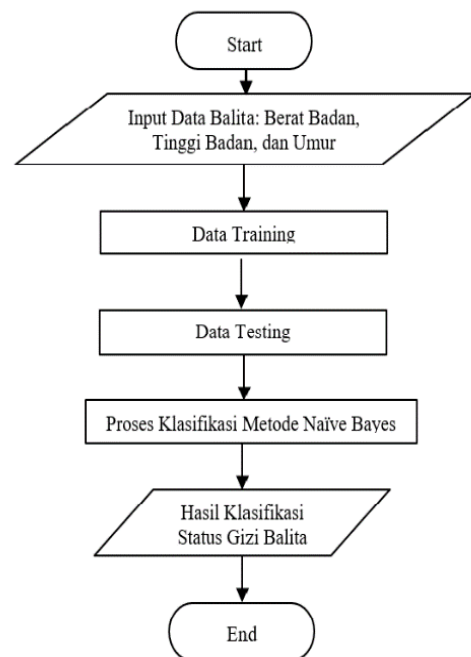
$$\mu = \frac{\sum_{i=1}^n x_i}{n} \dots\dots\dots(4)$$

Keterangan:

- μ* : Nilai rata-rata(mean)
- x_i* : Data ke *i*
- n* : Jumlah data

Skema Sistem

Skema sistem untuk pengolahan data dengan menggunakan metode naïve bayes sebagai berikut:



Gambar 1. Skema Sistem

HASIL DAN PEMBAHASAN

Data Training

Data training merupakan bahan baku informasi yang didapatkan berdasarkan data yang ada di lapangan. Pada penelitian ini data didapatkan dari Puskesmas Banda Sakti Kota Lhokseumawe. Data

training pada penelitian ini berjumlah 4072 data balita, dimana data training balita yang terdiri dari 28 balita berstatus gizi sangat kurus, 278 kurus, 3677 normal dan 89 gemuk. data training dapat dilihat pada tabel (1) berikut:

Tabel 1. Data Training Status Gizi Balita

No	Jenis Kelamin	Berat Badan	Tinggi Badan	Status Gizi
1	P	11	94	Kurus
2	P	14,4	97,6	Normal
3	L	17,1	99	Normal
...
...
4070	P	7,8	72	Normal
4071	P	3	51	Kurus
4072	P	2,8	51	Sangat Kurus

Data Testing

Data testing (data uji) merupakan bahan baku yang sengaja dibuat untuk membantu proses penelitian

dalam sistem data mining, secara lebih rinci data testing dapat di lihat pada tabel 2 berikut:

Tabel 2. Data Testing Balita

No	Jenis Kelamin	Berat Badan	Tinggi Badan	Status Gizi
1	L	7,4	64,6	?

Penentuan Hasil Klasifikasi

Apabila data testing telah dimasukkan, klasifikasi status gizi balita dapat ditentukan melalui langkah berikut:

seperti berat badan dan tinggi badan. Mean dan standar deviasi untuk masing-masing kelas atau label Sangat Kurus, Kurus, Normal dan Gemuk.

Menghitung Mean dan Standar Deviasi

Perhitungan mean dan standar deviasi dilakukan pada setiap atribut yang bersifat kontinu atau numerik

Untuk atribut berat badan, hasil perhitungan mean dan standar deviasi diperoleh sebagai berikut:

$$\mu_{sangat\ kurus} = \frac{3+5,8+2,8+2,7+3,2+\dots+3,2+3,2+7+2,8+2,8}{28} = 4,857142857$$

$$\mu_{kurus} = \frac{11+10+10,9+5+8+\dots+12+12+11+12+3}{278} = 8,492446043$$

$$\mu_{normal} = \frac{14,5+17,1+18+16,6+15,8+\dots+15+16,7+12+3,2+7,8}{3677} = 12,28550721$$

$$\mu_{gemuk} = \frac{19+9+4,5+5,8+6+\dots+12+20+17+20+17}{89} = 16,83258427$$

$$\sigma_{sangat\ kurus} = \sqrt{\frac{(3-4,857142857)^2+(5,8-4,857142857)^2+\dots+(2,8-4,857142857)^2}{28}} = 2,53650893$$

$$\sigma_{kurus} = \sqrt{\frac{(11-8,492446043)^2+(10-8,492446043)^2+\dots+(3-8,492446043)^2}{278}} = 3,534645656$$

$$\sigma_{normal} = \sqrt{\frac{(14,5-12,28550)^2+(17,1-12,28550)^2+\dots+(7,8-12,28550)^2}{3677}} = 46,10739623$$

$$\sigma_{gemuk} = \sqrt{\frac{(19-16,83258427)^2 + (9-16,83258427)^2 + \dots + (20-16,83258427)^2 + (17-16,83258427)^2}{89}} = 9,16501378$$

