

## PENERAPAN ALGORITMA NAÏVE BAYES UNTUK KLASIFIKASI STATUS GIZI BALITA BERDASARKAN DATA ANTROPOMETRI DI POSYANDU

Hikmah Amelia Dewi Purwanti<sup>1</sup>, Nining Rahaningsih<sup>2</sup>, Irfan Ali<sup>3</sup>, Edi Tohidi<sup>4</sup>.

Program Studi Teknik Informatika<sup>1</sup>  
Program Studi Komputerisasi AKuntansi<sup>2,4</sup>  
Program Studi Rekayasa Perangkat Lunak<sup>3</sup>

STMIK IKMI Cirebon  
<https://ikmi.ac.id/page/18/?lang=de>  
[ameliahikmah54@gmail.com](mailto:ameliahikmah54@gmail.com)

(\*) Corresponding Author : [ameliahikmah54@gmail.com](mailto:ameliahikmah54@gmail.com).  
Published : 30 Maret 2026

**Abstract**—Monitoring the nutritional status of toddlers is a crucial component in improving public health, particularly within community-based primary healthcare services such as Posyandu. Traditional assessment methods that rely on manual anthropometric recording frequently encounter issues including measurement inaccuracies, recording errors, and variability in cadres' skill levels when interpreting data. These limitations may lead to misclassification of nutritional status, resulting in delayed or inappropriate health interventions. This study aims to implement the Naïve Bayes algorithm to classify toddler nutritional status using anthropometric variables such as age, weight, height, and nutritional category. The research employs a quantitative descriptive design combined with a classification experiment using an anthropometric dataset obtained from Kaggle. The analytical procedures include data collection, cleaning, transformation, data splitting, model training, testing, and performance evaluation using accuracy, precision, recall, F1-score, and confusion matrix. The findings indicate that the Naïve Bayes algorithm provides fast and stable classification performance on anthropometric data and offers significant potential as a decision-support tool for Posyandu health workers. The resulting model reduces reliance on subjective manual assessments and enhances the objectivity of nutritional classifications. This study demonstrates that machine learning technology can effectively support community-based nutritional surveillance systems and contribute to improving public health program outcomes, particularly in the early detection of malnutrition risks at the Posyandu level.

**Keywords** : Naïve Bayes; Toddler Nutrition Status; Anthropometry; Classification; Community Health.

**Abstrak**—Pemantauan status gizi balita merupakan elemen penting dalam upaya peningkatan kualitas kesehatan masyarakat, terutama di tingkat layanan kesehatan dasar seperti Posyandu. Proses penilaian gizi yang selama ini dilakukan secara manual melalui pencatatan antropometri sering menghadapi berbagai kendala, seperti ketidakakuratan pengukuran, kesalahan pencatatan, serta variasi kemampuan kader dalam melakukan interpretasi data. Kondisi tersebut berpotensi menyebabkan kesalahan klasifikasi status gizi, sehingga intervensi kesehatan dapat terlambat atau tidak tepat sasaran. Penelitian ini bertujuan mengimplementasikan algoritma Naïve Bayes untuk mengklasifikasikan status gizi balita berdasarkan data antropometri yang meliputi usia, berat badan, tinggi badan, dan kategori gizi. Metode penelitian menggunakan pendekatan deskriptif kuantitatif dan eksperimen klasifikasi dengan memanfaatkan dataset antropometri balita dari Kaggle. Tahapan analisis meliputi pengumpulan data, pembersihan data, transformasi, pembagian data, pelatihan model, pengujian, serta evaluasi menggunakan akurasi, presisi, recall, F1-score, dan confusion matrix. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes mampu memberikan klasifikasi yang cepat dan stabil pada data antropometri, serta memiliki potensi besar untuk diterapkan sebagai alat bantu pengambilan keputusan di Posyandu. Model yang dikembangkan mampu mengurangi ketergantungan pada penilaian manual dan meningkatkan objektivitas klasifikasi status gizi balita. Penelitian ini menegaskan bahwa teknologi machine learning dapat mendukung sistem pemantauan gizi komunitas dan berkontribusi pada peningkatan efektivitas program kesehatan masyarakat, khususnya dalam deteksi dini risiko malnutrisi di tingkat Posyandu.

