

IMPLEMENTASI ALGORITMA NAIVE BAYES PADA APLIKASI WEB UNTUK KLASIFIKASI STATUS GIZI BALITA

Five Bunda De Putri Dopo¹, Sisilia Daeng Bakka Mau², Frengky Tedy³, Paul Filson M. Tengangatu⁴

^{1,2,3,4} Program Studi Ilmu Komputer, Fakultas Teknik, Universitas Katolik Widya Mandira, Kupang-Indonesia
¹dopoputri@gmail.com, ²sisiliamau@unwira.ac.id, ³frengkytedy@unwira.ac.id, ⁴paultengangatu@unwira.ac.id

Abstrak

Masalah gizi pada anak usia di bawah lima tahun hingga saat ini masih menjadi persoalan utama di Indonesia, termasuk di wilayah Kabupaten Manggarai Timur. Kebutuhan akan proses penilaian status gizi yang efektif dan efisien, serta masih terbatasnya jumlah tenaga kesehatan yang tersedia, menjadi faktor pendorong untuk diterapkannya pemanfaatan teknologi informasi. Tujuan dari penelitian ini, untuk mengimplementasikan metode Naive Bayes pada aplikasi web yang dapat mengklasifikasikan status gizi pada balita dengan menggunakan tiga indikator, yaitu Berat Badan per Umur (BB/U), Tinggi Badan per Umur (TB/U), dan Berat Badan per Tinggi Badan (BB/TB). Model klasifikasi dilatih dengan menggunakan dataset sebanyak 500 data rekam medis balita yang berasal dari Puskesmas Borong, Kabupaten Manggarai Timur, dari tahun 2019 -2023, dengan skenario 80% digunakan sebagai data latih (400 data) dan 20% sebagai data uji (100 data) pada masing-masing indikator. Variabel kontinu pada model diklasifikasikan menggunakan pendekatan distribusi Gaussian. Berdasarkan hasil pengujian yang di evaluasi menggunakan Confusion Matrix dan Classification Report, memperlihatkan penerapan metode Naive Bayes dalam platform web pada penelitian ini memiliki kinerja yang andal dalam melakukan klasifikasi dengan rata-rata akurasi sebesar 97.0%, presisi 0.97, recall 0.96, dan F1-score 0.97 sehingga dapat digunakan sebagai alat bantu keputusan (DSS) bagi tenaga kesehatan di Puskesmas dalam menentukan status gizi balita secara cepat, efektif dalam meminimalisir kesalahan, serta memastikan data tersimpan secara sistematis.

Kata kunci: *Naive Bayes*, Aplikasi Web, Klasifikasi, Status Gizi Balita, *Confusion Matrix*

1. Pendahuluan

Fase awal pertumbuhan seorang anak, khususnya ketika memasuki usia di bawah lima tahun, sering dianggap sebagai masa keemasan (*golden age*), yakni periode krusial yang berperan besar dalam membentuk kualitas perkembangan fisik, kesehatan dan kecerdasan anak di masa depan. Pada rentang waktu ini, pemenuhan gizi yang memadai menjadi faktor utama agar proses tumbuh kembang berjalan secara optimal. Akan tetapi, hingga saat ini Indonesia masih menghadapi masalah persoalan gizi ganda, baik berupa kekurangan maupun kelebihan gizi yang tercermin dari masih tingginya prevalensi stunting, wasting dan obesitas pada balita (Andriani et al., 2023; Wulandari et al., 2025).

Kondisi serupa juga ditemukan di Kabupaten Manggarai Timur, khususnya pada wilayah kerja Puskesmas Borong. Selama lima tahun terakhir, kasus yang berkaitan dengan status gizi balita menunjukkan fluktuasi dengan total akumulasi kasus yang signifikan. Pada tahun 2019 terdapat 314 kasus, kemudian menurun menjadi 175 kasus pada 2020. Akan tetapi, angka tersebut kembali meningkat menjadi 207 kasus pada 2021, 221 kasus pada 2022, hingga mencapai 307 kasus pada 2023. Dengan total populasi data historis mencapai 1.224 rekam medis dalam lima tahun terakhir, hal ini menyediakan basis

data yang memadai untuk penelitian komputasi. Namun, mengingat pentingnya kualitas data dalam pembelajaran mesin, proses seleksi data yang ketat diperlukan untuk menghasilkan dataset yang valid dan representatif guna melatih model klasifikasi yang presisi.

Di sisi lain, proses penentuan status gizi di Puskesmas Borong hingga kini masih mengandalkan perhitungan antropometri secara manual. Cara tersebut berpotensi menimbulkan keterlambatan serta inefisiensi dalam pengolahan data, sehingga informasi yang dihasilkan sering kali kurang optimal digunakan untuk pengambilan keputusan. Oleh karena itu, situasi ini menegaskan adanya urgensi untuk mengembangkan sistem informasi berbasis teknologi yang lebih andal, sehingga klasifikasi status gizi balita dapat dilakukan secara cepat, akurat dan dapat diakses baik oleh tenaga kesehatan maupun masyarakat. Dalam konteks kebutuhan akan sistem berbasis teknologi inilah, pemilihan algoritma klasifikasi yang tepat menjadi krusial.