Selanjutnya kita hitung mean dan standar deviasi pada atribut tinggi badan, hasil perhitungan diperoleh sebagai berikut:

$$\begin{aligned} \mu_{sangat\ kurus} &= \frac{52+69+52+53+53+\dots+53+53+76+51+51}{28} = 63,43928571 \\ \mu_{kurus} &= \frac{94+87,5+91,5+61+78,9+\dots+95+95+94+95+51}{278} = 79,27553957 \\ \mu_{normal} &= \frac{97,6+99+100,5+100,5+109,9+\dots+96+97,5+92+50+72}{3677} = 85,63937721 \\ \mu_{gemuk} &= \frac{94+65+51+52+56+\dots+78+99+96+98+95}{89} = 87,44269663 \end{aligned}$$

$$\sigma_{sangat\ kurus} = \sqrt{\frac{(52-63,43928571)^2 + (69-63,43928571)^2 + \dots + (51-63,43928571)^2 + (51-63,43928571)^2}{28}} = 15,10681822$$

$$\sigma_{kurus} = \sqrt{\frac{(94-79,27553957)^2 + (87-79,27553957)^2 + \dots + (95-79,27553957)^2 + (51-79,27553957)^2}{278}} = 18,13952456$$

$$\sigma_{normal} = \sqrt{\frac{(97,6-85,63937721)^2 + (99-85,63937721)^2 + \dots + (50-85,63937721)^2 + (72-85,63937721)^2}{3677}} = 14,09985787$$

$$\sigma_{gemuk} = \sqrt{\frac{(94-87,44269663)^2 + (65-87,44269663)^2 + \dots + (98-87,44269663)^2 + (95-87,44269663)^2}{89}} = 13,61679307$$

Menghitung Probabilitas

Menghitung P(Hi)

Perhitungan P(Hi) adalah perhitungan jumlah *class* / label. Jumlah *class* / label ada 4 yaitu sangat kurus, kurus, normal dan gemuk. Perhitungan dilakukan pada setiap label sehingga diperoleh hasil perhitungan dari keempat label sebagai berikut:

P(Y = label / class) = Jumlah label class / Total jumlah data

$$P(\text{Sangat Kurus}) = \frac{28}{4072} = 0,006876228$$

$$P(\text{Kurus}) = \frac{278}{4072} = 0,06827112$$

$$P(\text{Normal}) = \frac{3677}{4072} = 0,902996071$$

$$\begin{aligned} P(\text{Gemuk}) &= \frac{89}{4072} \\ &= 0,021856582 \end{aligned}$$

Menghitung P(X|Hi)

Perhitungan P(X|Hi) adalah perhitungan jumlah kasus yang sama dengan *class* yang sama. Pada penelitian ini perhitungan P(X|Hi) dilakukan pada atribut jenis kelamin, karena jenis kelamin merupakan atribut diskrit. Perhitungan juga dilakukan berdasarkan empat label, sehingga diperoleh hasil sebagai berikut:

Label class H = Sangat Kurus

Jumlah data dengan label class Sangat Kurus yaitu 28, hasil perhitungannya sebagai berikut:

$$P(\text{Jenis Kelamin} = \text{Laki} - \text{laki} \mid \text{Sangat Kurus}) = \frac{14}{28} = 0,5$$
$$P(\text{Jenis Kelamin} = \text{Perempuan} \mid \text{Sangat Kurus}) = \frac{14}{28} = 0,5$$

Label class H = Kurus

Jumlah data dengan label class Y = Kurus yaitu 278, hasil perhitungannya sebagai berikut:

$$P(\text{Jenis Kelamin} = \text{Laki} - \text{laki} \mid \text{Kurus}) = \frac{161}{278} = 0,579136691$$
$$P(\text{Jenis Kelamin} = \text{Perempuan} \mid \text{Kurus}) = \frac{117}{278} = 0,420863309$$