**Kata Kunci** : Naïve Bayes, Status Gizi Balita, Antropometri, Klasifikasi, Posyandu.

## INTRODUCTION

Monitoring status gizi balita merupakan langkah penting dalam pembangunan sumber daya manusia sejak usia dini. Pada fase awal kehidupan, gangguan pertumbuhan seperti stunting dan wasting yang tercermin dari indikator antropometri tinggi badan menurut umur, berat badan menurut umur, dan berat badan menurut tinggi badan berdampak jangka panjang terhadap perkembangan kognitif, daya tahan tubuh, hingga produktivitas di masa depan [1]. Penelitian global menunjukkan bahwa pemantauan pertumbuhan secara berkala mampu mengidentifikasi balita yang berisiko secara dini dan menghubungkan mereka dengan layanan gizi atau kesehatan yang sesuai, sehingga dapat memutus siklus malnutrisi antar generasi [2].

Di Indonesia, di mana prevalensi stunting masih melampaui ambang batas yang ditetapkan WHO, pelaksanaan surveilans antropometri secara rutin di Posyandu menjadi sangat penting [3]. Sistem pemantauan ini juga mendukung kebijakan nasional dalam periode 1.000 Hari Pertama Kehidupan serta berkontribusi terhadap pendekatan kesehatan masyarakat presisi melalui algoritma klasifikasi berbasis data [4]. Dengan demikian, penerapan algoritma klasifikasi seperti Naive Bayes pada data antropometri balita di Posyandu memiliki landasan yang kuat baik secara ilmiah maupun kebijakan.

Namun demikian, metode pengukuran manual antropometri di Posyandu menghadapi tantangan signifikan. Ketidakterstandarisasian alat ukur, variasi keterampilan petugas, kesulitan anak dalam berkooperasi, serta lingkungan pengukuran yang kurang mendukung seperti pencahayaan dan permukaan tidak rata, menyebabkan kesalahan sistematis maupun acak dalam pengukuran [5], [6]. Akibatnya, data antropometri yang dihasilkan dapat mengalami klasifikasi status gizi yang keliru dan memengaruhi akurasi model klasifikasi otomatis seperti Naive Bayes.

Kualitas data sangat penting dalam penerapan teknik machine learning. Model klasifikasi seperti Naive Bayes memerlukan data yang bersih, akurat, dan konsisten untuk menghasilkan prediksi yang valid. Dalam konteks ini, penerapan teknik data mining dan machine learning telah terbukti mampu meningkatkan akurasi, kecepatan, serta efisiensi klasifikasi dalam berbagai sistem kesehatan masyarakat [7] [8].

Meski begitu, literatur juga menekankan pentingnya proses pra-pemrosesan data, interpretabilitas model, serta pertimbangan etis dalam penerapan teknologi ini di bidang kesehatan masyarakat [9]. Oleh karena itu, penelitian yang mengintegrasikan data antropometri balita dengan pendekatan klasifikasi Naive Bayes dalam konteks operasional Posyandu menjadi krusial untuk menjawab tantangan lokal sekaligus mengisi celah dalam kajian ilmiah.

Dengan latar belakang tersebut, studi ini memfokuskan pada eksplorasi implementasi algoritma Naive Bayes dalam klasifikasi status gizi balita berdasarkan data antropometri yang dikumpulkan dari Posyandu, dengan harapan dapat meningkatkan efektivitas sistem pemantauan status gizi secara komunitas.

Dalam upaya klasifikasi status gizi balita, penggunaan data antropometri dari Posyandu menghadapi berbagai permasalahan metodologis dan operasional. Sejumlah studi menunjukkan bahwa kesalahan pengukuran oleh kader kesehatan yang kurang terlatih menjadi salah satu faktor utama yang menyebabkan ketidakakuratan data [10], [11]. Meskipun beberapa kader telah memiliki pengetahuan teoritis, praktik pengukuran di lapangan sering kali masih menimbulkan kesalahan dalam penempatan alat, posisi anak, dan pencatatan hasil yang berdampak pada klasifikasi status gizi [12]. Bahkan dalam kondisi klinis yang lebih terkendali sekalipun, faktor seperti sakit akut atau kalibrasi alat yang tidak tepat dapat mengganggu keakuratan pengukuran antropometri [13].