Sejalan dengan hal tersebut, penelitian-penelitian sebelumnya telah banyak mengeksplorasi penerapan algoritma Naive Bayes dalam klasifikasi status gizi balita, meskipun capaian akurasi yang diperoleh menunjukkan variasi yang beragam. Dimulai dari penelitian yang dilakukan oleh (Rozaq & Purnomo, 2022) pada salah satu

puskesmas di Kota Madiun, dengan mencatat akurasi sebesar 58%. Hasil tersebut berbeda cukup jauh dengan penelitian (Harliana & Anggraini, 2023) di Posyandu Desa Kalitengah yang justru memperoleh akurasi 87,33%. Penelitian lain bahkan melaporkan capaian yang lebih tinggi, yaitu yang dilakukan oleh (Saputra et al., 2023) dengan akurasi sebesar 95,77% ketika diuji pada data Posyandu Baros. Temuan ini hampir sebanding dengan penelitian (Hidayat et al., 2024) yang menggunakan data Bulan Penimbangan Balita di Kabupaten Majalengka dengan akurasi 94,65%. Tidak hanya menguji model, (Dirta et al., 2024) juga memperluas penerapan dengan merancang sistem berbasis PHP-MySQL dan hasil klasifikasinya menunjukkan konsistensi penuh antara perhitungan manual dan keluaran sistem. Konsistensi performa model semakin diperkuat oleh (Kurniawan et al., 2025) yang menganalisis data dua tahun berturut-turut dan berhasil mempertahankan rata-rata akurasi sebesar 91,8%. Penelitian (Arlita et al., 2025) menaruh perhatian pada permasalahan stunting di Kota Cirebon dan melaporkan akurasi 88%. Sementara itu, varian Gaussian Naive Bayes yang diuji oleh (Kurniawan et al., 2024) pada data dari Kota Samarinda menghasilkan akurasi sebesar 87,42%.

Berdasarkan berbagai hasil penelitian terdahulu, penerapan algoritma Naïve Bayes terbukti menampilkan kemampuan klasifikasi yang efektif pada pengelompokan data kondisi gizi anak usia dini, yang terlihat dari tingkat akurasi tinggi dan konsistensi hasil pada sebagian besar studi yang telah dilakukan. Namun demikian, sebagian besar penelitian masih terbatas pada tahap pengujian algoritma dan belum banyak yang mengembangkan sistem berbasis web yang dapat diimplementasikan secara langsung di lapangan, khususnya pada daerah dengan karakteristik gizi dan demografi yang berbeda, seperti pada Kabupaten Manggarai Timur, yang memiliki permasalahan gizi dengan variasi spasial serta dipengaruhi oleh faktor geografis, sosial dan demografi yang khas. Hasil penelitian (Supardianto & Mutawalli, 2025) menunjukkan adanya perbedaan risiko gizi antarwilayah di NTT yang menegaskan perlunya pendekatan yang disesuaikan dengan kondisi lokal. Penelitian (Karundeng et al., 2023) di Puskesmas Borong juga memperlihatkan bahwa kepatuhan konsumsi tablet zat besi pada ibu hamil merupakan faktor penting yang berdampak langsung pada kondisi gizi, sehingga konteks lokal perlu mendapat perhatian serius.

Di sisi lain, penelitian (DwiYanti & Saptari, 2023) telah merancang sebuah sistem informasi status gizi balita berbasis web dengan menggunakan metode Naive Bayes dan K-Nearest Neighbor, serta penelitian (Yoghatama & Pamungkas, 2025) yang mengembangkan platform web GiziMu Mijen untuk edukasi gizi dan deteksi stunting, kedua hasil penelitian tersebut membuktikan bahwa aplikasi web dapat menjadi sarana efektif untuk memperoleh

informasi gizi secara cepat, akurat dan mudah diakses tenaga kesehatan sehingga dalam penelitian ini memiliki tujuan untuk mengimplementasikan algoritma Naive Bayes pada aplikasi web guna mengklasifikasi status gizi balita di Puskesmas Borong, Kabupaten Manggarai Timur. Diharapkan sistem ini memberikan dukungan bagi tenaga kesehatan dalam proses identifikasi status gizi balita dengan efisiensi waktu yang lebih baik, serta tingkat akurasi yang cukup memadai pada tahap awal pengujian. Di sisi lain, sistem ini merupakan bentuk inovasi penerapan teknologi informasi dalam mendukung strategi lokal untuk pencegahan maupun penanggulangan masalah gizi balita.

2. Metode

Penelitian ini dilaksanakan melalui pendekatan kuantitatif eksperimental yang berlokasi di Puskesmas Borong, Kabupaten Manggarai Timur. Metode penelitian dirancang secara sistematis mulai dari akuisisi data rekam medis, penentuan kriteria label gizi, penerapan algoritma pembelajaran mesin, hingga evaluasi kinerja model. Tahapan ini bertujuan untuk memastikan bahwa sistem klasifikasi yang dibangun memiliki landasan data yang valid dan prosedur komputasi yang terukur. Rincian mengenai komponen data, algoritma, dan tahapan pengembangan sistem dijabarkan pada bagian-bagian berikut.

2.1. Pengumpulan dan Pengolahan Data

Data penelitian diperoleh dari rekam medis Puskesmas Borong periode tahun 2019 hingga 2023. Dari total populasi data yang terkumpul pada tahap awal berjumlah 1.224 rekam medis. Untuk menjamin kualitas model klasifikasi, dilakukan tahapan data preprocessing yang ketat melalui beberapa *fase* seleksi.

Fase pertama adalah pembersihan data duplikat dan data tidak lengkap (*missing values*), di mana sistem mengeliminasi 270 data yang teridentifikasi duplikat (pencatatan ganda pada pasien yang sama dalam periode yang sama) dan menghapus 404 data yang memiliki atribut tidak lengkap (*missing values*), seperti kekosongan pada kolom Tinggi Badan atau Berat Badan yang krusial untuk perhitungan.

