

Klasifikasi Status Gizi Balita Menggunakan Algoritma Naive Bayes Berdasarkan Pengukuran Antropometri

Classification of Toddler Nutrition Status Using Naive Bayes Algorithm Based on Anthropometric Measurements

Widya Kurniawan, Dihin Muriyatmoko, Safwa Alifiyah Taufiq

Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Darussalam Gontor
Jl. Raya Siman, km.6, Siman Ponorogo, 63471

Abstrak

Klasifikasi status gizi balita berdasarkan tinggi badan menurut umur (TB/U atau PB/U) merupakan masalah penting dalam upaya pemantauan dan penanganan kesehatan anak. Penelitian ini bertujuan untuk mengklasifikasikan status gizi balita di Puskesmas Jogorogo dengan menggunakan pengukuran antropometri, menggunakan algoritma Naive Bayes. Data yang digunakan berasal dari balita yang dikumpulkan di Puskesmas Jogorogo pada dua periode waktu, yaitu Agustus 2023 dengan jumlah 1784, dan Februari 2024 dengan jumlah 2147 data. Metode yang digunakan dalam penelitian ini adalah CRISP-DM (Cross-Industry Standard Process for Data Mining) untuk alur analisa data, dengan penerapan algoritma Naive Bayes untuk klasifikasi. Evaluasi model dilakukan menggunakan confusion matrix untuk mengukur kinerja klasifikasi. Hasil evaluasi menunjukkan bahwa algoritma Naive Bayes memberikan hasil yang konsisten pada kedua dataset. Akurasi yang diperoleh untuk data tahun 2023 adalah 91,83%, sedangkan untuk data tahun 2024 sedikit meningkat menjadi 91,84%. Hal ini menunjukkan bahwa algoritma Naive Bayes memiliki kinerja yang baik dalam mengklasifikasikan status gizi balita berdasarkan tinggi badan menurut umur.

Kata Kunci : Status gizi, TB/U, Klasifikasi, Naive Bayes, CRISP-DM, Puskesmas Jogorogo.

Abstract

Classification of nutritional status of toddlers based on height for age (H/A or H/A) is an important issue in efforts to monitor and treat child health. This study aims to classify the nutritional status of toddlers at the Jogorogo Health Center using anthropometric measurements, using the Naive Bayes algorithm. The data used came from toddlers collected at the Jogorogo Health Center in two time periods, namely August 2023 with a total of 1784, and February 2024 with a total of 2147 data. The method used in this study is CRISP-DM (Cross-Industry Standard Process for Data Mining) for the data analysis flow, with the application of the Naive Bayes algorithm for classification. Model evaluation was carried out using a confusion matrix to measure classification performance. The evaluation results showed that the Naive Bayes algorithm provided consistent results on both datasets. The accuracy obtained for the 2023 data was 91.83%, while for the 2024 data it increased slightly to 91.84%. This shows that the Naive Bayes algorithm has good performance in classifying the nutritional status of toddlers based on height according to age.

Keyword : Nutritional status, Height/Age, Classification, Naive Bayes, CRISP-DM, Jogorogo Health Center.

*Penulis Korespondensi. Telp: +62 82289067479

Alamat E-mail : safwataufiq42056@mhs.unida.gontor.ac.id

1. Pendahuluan

Bayi dibawah lima tahun atau disingkat menjadi balita, dimana saat masa ini otak anak sedang mengalami perkembangan pertumbuhan yang pesat[1]. Pada tahun 2021 data dari *global Nutrition Report* menunjukkan bahwa secara umum untuk diseluruh dunia terdapat 22,2 % balita *stunting* (kekerdilan), sebanyak 7,5% balita kurus, serta terdapat sekitar 5,6% balita gemuk[2]. Pengukuran antropometri adalah pengukuran yang digunakan untuk penentuan keadaan gizi seseorang yang berhubungan dengan pengukuran dimensi tubuh. Dalam regulasi Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak, terdapat 6 kategori status gizi[3]. Indeks Antropometri yang sering dipakai adalah berat badan dibagi umur (BB/U), berat badan dibagi tinggi badan (BB/TB) dan tinggi badan dibagi umur (TB/U). Dalam pengukuran usia dibawah 60 bulan cenderung memakai berat badan dibagi umur (BB/U)[4].