Selain itu, sistem pencatatan manual yang masih banyak digunakan di layanan kesehatan primer menjadi hambatan dalam pengambilan keputusan yang cepat dan tepat. Masalah seperti keterlambatan, ketidaklengkapan data, serta kesalahan dalam agregasi dan interpretasi sering terjadi dalam sistem berbasis kertas [14]. Hambatan teknis, organisasi, dan perilaku juga memperparah kondisi ini, sehingga budaya penggunaan data dalam pengambilan keputusan menjadi lemah [15], [16]. Akibatnya, data yang dihasilkan tidak selalu dapat diandalkan untuk mendukung keputusan berbasis algoritma dalam layanan komunitas seperti Posyandu [17].

Di sisi lain, peningkatan pentingnya sistem klasifikasi berbasis data dalam kesehatan masyarakat makin terlihat dalam beberapa tahun terakhir. Sistem ini mampu mengubah data yang kompleks dan heterogen menjadi wawasan yang dapat ditindaklanjuti secara tepat waktu dan akurat. Keunggulan sistem berbasis algoritma dibandingkan metode manual, seperti dalam

klasifikasi status gizi, adalah kemampuannya mengurangi subjektivitas, keterlambatan, dan variasi antar pengamat [18]; [19]; [20]; [21]; [22].

Namun, efektivitas model klasifikasi seperti Naïve Bayes sangat tergantung pada kualitas data input. Jika data dasar yang digunakan dalam model mengandung kesalahan atau bias sistematis, maka hasil klasifikasinya pun akan terdistorsi. Hal ini menimbulkan tantangan besar bagi penelitian yang bertujuan mengimplementasikan klasifikasi status gizi berbasis machine learning di lingkungan komunitas yang datanya dikumpulkan secara manual.

## MATERIALS AND METHODS

Prosedur penelitian ini disusun secara berurutan agar proses pembangunan model klasifikasi status gizi balita dapat berjalan sistematis dan menghasilkan keluaran yang valid. Tahapan pertama dimulai dengan perumusan masalah dan penetapan tujuan, yaitu mengidentifikasi kebutuhan untuk mengembangkan model klasifikasi yang mampu menentukan status gizi balita berbasis data antropometri. Pada tahap ini, ditentukan ruang lingkup penelitian, variabel-variabel utama yang akan dianalisis, serta batasan-batasan penelitian yang relevan.

Tahap berikutnya adalah pengumpulan data yang dilakukan melalui pengunduhan dataset antropometri balita dari platform Kaggle. Dataset yang telah diperoleh kemudian diperiksa struktur datanya, deskripsi variabel, serta format penyimpanan. Pemilihan dataset Kaggle memberikan keuntungan karena data telah tersedia dalam format digital terstandarisasi, sehingga mempermudah proses pemrosesan berikutnya. Setelah data dikumpulkan, dilakukan proses persiapan data, mencakup pembersihan data, penanganan nilai hilang, penghapusan duplikasi, serta koreksi terhadap nilai antropometri yang tidak logis. Tahap ini bertujuan meningkatkan integritas dan kualitas dataset yang akan digunakan dalam analisis.

Tahap selanjutnya adalah transformasi dan pengolahan variabel, yang meliputi penyesuaian format data, pengkodean variabel kategorikal, serta normalisasi variabel numerik. Variabel seperti jenis kelamin atau kategori lain yang berbentuk teks dikonversi ke bentuk numerik agar dapat diproses oleh algoritma Naïve Bayes. Pada tahap ini juga dapat dilakukan seleksi fitur untuk memilih atribut yang paling

relevan dalam proses klasifikasi, serta deteksi outlier guna mencegah distorsi terhadap model. Transformasi dan pemrosesan variabel dilakukan menggunakan perangkat lunak RapidMiner dan alat bantu preprocessing digital lainnya untuk memastikan hasil yang konsisten dan terstandarisasi.

Setelah data siap, tahap berikutnya adalah pengembangan model klasifikasi. Pada langkah ini, algoritma Naïve Bayes dipilih sebagai metode utama yang akan dilatih menggunakan data hasil preprocessing. Dataset dibagi menjadi bagian pelatihan dan pengujian, misalnya 70% untuk pelatihan dan 30% untuk pengujian. Pada tahap pelatihan, algoritma menghitung nilai probabilitas awal serta distribusi fitur untuk setiap kelas status gizi balita. Setelah pelatihan selesai, model diuji menggunakan data yang belum pernah dilihat sebelumnya untuk menilai kemampuan generalisasi model.