Fase kedua adalah eliminasi *outlier* untuk membuang data yang tidak logis secara medis atau kesalahan input (*human error*). Kriteria *outlier* yang ditetapkan dalam penelitian ini adalah data dengan nilai yang tidak mungkin secara biologis untuk balita, seperti Tinggi Badan > 120 cm untuk usia di bawah 24 bulan atau nilai Berat Badan yang ekstrem ($Z\text{-Score} > +5\text{ SD}$ atau $< -5\text{ SD}$). Pada tahap ini, sebanyak 50 data *outlier* di eliminasi.

Hasil dari proses seleksi tersebut menghasilkan 500 data balita bersih (*clean dataset*) yang digunakan sebagai dataset utama penelitian ini. Atribut yang

digunakan meliputi Jenis Kelamin, Umur (bulan), Berat Badan (kg), dan Tinggi Badan (cm).

Untuk keperluan eksperimen dan validasi model, dataset dibagi menggunakan metode *Hold-out Validation* dengan rasio 80:20, sebagaimana dirincikan pada Tabel 1.

Tabel 1. Distribusi Dataset Penelitian		
Keterangan	Jumlah Data	Persentase
Data Latih	400	80%
Data Uji	100	20%
Total	500	100%

2.2. Kriteria Status Gizi (Labeling)

Penentuan kelas atau label klasifikasi merujuk pada Standar Antropometri Anak sesuai Peraturan Menteri Kesehatan RI No. 2 Tahun 2020 yang mengadopsi standar WHO (Kementerian Kesehatan, 2020.). Kriteria status gizi dibagi berdasarkan Z-Score untuk tiga indikator utama seperti terlihat pada Tabel 2.

Tabel 2. Kriteria Klasifikasi Status Gizi		
Indikator	Z-Score (SD)	Klasifikasi (Kelas)
BB/U (Berat/Umur)	< -3 SD	Berat Badan Sangat Kurang
	-3 SD s.d. < -2 SD	Berat Badan Kurang
	-2 SD s.d. +1 SD	Berat Badan Normal
	> +1 SD	Risiko Berat Badan Lebih
TB/U (Tinggi/Umur)	< -3 SD	Sangat Pendek (<i>Severely Stunted</i>)
	-3 SD s.d. < -2 SD	Pendek (<i>Stunted</i>)
	-2 SD s.d. +3 SD	Normal
	> +3 SD	Tinggi
BB/TB (Berat/Tinggi)	< -3 SD	Gizi Buruk
	-3 SD s.d. < -2 SD	Gizi Kurang
	-2 SD s.d. +1 SD	Gizi Baik
	> +1 SD s.d. +2 SD	Risiko Gizi Lebih
	> +2 SD s.d. +3 SD	Gizi Lebih
	> +3 SD	Obesitas

2.3. Algoritma Gaussian Naive Bayes

Metode yang digunakan adalah Naive Bayes, sebuah metode klasifikasi probabilistik berdasarkan Teorema Bayes. Karena atribut data antropometri (berat, tinggi) bersifat numerik kontinu, maka digunakan pendekatan distribusi Gaussian.

Secara umum, Teorema Bayes dinyatakan sebagai:

$$P(h|D) = \frac{P(D|h) \cdot P(h)}{P(D)} \quad (1)$$

Dengan $P(h|D)$ adalah probabilitas posterior (peluang kelas h benar setelah data D diamati), $P(D|h)$ adalah *likelihood* (peluang munculnya data D jika kelas h benar), $P(h)$ merupakan prior *probability* (peluang awal kelas h), dan $P(D)$ adalah *evidence* atau probabilitas data.

Karena data bersifat kontinu, *Likelihood* dihitung menggunakan fungsi densitas probabilitas Gaussian, yang secara umum dirumuskan sebagai:

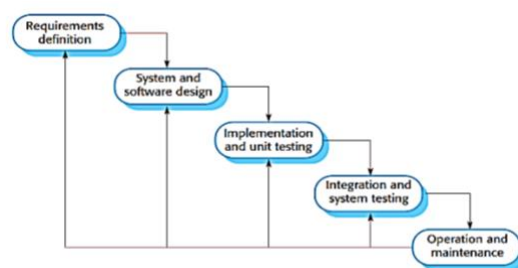
$$P(x_i|h) = \frac{1}{\sqrt{2\pi\sigma_h^2}} \exp\left(-\frac{(x_i - \mu_h)^2}{2\sigma_h^2}\right) \quad (2)$$

Dimana μ_h adalah rata-rata (*mean*) dan σ_h^2 adalah varians dari atribut pada kelas h . Nilai prediksi kelas akhir ($\hat{h}$) ditentukan dengan memilih probabilitas posterior tertinggi menggunakan rumus:

$$\hat{h} = \arg \max_h P(h) \prod_{i=1}^n P(x_i|h) \quad (3)$$

2.4. Metode Pengembangan Sistem

Kegiatan penelitian ini dilaksanakan di area pelayanan Puskesmas Borong, Kabupaten Manggarai Timur. Subjek yang dianalisis adalah balita dengan data status gizi yang terekam secara rutin pada kurun waktu 2019 hingga 2023. Data utama yang digunakan adalah hasil pengukuran antropometri, yang mencakup jenis kelamin, berat badan, tinggi badan dan umur, yang kemudian dijadikan dasar dalam proses pengembangan sistem klasifikasi status gizi. Untuk mendukung pengembangan tersebut, alur penelitian dirancang dengan mengacu pada model *Waterfall* sebagaimana ditunjukkan pada Gambar 1 (Sommerville, 2016). Model ini dipilih karena menyajikan tahapan pengembangan yang sistematis dan berurutan, sehingga setiap keluaran dari suatu tahap dapat digunakan sebagai masukan bagi tahap berikutnya.