Label class H = Normal

Jumlah data dengan label class Y = Normal yaitu 3677, hasil perhitungannya sebagai berikut:

$$P(\text{Jenis Kelamin} = \text{Laki} - \text{laki} \mid \text{Normal}) = \frac{2031}{3677} = 0,552352461$$
$$P(\text{Jenis Kelamin} = \text{Perempuan} \mid \text{Normal}) = \frac{1646}{3677} = 0,447647539$$

Label class H = Gemuk

Jumlah data dengan label class Y = Gemuk yaitu 89, hasil perhitungannya sebagai berikut:

$$P(\text{Jenis Kelamin} = \text{Laki} - \text{laki} \mid \text{Gemuk}) = \frac{53}{89} = 0,595505618$$
$$P(\text{Jenis Kelamin} = \text{Perempuan} \mid \text{Gemuk}) = \frac{36}{89} = 0,404494382$$

Pada data numerik atau data kontinu, perhitungan probabilitas digantikan dengan densitas gauss. Perhitungan dilakukan berdasarkan data testing numerik yang telah ada sebelumnya pada tabel 2. Perhitungan densitas gauss hanya dilakukan pada atribut berat badan dan tinggi badan.

Untuk atribut berat badan = 7,4 kg, proses dan hasil perhitungan densitas gauss diperoleh sebagai berikut:

$$f(\text{Berat Badan} = 7,4 \mid \text{Label Sangat Kurus}) = \frac{1}{\sqrt{2\pi(2,53650893)}} e^{-\frac{(7,4-4,857142857)^2}{2(2,53650893)^2}} = 0,09518056$$
$$f(\text{Berat Badan} = 7,4 \mid \text{Label Kurus}) = \frac{1}{\sqrt{2\pi(3,534645656)}} e^{-\frac{(7,4-8,492446043)^2}{2(3,534645656)^2}} = 0,1076296$$
$$f(\text{Berat Badan} = 7,4 \mid \text{Label Normal}) = \frac{1}{\sqrt{2\pi(46,10739623)}} e^{-\frac{(7,4-12,28550721)^2}{2(46,10739623)^2}} = 0,008606203$$
$$f(\text{Berat Badan} = 7,4 \mid \text{Label Gemuk}) = \frac{1}{\sqrt{2\pi(9,16501378)}} e^{-\frac{(7,4-16,83258427)^2}{2(9,16501378)^2}} = 0,025637496$$

Selanjutnya untuk atribut tinggi badan = 64,4 cm, proses dan hasil perhitungan densitas gauss diperoleh sebagai berikut :

$$f(\text{Tinggi Badan} = 64,6 \mid \text{Label} = \text{Sangat Kurus}) = \frac{1}{\sqrt{2\pi(15,10681822)}} e^{-\frac{(64,6-63,43928571)^2}{2(15,10681822)^2}}$$
$$= 0,026336937$$
$$f(\text{Tinggi Badan} = 64,6 \mid \text{Label Kurus}) = \frac{1}{\sqrt{2\pi(18,13952456)}} e^{-\frac{(64,6-79,27553957)^2}{2(18,13952456)^2}}$$
$$= 0,015858521$$

$$f(\text{Tinggi Badan} = 64,6 \mid \text{Label Normal}) = \frac{1}{\sqrt{2\pi(14,09985787)}} e^{-\frac{(64,6-85,63937721)^2}{2(14,09985787)^2}}$$
$$= 0,00929634$$

$$f(\text{Tinggi Badan} = 64,6 \mid \text{Label Gemuk}) = \frac{1}{\sqrt{2\pi(13,61679307)}} e^{-\frac{(64,6-87,44269663)^2}{2(13,61679307)^2}}$$
$$= 0,007175687$$

Menghitung Probabilitas Akhir

Perhitungan probabilitas akhir dilakukan dengan cara mengalikan semua hasil nilai probabilitas atribut dan juga hasil densitas gauss. Hasil nilai probabilitas akhir yang diperoleh sebagai berikut:

- $P(H = \text{Sangat kurus}) = P(\text{Berat badan} \mid \text{Sangat kurus}) * P(\text{Tinggi badan} \mid \text{Sangat kurus}) * P(\text{Jenis kelamin} \mid \text{Sangat kurus})$
 $= 0,006876228 \times 0,09518056 \times 0,026336937 \times 0,5$
 $= 0,00000861854163181961$
- $P(H = \text{Kurus}) = P(\text{Berat badan} \mid \text{Kurus}) * P(\text{Tinggi badan} \mid \text{Kurus}) * P(\text{Jenis kelamin} \mid \text{Kurus})$
 $= 0,06827112 \times 0,1076296 \times 0,015858521 \times 0,579136691$
 $= 0,0000674858158701352$
- $P(H = \text{Normal}) = P(\text{Berat badan} \mid \text{Normal}) * P(\text{Tinggi badan} \mid \text{Normal}) * P(\text{Jenis kelamin} \mid \text{Normal})$
 $= 0,902996071 \times 0,008606203 \times 0,00929634 \times 0,552352461$
 $= 0,00003990614259853130$
- $P(H = \text{Gemuk}) = P(\text{Berat badan} \mid \text{Gemuk}) * P(\text{Tinggi badan} \mid \text{Gemuk}) * P(\text{Jenis kelamin} \mid \text{Gemuk})$
 $= 0,021856582 \times 0,025637496 \times 0,007175687 \times 0,595505618$
 $= 0,00000239445784433354$

Setelah hasil perkalian diperoleh, kemudian dilakukan normalisasi terhadap hasil tersebut sehingga jumlah nilai yang diperoleh = 1.

$$\text{Sangat Kurus} = \frac{0,0000086185}{0,000008618+0,000067485+0,000039904+0,000002394} = 0,07278948$$

$$\text{Kurus} = \frac{0,000067485}{0,000008618+0,000067485+0,000039904+0,000002394} = 0,56996389$$

$$\text{Normal} = \frac{0,000039904}{0,000008618+0,000067485+0,000039904+0,000002394} = 0,33702379$$

$$\text{Gemuk} = \frac{0,000002394}{0,000008618+0,000067485+0,000039904+0,000002394} = 0,02022283$$

Membandingkan Hasil Normalisasi

Setelah hasil normalisasi diperoleh, kita bandingkan semua hasil normalisasi nilai *probabilitas* yang menunjukkan nilai terbesar. Diperoleh *probabilitas* kelas kurus paling tinggi dengan nilai 0,56996389, maka data testing masuk kedalam klasifikasi kurus.

KESIMPULAN

Hasil pengujian data balita dengan jenis kelamin "laki-laki", berat badan 7,4 kg dan tinggi badan 64,6 cm mendapatkan hasil normalisasi yang terdiri dari sangat kurus sebesar 0,07278948, kurus sebesar

0,56996389, normal sebesar 0,33702379 dan gemuk sebesar 0,02022283. Setiap hasil normalisasi dibandingkan berdasarkan nilai terbesar, maka didapatkan nilai 0,56996389 yang paling besar sehingga hasil klasifikasi status gizi balita adalah "Kurus".

DAFTAR PUSTAKA

- Dinas Kesehatan Aceh. (2018). *Profil Kesehatan Aceh 2018*.
- Guntur, M., Santony, J., & Yuhandri. (2018). Prediksi Harga Emas dengan Menggunakan Metode Naïve Bayes dalam Investasi untuk Meminimalisasi Resiko. *Jurnal RESTI*

- (*Rekayasa Sistem Dan Teknologi Informasi*),
2(1), 354–360.
- Hari, F. R., & Pamungkas, D. P. (2016).
Implementasi Naïve Bayes Classifier Untuk
diagnose Status Gizi Balita. *Seminar Nasional
Teknologi Informasi Dan Multimedia*, 6–13.
- Putri, R. A., Sendari, S., & Widiyaningtyas, T. (2018).
Classification of Toddler Nutrition Status with
Anthropometry Calculation Using Naïve Bayes
Algorithm. *3rd International Conference on
Sustainable Information Engineering and
Technology, SIET 2018*, 66–70.
- Rahmawati, E. (2017). Implementasi Program
(Gentasibu) Gerakan Pengentasan Gizi Buruk
Di Puskesmas Kecamatan Tanjunganom
Kabupaten Nganjuk. *Publika*, 5(3).
- Widiastuti, N. A., Santosa, S., & Supriyanto, C.
(2014). Algoritma Klasifikasi Data Mining
Naïve Bayes Berbasis Particle Swarm Particle
Swarm Optimization Untuk Deteksi Penyakit
Jantung. *Pseudocode*, 1(1), 11–14.