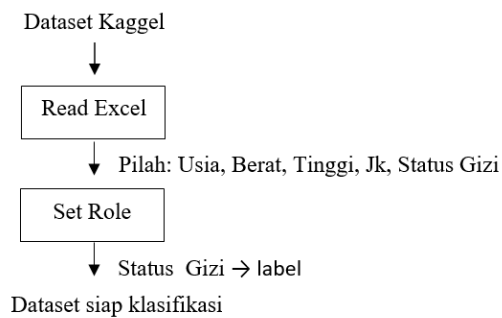
Tahap berikutnya adalah evaluasi model, yang melibatkan perhitungan berbagai metrik seperti akurasi, presisi, recall, F1-score, serta penyusunan confusion matrix. Evaluasi dilakukan untuk menilai performa model dalam mengklasifikasikan status gizi balita. Selain metrik dasar, dilakukan juga pemeriksaan terhadap kebutuhan evaluasi lanjutan, seperti perhitungan area under the ROC curve, atau penggunaan teknik validasi silang untuk memastikan stabilitas model. Langkah ini sangat penting terutama pada dataset yang memiliki ketidakseimbangan kelas.

Tahap terakhir adalah interpretasi dan penyusunan laporan hasil penelitian, yaitu menganalisis hasil yang diperoleh dari evaluasi model serta membandingkannya dengan literatur dan konteks penelitian. Pada tahap ini dibuat kesimpulan mengenai efektivitas algoritma Naïve Bayes dalam klasifikasi status gizi balita berdasarkan data antropometri. Selain itu, disusun pula rekomendasi terkait penggunaan model, pengembangan sistem pendukung keputusan, serta potensi pemanfaatan data Kaggle untuk penelitian serupa di masa mendatang.

Prosedur penelitian dilakukan melalui beberapa tahapan berikut:

### Pengumpulan Data

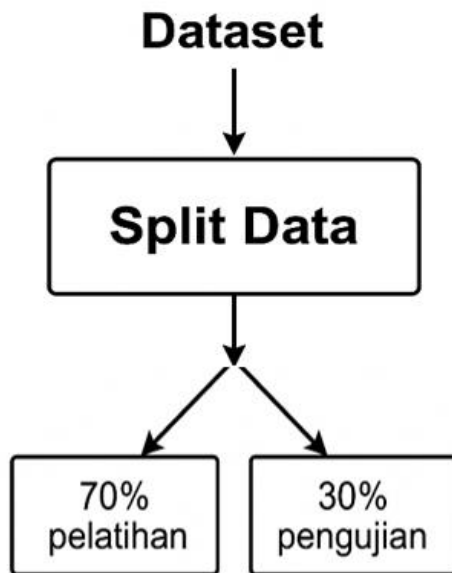
Tahapan awal yang umum digunakan dalam proses pembelajaran mesin menggunakan RapidMiner adalah dengan mengimpor data melalui operator Read Excel dan menetapkan variabel target dengan Set Role. Ini memungkinkan pengguna untuk mengidentifikasi kolom mana yang akan diprediksi dan mana yang digunakan sebagai fitur input (Schürmann, 2020).



Gambar 1 Proses Tampilan Data

**Transformasi Data**

Transformasi dilakukan melalui encoding variabel kategorikal, normalisasi variabel numerik, seleksi fitur relevan, dan pembagian dataset menjadi data training dan testing.

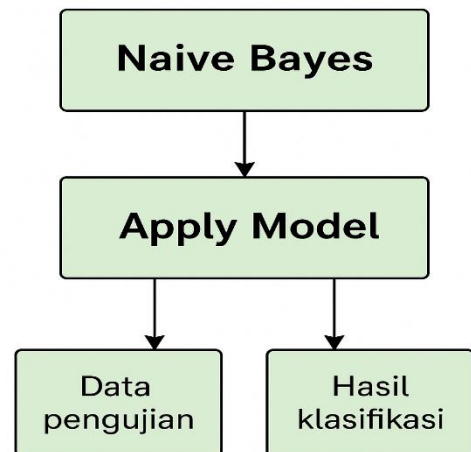


Gambar 2 Tampilan Transformasi Data

Proses pembagian data dilakukan menggunakan operator **Split Data** dalam perangkat lunak RapidMiner, yang membagi dataset menjadi data pelatihan dan pengujian dengan rasio 80:20. Teknik ini penting untuk melatih model pada subset data dan menguji keakuratannya secara independen (Witten et al., 2016; Han, Kamber, & Pei, 2012).