Gambar 1. Model *Waterfall*

Proses penelitian ini dilaksanakan melalui beberapa tahapan yang dijabarkan pada bagian berikut:

1. Requirements Definition

Tahap *Requirements Definition* diawali dengan mengkaji permasalahan utama dan merumuskan kebutuhan sistem berdasarkan hasil wawancara dengan tenaga kesehatan serta melakukan telaah laporan status gizi balita di Puskesmas Borong. Dari kajian tersebut, ditetapkan kebutuhan fungsional sistem, seperti fitur input data, klasifikasi status gizi, pelaporan hasil dan kebutuhan non-fungsional yang menekankan pada aspek kecepatan, kemudahan akses, dan akurasi, yang selanjutnya dijadikan dasar dalam pengembangan aplikasi.

2. System and Software Design

Dalam tahapan ini dilakukan perancangan menyeluruh terhadap komponen utama aplikasi. Kegiatan ini meliputi permodelan arsitektur perangkat lunak, pemetaan proses informasi menggunakan diagram alir data (*Data Flow Diagram*), serta penyusunan struktur basis data melalui diagram entitas dan relasi (*Entity Relationship Diagram*). Selain itu, dirancang pula tampilan antarmuka berbasis web agar sistem dapat digunakan secara mudah oleh petugas kesehatan maupun pengguna umum.

3. *Implementation and Unit Testing*

Rancangan yang telah disusun kemudian diterjemahkan ke bentuk kode program menggunakan bahasa PHP dan sistem pengelolaan basis data MySQL. Di dalam aplikasi ini, algoritma Naive Bayes diterapkan untuk melakukan proses pengelompokan kondisi gizi balita. Selama tahap ini, setiap modul diuji secara terpisah (*unit testing*) guna memastikan fungsi tiap komponen berjalan sesuai rancangan.

4. *Integration and System Testing*

Setelah seluruh modul berhasil melalui tahap uji individual, modul-modul tersebut digabungkan menjadi satu kesatuan sistem. Pengujian menyeluruh dilakukan menggunakan pendekatan black-box testing untuk menilai fungsionalitas aplikasi, sekaligus melakukan validasi hasil klasifikasi dengan perhitungan manual sehingga dapat diketahui konsistensi kinerja model. Selain menilai performa berdasarkan tingkat ketepatan (*accuracy*), evaluasi sistem juga diperluas dengan analisis presisi (*metrik precision*), sensitivitas (*recall*), dan nilai F1 (*F1-score*) agar kemampuan model dapat dinilai secara lebih menyeluruh. *Metrik precision* mengukur perbandingan antara prediksi benar terhadap seluruh hasil prediksi positif, *recall* menggambarkan sejauh mana model mampu mengenali seluruh data positif aktual, sedangkan nilai F1 menunjukkan rata-rata harmonis dari kedua metrik tersebut. Setiap nilai dihitung berdasarkan hasil prediksi terhadap data uji menggunakan formula berikut:

$$\text{Precision} = \frac{TP}{TP+FP}, \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN}, \quad (5)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

dengan TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*) merupakan nilai yang diperoleh dari hasil perbandingan antara keluaran prediksi sistem dan data aktual pada masing-masing kategori status gizi.

5. *Operation and Maintenance*

Setelah melewati tahap pengujian, sistem belum langsung diterapkan pada lingkungan nyata, melainkan masih diuji melalui simulasi dengan menggunakan data status gizi balita yang telah tersedia. Tahap ini dimaksudkan untuk mengevaluasi apakah fungsi sistem berjalan sesuai rancangan serta menilai tingkat keakuratan hasil klasifikasi. Dengan demikian, penelitian ini berhenti pada tahap simulasi pengujian, sehingga proses pemeliharaan di lapangan belum dilakukan, meskipun peluang pengembangan lebih lanjut tetap terbuka untuk kebutuhan implementasi nyata di Puskesmas Borong.

3. Hasil dan Pembahasan

Hasil penelitian ini dipaparkan dalam dua perspektif utama, yaitu hasil pengembangan perangkat lunak dan hasil analisis kinerja algoritma. Bagian implementasi sistem menampilkan antarmuka aplikasi web yang telah dibangun untuk memfasilitasi pengguna di Puskesmas Borong. Selanjutnya, bagian evaluasi kinerja model menyajikan analisis kuantitatif berdasarkan pengujian terhadap dataset baru, yang mencakup perhitungan *Confusion Matrix* serta laporan klasifikasi (*Classification Report*) untuk mengukur tingkat keberhasilan metode Gaussian Naive Bayes dalam mengenali pola status gizi balita.

3.1. Evaluasi Kinerja Model

Untuk memvalidasi keandalan sistem secara menyeluruh, pengujian dilakukan menggunakan 100 data uji (20% dari total dataset) yang dialokasikan secara acak dan belum pernah dilihat oleh model selama proses pelatihan. Evaluasi ini bertujuan untuk mengukur tingkat presisi dan akurasi sistem dalam memprediksi kelas status gizi pada ketiga indikator antropometri standar, yaitu Berat Badan menurut Umur (BB/U), Tinggi Badan menurut Umur (TB/U), dan Berat Badan menurut Tinggi Badan (BB/TB). Analisis hasil klasifikasi untuk setiap indikator disajikan secara rinci melalui *Confusion Matrix* guna memetakan distribusi prediksi benar dan salah pada masing-masing kelas, yang kemudian dirangkum dalam rekapitulasi performa akhir.