Dari beberapa penelitian terdahulu tentang klasifikasi status gizi balita menggunakan metode *Naive Bayes* pada tahun 2020 penelitian yang ditulis oleh Nuraisana dkk tentang status gizi berdasarkan indeks BB/U dan TB/U dengan hasil akurasi yang diperoleh yaitu 90,91% untuk status gizi BB/U dan 94,42% untuk status gizi TB/U[5]. Pada tahun 2022 penelitian yang ditulis oleh Heru Budi Setiawan dkk tentang status gizi dengan tujuan memudahkan status gizi bati lima tahun menggunakan data mining dengan algoritma *Naive Bayes* dengan hasil pengujian sebesar 72% (BB/U), 70,8% (PB/U) dan 71,02% (BB/TB)[1]. Pada tahun 2022 penelitian yang ditulis oleh Raya Sikandi dkk tentang klasifikasi status gizi balita menggunakan penerapan *Logika Fuzzy* dengan variabel usia, berat badan, dan tinggi badan dengan hasil nilai akurasi sebesar 88,3%[6].

2. Teori Dasar

a. Klasifikasi

Klasifikasi adalah proses untuk menentukan mode atau fungsi yang menjelaskan dan membedakan konsep keas data untuk memperkirakan kelas yang tidak diketahui untuk mengklasifikasikan objek atau instansi ke dalam kelompok atau kategori yang telah ditentukan sebelumnya dalam pengklasifikasian data terdapat dua yang dilakukan yaitu proses training dan proses testing[7].

b. Pengukuran Antropometri

Standar Antropometri anak digunakan untuk menilai atau menentukan status gizi anak. Penilaian status gizi anak dilakukan dengan membandingkan hasil pengukuran berat badan dan panjang/tinggi badan dengan standar Antropometri anak. Bahwa standar ukuran pertumbuhan anak sebagaimana diatur dalam

Keputusan Menteri Kesehatan Nomor 1995/MENKES/SK/XII/2010[4].

c. *Naive Bayes*

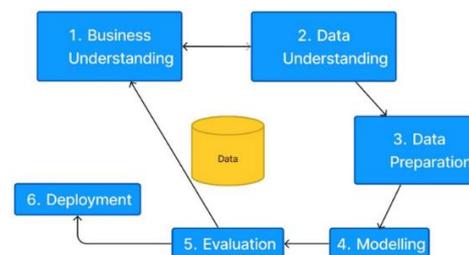
Naive Bayes yaitu suatu metode klasifikasi probalistik yang sederhana untuk menghitung beberapa kemungkinan dengan dijumlahkan frekuensi dan kombinasi nilai dari suatu dataset. Teorema Bayes diasumsikan bahwa atribut yang independent atau tidak saling bergantung diberikan oleh nilai pada variabel class. Arti lain dari *Naive Bayes* yang dikembangkan oleh Thomas Bayes (ilmuwan Inggris) yaitu metode klasifikasi statistik untuk memperkirakan peluang yang akan terjadi di masa yang akan datang berdasarkan pengalaman dari masa lalu[7].

d. Confusion Matrix

Confusion matrix digunakan untuk mempresentasikan kinerja model dalam 4 parameter yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP), *false negative* (FN). Keempat parameter ini digunakan sebagai landasan dalam rumus untuk menentukan *performance metrics* seperti *accuracy*, *precision*, *recall*, dan *f1-score*. [8]

3. Metodologi

Penelitian ini bertujuan untuk mengimplementasikan Algoritma *Naive Bayes* pada klasifikasi status gizi berdasarkan pengukuran Antropometri. Penelitian ini juga mengukur tingkat akurasi pada metode klasifikasi yang digunakan yaitu algoritma *Naive Bayes*. Dibawah berikut adalah tahapan metodologi penelitian pada gambar (1) :



Gambar 1. Tahapan Penelitian

a. *Business Understanding*

Langkah awal yang penting dalam sebuah penelitian adalah mengidentifikasi masalah. Pada fase ini, kegiatan penelitian dilakukan untuk merinci dan memahami esensi permasalahan serta mencari solusi yang sesuai. Fokus permasalahan dalam penelitian ini adalah melakukan pengklasifikasian status gizi pada balita untuk memberikan informasi dari data status gizi mengenai tinggi badan menurut umur (TB/U atau PB/U).

b. *Data Understanding*

Data yang diperoleh yaitu mencakup dua periode waktu: Agustus 2023 dan Februari

2024. Masing-masing dataset memiliki 32 kolom. Tahun 2023 berisi 1783 dataset dan 2024 berisi 2147 dataset. Setelah dilakukannya proses *understanding*, terdapat 8 *record* missing value pada Agustus 2023 dan 6 *record* pada Februari 2024.

c. Data Preparation

Semua data yang mencakup semua kegiatan pada persiapan data yang diteliti untuk membangun dataset status gizi yang diterapkan ke dalam alat pemodelan, dari data mentah berupa dataset status gizi dan selanjutnya akan dilakukan proses data mining. Tahap selanjutnya dilakukan tahap pemodelan klasifikasi pada data mining.