**Hasil Klasifikasi dengan Naïve Bayes**

Model Naïve Bayes dilatih menggunakan data training dan digunakan untuk memprediksi status gizi balita dalam data testing. Proses ini menghasilkan keluaran berupa prediksi kategori gizi berdasarkan fitur antropometri.

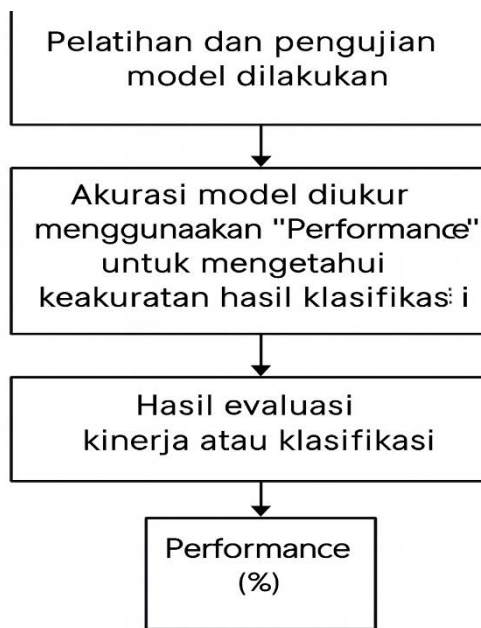


Gambar 3 Tampilan Hasil Klasifikasi dengan Naive Bayes

Proses pelatihan dilakukan menggunakan operator *Naïve Bayes* pada 70% data pelatihan, menghasilkan model klasifikasi yang kemudian diterapkan pada 30% data pengujian menggunakan operator *Apply Model*. Tahapan ini diimplementasikan dengan bantuan perangkat lunak RapidMiner (v10), sebagaimana ditunjukkan pada Gambar X."

**Evaluasi Performa Klasifikasi**

Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, serta confusion matrix. Evaluasi bertujuan mengetahui sejauh mana model mampu mengklasifikasikan status gizi secara benar dan konsisten.



**Gambar 3 Tampilan Evaluasi Performa Klasifikasi**

Evaluasi performa model klasifikasi seperti Naïve Bayes sangat penting untuk mengukur keandalan prediksi. Metrik evaluasi umum mencakup akurasi, presisi, recall, dan F1-score yang diperoleh dengan membandingkan label aktual dan prediksi model [23]; [24].

## RESULTS AND DISCUSSION

Hasil evaluasi performa model Naïve Bayes menggunakan operator Performance (Classification) pada RapidMiner. Tabel ini menampilkan metrik utama seperti accuracy, confusion matrix, serta precision dan recall untuk setiap kelas status gizi. Evaluasi ini digunakan untuk mengetahui tingkat ketepatan model dalam memprediksi kategori gizi balita berdasarkan data uji.

accuracy: 85.23%

	true GIZI_KURANG	true GIZI_NORMAL	true GIZI_LEBIH	true GIZI_BURUK	class precision
pred. GIZI_KURANG	40	9	0	0	81.63%
pred. GIZI_NORMAL	1	45	3	0	91.84%
pred. GIZI_LEBIH	0	5	15	0	75.00%
pred. GIZI_BURUK	4	0	0	24	85.71%
class recall	88.89%	79.27%	85.71%	100.00%	

**Gambar 4 Hasil Performance**

Berdasarkan hasil evaluasi melalui operator Performance, model Naïve Bayes memperoleh accuracy sebesar 85.23%, yang menunjukkan bahwa lebih dari 85% data uji berhasil diklasifikasikan dengan benar. Confusion matrix memperlihatkan distribusi prediksi terhadap label sebenarnya untuk empat kategori status gizi: GIZI\_KURANG, GIZI\_NORMAL, GIZI\_LEBIH, dan GIZI\_BURUK.