A. Evaluasi Indikator Berat Badan Menurut Umur (BB/U)

Indikator BB/U digunakan untuk mendeteksi masalah berat badan secara umum yang sifatnya labil dan mudah berubah (sensitif). Tantangan utama pada indikator ini adalah variabilitas berat badan balita yang dinamis akibat faktor kesehatan sesaat. Hasil pengujian performa model terhadap 100 data uji untuk indikator ini disajikan secara rinci pada Tabel 3 (*Confusion Matrix*) dan Tabel 4 (*Classification Report*).

Berdasarkan kedua tabel tersebut, algoritma Naive Bayes mampu mengidentifikasi pola distribusi

berat badan dengan tingkat akurasi sebesar 96,0%. Secara spesifik pada Tabel 3, meskipun mayoritas data terklasifikasi dengan benar, terdapat beberapa kesalahan prediksi (*misclassification*) yang perlu dicermati. Misalnya, terdapat 1 data aktual "Sangat Kurang" yang diprediksi sebagai "Kurang", dan sebaliknya 1 data aktual "Kurang" diprediksi "Sangat Kurang". Kesalahan deteksi ini terjadi karena nilai Z-Score data tersebut berada tepat di area ambang batas (*boundary*) atau perbatasan antar kelas (misalnya mendekati nilai -3 SD). Pada area perbatasan ini, kurva distribusi probabilitas Gaussian dari kedua kelas cenderung saling beririsan (*overlapping*), sehingga probabilitas posterior yang dihasilkan memiliki selisih yang sangat tipis yang menyebabkan pergeseran prediksi ke kelas terdekat. Meskipun terdapat sedikit kesalahan prediksi (*misclassification*) pada kelas yang berdekatan, nilai rata-rata *Weighted Avg* untuk presisi dan *recall* yang mencapai 0.96 menunjukkan bahwa model tetap konsisten dalam membedakan kategori berat badan normal maupun risiko masalah gizi.

Tabel 3. Confusion Matrix Indikator BB/U

	Prediksi Sangat Kurang	Prediksi Kurang	Prediksi Normal	Prediksi Resiko Lebih
Aktual Sangat Kurang	12	1	0	0
Aktual Kurang	1	18	1	0
Aktual Normal	0	1	56	0
Aktual Risiko Lebih	0	0	0	10

Tabel 4. Classification Report Indikator BB/U

Kelas	Precision	Recall	F1-Score
Sangat Kurang	0.92	0.92	0.92
Kurang	0.90	0.90	0.90
Normal	0.98	0.98	0.98
Resiko Lebih	1.00	1.00	1.00
Accuracy			0.96
Macro Avg	0.95	0.95	0.95
Weighted Avg	0.96	0.96	0.96

B. Evaluasi Indikator Tinggi Badan Menurut Umur (TB/U)

Berbeda dengan berat badan, indikator TB/U merepresentasikan status gizi kronis yang menggambarkan pertumbuhan linear anak. Distribusi data tinggi badan cenderung lebih stabil, memiliki pola deterministik yang kuat, dan jarang mengalami fluktuasi drastis dalam waktu singkat. Karakteristik data yang stabil ini tercermin dari hasil evaluasi yang ditampilkan pada Tabel 5 dan Tabel 6.

Sistem mencatatkan performa klasifikasi tertinggi pada indikator ini dengan akurasi mencapai 98,0%. Analisis pada Tabel 5 menunjukkan kesalahan prediksi yang sangat minim, yaitu hanya terjadi pada perbatasan kelas "Pendek" dan "Normal". Terdapat 1 kasus balita "Pendek" yang terdeteksi "Normal", dan 1 kasus "Normal" terdeteksi "Pendek". Hal ini mengindikasikan bahwa meskipun data tinggi badan cenderung stabil, tantangan tetap muncul pada

data yang memiliki nilai tepat di garis batas Z-Score -2 SD. Minimnya kesalahan prediksi pada *Confusion Matrix* serta nilai *Precision* dan *Recall* yang nyaris sempurna pada *Classification Report* menegaskan bahwa algoritma Gaussian Naive Bayes sangat efektif dalam memetakan kurva pertumbuhan tinggi badan yang mengikuti distribusi normal baku.

Tabel 5. Confusion Matrix Indikator TB/U

	Prediksi Sangat Pendek	Prediksi Pendek	Prediksi Normal	Prediksi Tinggi
Aktual Sangat Pendek	8	0	0	0
Aktual Pendek	0	14	1	0
Aktual Normal	0	1	68	0
Aktual Tinggi	0	0	0	8

Tabel 6. Classification Report Indikator TB/U

Kelas	Precision	Recall	F1-Score
Sangat Pendek	1.00	1.00	1.00
Pendek	0.93	0.93	0.93
Normal	0.99	0.99	0.99
Tinggi	1.00	1.00	1.00
Accuracy			0.98
Macro Avg	0.98	0.97	0.98
Weighted Avg	0.98	0.98	0.98

C. Evaluasi Indikator Berat Badan Menurut Tinggi Badan (BB/TB)

Indikator BB/TB merupakan parameter paling kritis dan kompleks karena melibatkan interaksi dua variabel pengukuran fisik sekaligus untuk menentukan status gizi akut (seperti gizi buruk/wasting) serta kegemukan (obesitas). Oleh karena itu, evaluasi pada indikator ini menjadi tolok ukur utama keandalan sistem. Detail kinerja model untuk indikator BB/TB dapat dilihat pada Tabel 7 dan Tabel 8.