d. Modelling

Pada tahap ini akan digunakan pemodelan menggunakan metode klasifikasi akan dibagi menjadi 4 kategori yaitu sangat pendek, pendek, normal, dan tinggi.

e. Evaluasi

Fase evaluasi merupakan tahap untuk mengetahui apakah model yang dirancang telah sesuai atau belum dengan tujuan pada fase awal dan merupakan langkah yang menentukan seberapa baik model yang dibangun dapat memprediksi hasil dengan akurasi yang tinggi. Tujuan awal dirancangnya model ini yaitu agar menghasilkan nilai akurasi yang tinggi, sehingga dapat membuktikan bahwa penelitian yang dilakukan telah berhasil.

f. Deployment

Tahap selanjutnya dari penelitian yang sudah dilakukan yaitu tahap penyebaran hasil dan dijadikan sebagai presentasi atau laporan dari pengetahuan yang sudah didapat berdasarkan evaluasi dan pemodelan pada proses data mining. Pada tahap ini model skenario diimplementasikan ke dalam sebuah bentuk visualisasi data berupa *decision boundary plot*.

4. Hasil dan Pembahasan

a. Pengumpulan Data

Pada penelitian ini menggunakan pengumpulan data Puskesmas Jogorogo pada Agustus 2023 dan Februari 2024. Data yang diperoleh menunjukkan bahwa atribut yang dapat mempengaruhi penjualan adalah Jenis Kelamin, Usia saat Diukur, Tinggi, ZS TB/U, TB/U. Dengan demikian data status gizi yang diperoleh tersebut akan dilakukan pengolahan dengan teknik klasifikasi data mining yang bertujuan untuk mengimplementasikan algoritma *Naive Bayes* pada klasifikasi status gizi berdasarkan Pengukuran Antropometri. Gambar (2) merupakan bentuk data status gizi yang akan diolah.

JK	Usia Saat Ukur	Tinggi	ZS TB/U	TB/U
P	0 Tahun - 4 Bulan - 22 Hari	67	1.57	Normal
L	0 Tahun - 7 Bulan - 20 Hari	67	-1.44	Normal
L	0 Tahun - 9 Bulan - 1 Hari	67	-2.23	Pendek
P	0 Tahun - 6 Bulan - 6 Hari	68	-1.68	Normal
L	0 Tahun - 7 Bulan - 19 Hari	69	-0.51	Normal
P	0 Tahun - 8 Bulan - 25 Hari	69	-0.38	Normal
L	0 Tahun - 6 Bulan - 14 Hari	70	0.76	Normal

Gambar 2. Data Status Gizi TB/U sebelum diolah.

Namun, untuk memudahkan proses dalam mesin membaca data, terdapat salah 1 atribut Usia saat ukur masih terdapat data dalam bentuk type data *string* dengan contoh, 0 tahun – 4 bulan – 22 hari. Bentuk itu mungkin cukup agak sulit dibaca oleh mesin, maka dari itu atribut Usia saat ukur type data akan diubah dengan bentuk desimal atau *float*. Dimana mesin mungkin akan lebih mudah memahami konteks sebuah data yang akan dibaca. Untuk mengubah bentuk type data string menjadi *float*, digunakan *google colab* terlihat pada gambar (3) atribut Usia saat ukur sudah menjadi bentuk type data *float* atau desimal. Satuan yang digunakan untuk atribut Usia saat ukur adalah dalam satuan tahun.

JK	(usia saat ukur)	Tinggi	ZS TB/U	TB/U
P	0.39	67	1.57	Normal
L	0.63	67	-1.44	Normal
L	0.74	67	-2.23	Pendek
P	0.51	68	-1.68	Normal
L	0.63	69	-0.51	Normal
P	0.53	69	-0.38	Normal
L	0.53	70	0.76	Normal
L	1.12	71.9	-2.24	Pendek
P	1.3	72	-2.26	Pendek

Gambar 3. Data Status Gizi TB/U setelah diolah.