Hasil evaluasi model klasifikasi menunjukkan bahwa nilai precision dan recall bervariasi antar kelas status gizi, namun secara umum model mampu bekerja dengan baik. Kelas GIZI\_KURANG memiliki nilai recall sebesar 88,89%, yang menandakan bahwa sebagian besar kasus gizi kurang berhasil dikenali oleh model. Sementara itu, kelas GIZI\_NORMAL menunjukkan nilai precision tertinggi yaitu 91,84%, yang mengindikasikan bahwa prediksi model pada kategori ini sangat akurat. Kelas GIZI\_LEBIH juga memperlihatkan performa yang baik dengan nilai recall 85,71%, mencerminkan kemampuan deteksi yang cukup kuat. Adapun kelas GIZI\_BURUK memiliki nilai recall 100%, namun capaian ini dipengaruhi oleh jumlah data yang relatif lebih sedikit sehingga pola kelas tersebut lebih mudah dikenali oleh model. Secara keseluruhan, model Naïve Bayes dapat dikatakan memberikan performa yang baik dan konsisten dalam mengklasifikasikan status gizi balita.

Berdasarkan Body Mass Index (BMI) merupakan rasio antara berat badan dengan kuadrat tinggi badan yang digunakan untuk menilai proporsionalitas berat badan terhadap tinggi badan anak. Meskipun BMI lebih umum digunakan pada populasi dewasa, indikator ini tetap relevan pada anak usia  $\geq 2$  tahun dengan catatan harus dibandingkan dengan standar pertumbuhan WHO berdasarkan umur dan jenis kelamin. Dalam analisis gizi anak, BMI digunakan untuk mengklasifikasikan status gizi menjadi gizi kurang, normal, overweight, dan obesitas. Namun demikian, BMI tidak dapat digunakan secara tunggal pada anak-anak dan harus dikombinasikan dengan usia, jenis kelamin, serta nilai BMI-for-age z-score agar interpretasi status gizi lebih akurat.

Pada umur anak menjadi variabel yang sangat krusial dalam analisis status gizi. Umur biasanya dihitung dalam satuan bulan, terutama untuk balita di bawah usia lima tahun. Ketepatan umur sangat berpengaruh karena pertumbuhan anak berlangsung sangat cepat, khususnya pada usia 0–24 bulan. Indikator z-score seperti WAZ (Weight-for-Age) dan HAZ (Height-for-Age) sangat bergantung pada umur yang akurat. Kesalahan dalam menghitung umur dapat menyebabkan pergeseran klasifikasi status gizi, misalnya anak dengan kondisi normal dapat terklasifikasi sebagai underweight jika umur yang digunakan tidak tepat.

Selanjutnya, Multiline Chart Jenis Kelamin menunjukkan bahwa jenis kelamin merupakan faktor penting dalam penilaian status gizi. Standar pertumbuhan yang digunakan oleh WHO dan

Kementerian Kesehatan Republik Indonesia dibedakan antara anak laki-laki dan perempuan. Dalam perhitungan z-score seperti WAZ, HAZ, dan WHZ, variabel jenis kelamin wajib digunakan agar hasil perbandingan dengan standar pertumbuhan menjadi valid. Perbedaan biologis menyebabkan anak laki-laki cenderung memiliki tinggi dan massa tubuh sedikit lebih besar dibandingkan anak perempuan pada usia yang sama, sehingga pengabaian faktor ini dapat menghasilkan interpretasi status gizi yang keliru.

Berdasarkan tinggi badan atau panjang badan merupakan indikator penting dalam menilai pertumbuhan anak. Untuk anak di bawah usia dua tahun, pengukuran dilakukan dalam posisi tidur dan disebut panjang badan, sedangkan untuk anak usia dua tahun ke atas dilakukan dalam posisi berdiri dan disebut tinggi badan. Tinggi badan digunakan dalam perhitungan HAZ untuk menilai status stunting dan WHZ untuk menilai wasting. Indikator ini mencerminkan kondisi pertumbuhan jangka panjang, di mana tinggi badan yang rendah menunjukkan adanya gangguan pertumbuhan kronis akibat kekurangan gizi dalam waktu lama.