Hasil pengujian menunjukkan bahwa sistem berhasil menjaga keseimbangan performa dengan akurasi sebesar 97,0%. Berdasarkan Tabel 7, terlihat adanya 1 data aktual "Gizi Kurang" yang terklasifikasi sebagai "Gizi Baik", serta sedikit ambiguitas antara kelas "Gizi Baik" dan "Gizi Lebih". Sama seperti indikator sebelumnya, kesalahan klasifikasi ini bersifat *borderline error*. Data tersebut memiliki karakteristik fisik yang berada di area transisi antara status normal dan risiko penyimpangan ringan, sehingga model memprediksi ke kelas mayoritas terdekat. Sebagaimana terlihat pada *Classification Report*, nilai *F1-Score* rata-rata sebesar 0.96 mengindikasikan bahwa model tidak hanya akurat pada kelas mayoritas (Gizi Baik), tetapi juga sangat sensitif dalam mendeteksi kelas ekstrem seperti Gizi Buruk dan Obesitas dengan presisi sempurna (1.00). Hal ini membuktikan bahwa sistem layak digunakan sebagai alat deteksi dini kasus gizi prioritas.

Tabel 7. Confusion Matrix Indikator BB/TB

Prediksi Gizi Buruk	Prediksi Gizi Kurang	Prediksi Gizi Baik	Prediksi Gizi Lebih	Prediksi Obesitas
---------------------	----------------------	--------------------	---------------------	-------------------

Aktual Gizi Buruk	8	0	0	0	0
Aktual Gizi Kurang	0	14	1	0	0
Aktual Gizi Baik	0	0	58	1	0
Aktual Gizi Lebih	0	0	1	10	0
Aktual Obesitas	0	0	0	0	7

Tabel 8. Classification Report BB/TB

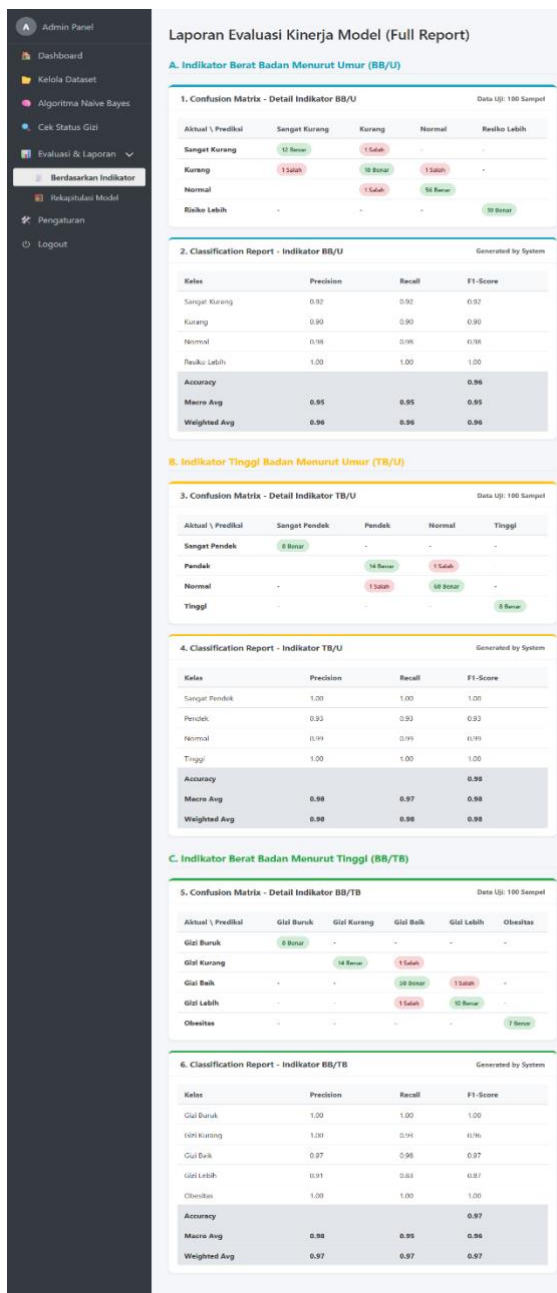
Kelas	Precision	Recall	F1-Score
Gizi Buruk	1.00	1.00	1.00
Gizi Kurang	1.00	0.93	0.96
Gizi Baik	0.97	0.98	0.97
Gizi Lebih	0.91	0.83	0.87
Obesitas	1.00	1.00	1.00
Accuracy			0.97
Macro Avg	0.98	0.95	0.96
Weighted Avg	0.97	0.97	0.97

3.2. Rekapitulasi Akhir Performa Model

Sebelum masuk pada evaluasi statistik, penting untuk memastikan bahwa alur data pada aplikasi berjalan sesuai rancangan. Proses klasifikasi dimulai dari input data balita pada antarmuka aplikasi sebagaimana terlihat pada Gambar 2. Data ini kemudian diproses oleh sistem menggunakan algoritma Gaussian Naive Bayes yang telah dilatih.

Gambar 2. Halaman Input Data Balita

Setelah dilakukan evaluasi mendalam, Tabel 9 menyajikan ringkasan kinerja sistem secara keseluruhan. Seluruh metrik statistik yang tercantum dalam rekapitulasi ini dihasilkan dan direkam secara otomatis oleh fitur evaluasi pada aplikasi web. Hal ini memastikan bahwa evaluasi kinerja dilakukan secara transparan dan terukur, serta menjawab kebutuhan validasi data secara *real-time*. Bukti otentikasi perhitungan sistem yang mencakup *Confusion Matrix* dan *Classification Report* lengkap untuk ketiga indikator (BB/U, TB/U, dan BB/TB) ditampilkan pada Gambar 3.