b. Implementasi Naive Bayes

Implementasi menggunakan algoritma *Naive Bayes* terhadap data status gizi di Puskesmas Jogorogo. Data yang terdiri dari 1783 *record* pada Agustus 2023 dan 2147 *record* pada Februari 2024. Hanya menggunakan 5 atribut terdiri dari Jenis Kelamin (JK), Usia saat Ukur, Tinggi, ZS TB/U, TB/U. setelah melakukan penghapusan *missing value* data diperoleh menjadi 1775 *record* untuk Agustus 2023 dan 2141 *record* Februari 2024, kemudian data diolah menjadi data *training* dan data *testing* pembagian data dengan perbandingan 8:2, pengimplementasian algoritma *Naive Bayes* pada klasifikasi status gizi berdasarkan pengukuran antropometri dilakukan menggunakan *Google Colab* sehingga diperoleh hasil performance pada gambar (4) dan (5) sebagai berikut:

	precision	recall	f1-score	support
Normal	0.93	1.00	0.96	306
Pendek	0.80	0.39	0.52	31
Sangat Pendek	0.57	0.36	0.44	11
Tinggi	0.83	0.71	0.77	7
accuracy			0.92	355
macro avg	0.78	0.62	0.67	355
weighted avg	0.91	0.92	0.91	355

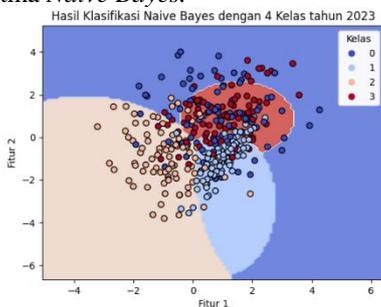
Gambar 4. Hasil Performance untuk data 2023

	precision	recall	f1-score	support
Normal	0.93	1.00	0.96	384
Pendek	0.00	0.00	0.00	23
Sangat Pendek	0.78	0.47	0.58	15
Tinggi	1.00	0.43	0.60	7
accuracy			0.92	429
macro avg	0.68	0.47	0.54	429
weighted avg	0.88	0.92	0.89	429

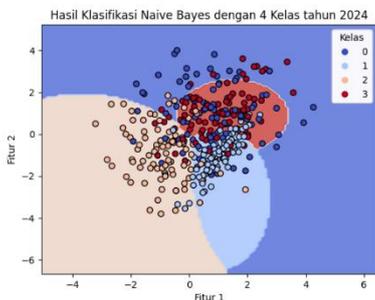
Gambar 5. Hasil Performance untuk data 2024

c. Analisa Hasil

Pada analisa hasil akan dijelaskan hasil implementasi algoritma *Naive Bayes* dalam klasifikasi status gizi TB/U yang telah diperoleh dari Puskesmas Jogorogo. Berdasarkan data status gizi, data tersebut di olah menggunakan *tools Google Colab* dengan bahasa pemograman *Python*. Secara rinci hasil prediksi tersebut dapat dilihat pada gambar 5 dan 6 yaitu gambar visualisasi hasil implementasi algoritma *Naive Bayes* untuk klasifikasi status gizi TB/U. berdasarkan data status gizi TB/U tahun 2023 dan 2024 memiliki 4 kelas yaitu kelas 0 dengan warna biru tua “Normal”, kelas 1 dengan warna biru muda ”pendek”, kelas 2 dengan warna cream “sangat pendek” dan kelas 3 dengan warna merah “tinggi”. Berikut hasil pada gambar visualisasi hasil klasifikasi status gizi TB/U untuk data 2023 dan 2024 di analisa bahwa hasil dari implementasi algoritma *Naive Bayes*.



Gambar 6. Visualisasi hasil klasifikasi untuk data 2023



Gambar 7. Visualisasi hasil klasifikasi untuk data 2024

Hasil pada visualisasi klasifikasi untuk data 2023 dan 2024 dapat dianalisa bahwa hasil implementasi algoritma *Naive Bayes* untuk klasifikasi status gizi TB/U pada Puskesmas Jogorogo adalah berupa nilai performance dengan

akurasi yang baik untuk di tahun 2023 91,83% dan pada tahun 2024 91,84%. Nilai akurasi yang dihasilkan dapat memberikan kemudahan untuk penggalian informasi mengenai klasifikasi tinggi badan menurut umur melalui data status gizi TB/U dari tahun 2023 dan 2024 agar seorang ibu yang mempunyai balita mengetahui seorang anak yang telah mencapai usianya mengetahui berapa tinggi badanya yang sesuai. Pemisah kelas untuk model *Naive Bayes* mampu memisahkan empat kelas (0, 1, 2, 3) dengan cukup baik pada gambar (6) dan gambar (7). Setiap area berwarna menunjukkan wilayah keputusan untuk masing-masing kelas. Distribusi data pada titik-titik data tersebar dalam area yang sesuai dengan kelas yang telah diidentifikasi. Namun, beberapa titik berada di luar area keputusan yang diharapkan, yang mengindikasikan kesalahan klasifikasi atau overlap antara kelas yang berbeda. Tidak seimbang kelas pada suatu dataset. Terlihat ada sedikit overlap antar kelas, terutama pada bagian antara kelas 1 dan 3, serta kelas 0 dan 2. Ini menunjukkan beberapa kesamaan dalam distribusi fitur antar kelas yang dapat menjadi tantangan dalam memisahkan mereka secara sempurna. Berdasarkan visualisasi ini, model cenderung memberikan hasil klasifikasi yang baik, namun perlu evaluasi lebih lanjut menggunakan metrik akurasi atau confusion matrix untuk mengukur seberapa baik performa model dalam mengklasifikasi seluruh data.