Gambar 3. Tampilan Output Program Evaluasi Kinerja (Full Report: Confusion Matrix & Classification Report)

Berdasarkan *report* yang dihasilkan sistem di atas, rata-rata akurasi sistem untuk ketiga indikator berada pada angka 97,0%. Angka ini menunjukkan bahwa metode Gaussian Naive Bayes telah sejalan dengan formulasi dalam kerangka Teorema Bayes (Murphy, 2012). Tingginya konsistensi nilai *Precision*, *Recall*, dan *F1-Score* mengonfirmasi bahwa penambahan jumlah dataset latih menjadi 400 data dan data uji 100 data telah berhasil meningkatkan kemampuan generalisasi model secara signifikan.

Ringkasan angka-angka tersebut kemudian diagregasi oleh sistem ke dalam halaman rekapitulasi seperti terlihat pada Gambar 4, yang datanya disalin kembali ke dalam Tabel 9 untuk memudahkan pembacaan.

3. Rekapitulasi Performa Model (Semua Indikator)		Global Summary		
Indikator	Akurasi	Rata-Rata Precision	Rata-Rata Recall	Rata-Rata F1-Score
BB/U (Berat/Umar)	96.0%	0.95	0.95	0.95
TB/U (Tinggi/Umar)	98.0%	0.96	0.97	0.98
BB/TB (Berat/Tinggi)	97.0%	0.97	0.97	0.97
RATA-RATA SISTEM	97.0%	0.97	0.96	0.97

Gambar 4. Tampilan Dasbor Rekapitulasi Model pada Aplikasi

Tabel 9. Rekapitulasi Performa Model

Indikator	Akurasi	Rata-Rata Precision	Rata-rata Recall	Rata-rata F1-Score
BB/U	96.0%	0.95	0.95	0.95
TB/U	98.0%	0.98	0.97	0.98
BB/TB	97.0%	0.97	0.97	0.97
Rata-Rata Sistem	97.0%	0.97	0.96	0.97

Hasil akurasi sebesar 97,0% ini merupakan capaian yang kompetitif jika dibandingkan dengan penelitian-penelitian terdahulu. Sebagai pembanding, Saputra et al., (2023) mencatat akurasi 95,77% pada data Posyandu Baros, dan Hidayat et al., (2024) memperoleh akurasi 94,65% di Kabupaten Majalengka. Sementara Kurniawan et al., (2025) mampu mempertahankan rata-rata akurasi 91,8% selama dua tahun berturut-turut. Hal ini membuktikan bahwa dengan proporsi data latih dan uji yang memadai, model yang dikembangkan dalam penelitian ini mampu melampaui atau setidaknya menyamai standar performa klasifikasi status gizi yang ada saat ini.

Secara praktis, tingginya akurasi ini berimplikasi positif terhadap operasional di lapangan. Aplikasi berbasis web yang dikembangkan berpotensi besar dimanfaatkan sebagai media pendukung keputusan (DSS) di Puskesmas Borong, Kabupaten Manggarai Timur. Kehadiran sistem ini memungkinkan proses deteksi status gizi berlangsung lebih cepat, terukur, serta terdokumentasi dengan baik dibandingkan cara manual yang umumnya membutuhkan waktu lebih panjang dan rentan *human error*.

Terlepas dari capaian dan potensi manfaat tersebut, penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil. Pertama, dataset yang digunakan masih menghadapi tantangan ketidakseimbangan kelas (*class imbalance*), di mana jumlah data pada kategori status gizi normal jauh lebih dominan dibandingkan kategori ekstrem (seperti gizi buruk atau obesitas). Pada algoritma Naive Bayes, kondisi ini berpotensi memengaruhi nilai probabilitas prior, yang dapat menyebabkan bias prediksi ke arah kelas mayoritas jika ditemukan data

dengan karakteristik yang ambigu. Kedua, validasi model saat ini hanya menerapkan skenario pembagian data tunggal (*single-split*) dengan rasio 80:20 dan belum menggunakan metode validasi silang (*K-Fold Cross Validation*), sehingga stabilitas performa model terhadap berbagai variasi acak data belum teruji sepenuhnya. Terakhir, karena data hanya bersumber dari satu lokasi spesifik (Puskesmas Borong), generalisasi model terhadap populasi balita di wilayah geografis lain mungkin memerlukan penyesuaian ulang.

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan algoritma Naive Bayes pada aplikasi web untuk klasifikasi status gizi balita di Puskesmas Borong dengan memanfaatkan dataset sebanyak 500 rekam medis. Berdasarkan hasil evaluasi menggunakan *Confusion Matrix* dan *Classification Report* pada 100 data uji, sistem menunjukkan kinerja yang andal dengan rata-rata akurasi mencapai 97,0% pada ketiga indikator (BB/U, TB/U, dan BB/TB). Kualitas prediksi model juga terbukti seimbang, ditandai dengan nilai rata-rata presisi 0,97, recall 0,96, dan F1-Score 0,97. Angka-angka ini menegaskan bahwa metode Gaussian Naive Bayes efektif dalam meminimalisir kesalahan prediksi sekaligus sensitif dalam mengenali berbagai kelas status gizi, baik kelas mayoritas maupun minoritas.

Untuk penelitian selanjutnya, disarankan agar memperluas cakupan dataset dengan mengambil sampel dari beberapa Puskesmas berbeda di wilayah Nusa Tenggara Timur agar model menjadi lebih general. Selain itu, perlu dilakukan komparasi algoritma dengan metode lain seperti *K-Nearest Neighbor* (KNN) atau *Decision Tree* untuk mencari performa komputasi terbaik, serta penambahan fitur visualisasi grafik pertumbuhan anak (*growth chart*) pada aplikasi web agar pemantauan gizi dapat dilakukan secara *real-time* dan lebih interaktif bagi pengguna.

Daftar Pustaka:

- Andriani, H., Friska, E., Arsyi, M., Sutrisno, A. E., Waits, A., & Rahmawati, N. D. (2023): A Multilevel Analysis of the Triple Burden of Malnutrition in Indonesia: Trends and Determinants from Repeated Cross-sectional Surveys, *BMC Public Health*, Vol. 23:1836, DOI: <https://doi.org/10.1186/s12889-023-16728-y>.
- Wulandari, D. T., Azizah, N., Rahmawati, V. E., & Annuchasari, H. (2025): Community-Based Insights on Maternal Nutrition Knowledge and Its Role in Preventing Stunting and Wasting in Rural Indonesian Children, *Journal of Current Health Sciences*, Vol. 5, No. 3, 2025, pp. 145-154, DOI: <https://doi.org/10.47679/jchs.2025128>.

- Purnomo, A. J., & Rozaq, A. (2022): Classification of Stunting Status in Toddlers Using Naive Bayes Method in the City of Madiun Based on Website, *Techno Nusa Mandiri: Journal of Computing and Information Technology*, Vol. 19, No. 2, September 2022, pp. 69-76, DOI: <https://doi.org/10.33480/techno.v19i2.3337>
- Harliana, & Anggraini, D. (2023): Penerapan Algoritma Naïve Bayes Pada Klasifikasi Status Gizi Balita di Posyandu Desa Kalitengah, *FAHMA – Jurnal Informatika Komputer, Bisnis dan Manajemen*, Vol. 21, No. 2, Mei 2023, pp. 38-45, ISSN: 2715-2944.
- Saputra, Y., Khoirunnisa, N., & Haqq, S. A. (2023): Classification of Health and Nutritional Status of Toddlers Using the Naïve Bayes Classification, *CoreID Journal*, Vol. 1, No. 2, Juli 2023, pp. 49-57, DOI: <https://doi.org/10.60005/coreid.v1i1.8>.
- Hidayat, F. M., Kusriani, & Yaqin, A. (2024): Penerapan Algoritma Naïve Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita, *Dielektrika – Jurnal Ilmiah Kajian Teori dan Aplikasi Teknik Elektro*, Vol. 11, No. 2, Agustus 2024, pp. 107-118.
- Dirta, E. R., Maharani, D., & Syahputra, A. K. (2024): Pemanfaatan Metode Naïve Bayes Untuk Klasifikasi Status Gizi Balita Pada Kelurahan Karang Anyer, *DECODE: Jurnal Pendidikan Teknologi Informasi*, Vol. 4, No. 2, Juli 2024, pp. 392-405, DOI: <http://dx.doi.org/10.51454/decode.v4i2.279>
- Kurniawan, W., Muriyatmoko, D., & Taufiq, S. A. (2025): Klasifikasi Status Gizi Balita Menggunakan Algoritma Naive Bayes Berdasarkan Pengukuran Antropometri, *Prosiding Universitas Darussalam Gontor*, Agustus 2023 – Februari 2024, pp. 1-10.
- Arlita, A., Prihartono, W., & Rohmat, C. L. (2025): Optimalisasi Upaya Pencegahan Stunting Pada Balita di Kota Cirebon Menggunakan Metode Naïve Bayes, *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, Vol. 13, No. 2, April 2025, pp. 9-15, DOI: <http://dx.doi.org/10.23960/jitet.v13i2.5868>.
- Kurniawan, H., Rahim, A., & Siswa, T. A. Y. (2024): Implementasi Algoritma Gaussian Naïve Bayes Dalam Klasifikasi Status Gizi Pada Balita, *Building of Informatics, Technology and Science (BITS)*, Vol. 6, No. 2, September 2024, pp. 627-634, DOI: <https://doi.org/10.47065/bits.v6i2.5493>.
- Supardianto, & Mutawalli, L. (2025): Analisis Klaster Data Kesehatan Bayi dan Anak Menggunakan K-Means Studi Kasus Provinsi Nusa Tenggara Timur, *Journal Computer and Technology*, Vol. 3, No. 1, Juli 2025, hlm. 57-64, DOI: <https://doi.org/10.69916/comtechno.v3i1.343>.
- Karundeng, S. E. S., Picauly, I., & Toy, S. M. (2023): Faktor-Faktor yang Berhubungan dengan Kepatuhan Mengonsumsi Tablet Zat Besi pada Ibu Hamil di Puskesmas Borong Kecamatan Borong Manggarai Timur, *Jurnal PAZIH PERGIZI PANGAN DPD NTT*, Vol. 12, No. 2, Oktober 2023, E-ISSN: 2745-6536.
- Dwiyanti, D. S., & Saptari, M. A. (2023): Perancangan Sistem Informasi Status Gizi Balita (Stunting) di UPTD Puskesmas Muara Satu dan Muara Dua Menggunakan Metode Naïve Bayes Classification dan K-Nearest Neighbor Berbasis Web, *SISFO: Jurnal Ilmiah Sistem Informasi*, Vol. 7, No. 2, 2023, ISSN (P): 2598-599X, ISSN (E): 2599-0330.
- Yoghatama, A., & Pamungkas, E. W. (2025): Development of a Web-Based Platform for Nutrition Education and Early Stunting Detection, *Bit-Tech (Binary Digital–Technology)*, Vol. 8, No. 1, Agustus 2025, DOI: <https://doi.org/10.32877/bt.v8i1.2559>, ISSN: 2622-2728 (online), 2622-271X (print).
- Kementerian Kesehatan Republik Indonesia, (2020) Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak.
- Sommerville, I. (2011): *Software Engineering*, 10th Edition, Addison-Wesley, Boston, USA.
- Murphy, K. P. (2012): *Machine Learning: A Probabilistic Perspective*, The MIT Press, Cambridge, Massachusetts, London, England