5. Kesimpulan

Berdasarkan hasil penelitian dan pengolahan data terhadap data status gizi balita pada Puskesmas Jogorogo dapat menyimpulkan bahwa metode *Naive Bayes* bisa digunakan untuk mengklasifikasikan dalam status gizi balita TB/U dan dapat digunakan sebagai informasi seorang ibu yang mempunyai balita mengetahui seorang anak yang telah mencapai usianya mengetahui berapa tinggi badanya yang sesuai. Pengujian data 1775 untuk tahun 2023 dan 2147 untuk tahun 2024 dengan perhitungan *google colab* dihasilkan nilai akurasi 91, 83% untuk tahun 2023 dan 91, 84% untuk tahun 2024. Dan ketidakseimbangannya sebuah kelas bisa juga dapat menyebabkan overlap antar kelas dikarenakan Ini menunjukkan beberapa kesamaan dalam distribusi fitur antar kelas yang dapat menjadi tantangan dalam memisahkan mereka secara sempurna.

Ucapan Terima Kasih

Ucapan terimakasih saya haturkan kepada Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan paper tepat pada waktunya, ucapan terimakasih juga saya haturkan kepada bapak Widya Kurniawan dan bapak Dihin

Muriyatmoko atas segala arahan, bimbingan serta masukan selama penulisan paper ini berlangsung, juga kepada kedua orang tua atas dukungan moral dan materil, serta kawan-kawan satu perjuangan Teknik Informatika.

Daftar Pustaka

- [1] H. B. Setiawan and G. P. Utama, "Klasifikasi Status Gizi Balita Menggunakan Metode Naïve Bayes Classifier Classification Of The Nutritional Status Of Toddlers Using The Nave Bayes Classifier Method," no. September, pp. 707–715, 2022.
- [2] E. Haerani, F. Syafria, and L. Oktavia, "Penerapan Algoritma Naïve Bayes Classifier Dalam Klasifikasi Status Gizi Balita dengan Pengujian K-Fold Cross Validation," vol. 4, no. 3, pp. 578–586, 2023, doi: 10.47065/josyc.v4i3.3414.
- [3] "KLASIFIKASI GIZI BURUK PADA BALITA MENGGUNAKAN ALGORITMA C5 . 0 (STUDI KASUS : DINAS KESEHATAN ACEH TIMUR) SKRIPSI Disusun Sebagai Syarat Memperoleh Gelar Sarjana Komputer Prodi Teknik Informatika Fakultas Teknik Universitas Malikussaleh DISUSUN OLEH : ," vol. 0, 2024.
- [4] P. N. 2 T. 2020 T. S. A. Anak, "No Title," no. 3, pp. 1–78, 2020.
- [5] A. M. Majid, R. Nuraeni, A. H. Anshor, T. Informatika, and U. P. Bangsa, "Jurnal Teknologi Pelita Bangsa Penerapan Data Mining Untuk Penentuan Status Gizi Balita Menggunakan Algoritma Naive Bayes (Studi Kasus Posyandu Seruni XII Kelurahan Karangraharja)," vol. 12, no. 1, pp. 1–7, 2020.
- [6] D. H. Ramadhani, R. Srikandi, M. Ikhwan, and R. A. Saputra, "Penerapan Logika Fuzzy Dalam Klasifikasi Status Gizi Balita Di Puskesmas Pondidaha Menggunakan Metode Fuzzy Tsukamoto," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4017.
- [7] Y. Mulyanto, F. Idifitriani, A. Wati, U. T. Sumbawa, D. Mining, and K. P. Tano, "Vol 7 No 2 , September 2024 KLASIFIKASI DATA MINING UNTUK PENENTUAN STUNTING," vol. 7, no. 2, pp. 119–125, 2024.
- [8] H. I. Islam, M. Khandava Mulyadien, U. Enri, U. Singaperbangsa, and K. Abstract, "Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 10, pp. 116–125, 2022, [Online]. Available: <https://doi.org/10.5281/zenodo.6791722>