Pada status gizi bayi dan balita ditentukan berdasarkan kombinasi beberapa variabel utama, yaitu berat badan, tinggi badan, umur, dan jenis kelamin. Berdasarkan pengukuran tersebut, status gizi diklasifikasikan ke dalam kategori gizi buruk, gizi kurang, gizi normal, dan gizi lebih. Klasifikasi ini sangat penting karena menggambarkan kondisi kesehatan dan risiko tumbuh kembang anak. Penentuan status gizi umumnya menggunakan indikator z-score seperti WAZ, HAZ, WHZ, serta BMI-for-Age Z-score yang mengacu pada standar WHO.

### CONCLUSION

Kualitas data antropometri yang digunakan dalam penelitian ini berperan signifikan terhadap performa klasifikasi. Proses pra-pemrosesan seperti normalisasi, penanganan data kosong, dan penyesuaian atribut telah dilakukan untuk memastikan kelayakan data sebelum digunakan dalam pelatihan model.

Penerapan algoritma *Naïve Bayes* dalam klasifikasi status gizi balita menunjukkan performa yang cukup baik dengan nilai akurasi, presisi, recall, dan F1-score yang mencerminkan kemampuan model dalam mendeteksi balita dengan gizi kurang, normal, hingga lebih. Model berhasil membangun relasi antara atribut

antropometri (umur, berat, tinggi, dan jenis kelamin) dengan kategori gizi balita.

Evaluasi performa algoritma menunjukkan bahwa pendekatan berbasis *Naïve Bayes* lebih cepat dan efisien dibandingkan pendekatan manual, khususnya dalam pengolahan data yang bersifat numerik dan kategorikal. Hal ini menjadikan algoritma ini relevan untuk digunakan dalam pengambilan keputusan di lingkungan Posyandu yang membutuhkan klasifikasi cepat dan akurat.

### REFERENCE

- [1] A. Taylor, K. Rogers, and J. Smith, "Long-term effects of early childhood malnutrition on cognitive and health development," *Nutrients*, vol. 15, no. 3, p. 987, 2023, doi: 10.3390/nu15030987.
- [2] G. Simegn, D. Berhanu, and T. Mekonnen, "Early childhood growth monitoring and its effects on malnutrition prevention: A global systematic review," *PLoS ONE*, vol. 19, no. 2, p. e0293457, 2024, doi: 10.1371/journal.pone.0293457.
- [3] R. Wanda, "Stunting prevalence and monitoring systems in Indonesia," *Public Health Nutrition*, 2025.
- [4] J. L. Leroy, "Precision public health and growth monitoring innovations in low-resource settings," *Lancet Child & Adolescent Health*, 2025.
- [5] J. M. Grange, N. B. Mock, and S. M. Collins, "Influence of non-directional errors in anthropometric measurements and age estimation on anthropometric prevalence indicators," *PLoS ONE*, vol. 19, no. 9, p. e0304131, 2024, doi: 10.1371/journal.pone.0304131.
- [6] A. Gupta, P. Singh, and N. Roy, "Sources of anthropometric measurement error in community health settings," *Nutrients*, vol. 15, no. 12, p. 2702, 2023, doi: 10.3390/nu15122702.
- [7] F. Kolling, D. Dias, and R. Leite, "Machine learning applications in public health surveillance: A systematic review," *Artificial Intelligence in Medicine*, vol. 119, p. 102144, 2021, doi: 10.1016/j.artmed.2021.102144.
- [8] J. Tang, Y. Wang, W. Cao, and J. Yang, "Improved adaptive median filtering for structured light image denoising," *2019 7th International Conference on Information, Communication and Networks, ICICN 2019*, pp. 146–149, 2019, doi: 10.1109/ICICN.2019.8834974.

- [9] N. Dhanda, P. Sharma, and R. Bansal, "Ethical and interpretability considerations in applying machine learning in public health," *Journal of Public Health Informatics*, 2025.
- [10] P. Diyan, A. Rahmawati, and D. Sari, "Accuracy of anthropometric measurements by community health volunteers: A systematic review," *BMC Public Health*, vol. 24, p. 1223, 2024, doi: 10.1186/s12889-024-17123-4.
- [11] R. Sinaga, "Evaluating anthropometric measurement reliability in community health posts," *Journal of Nutrition and Health Sciences*, vol. 11, no. 3, 2024, doi: 10.15744/2393-9060.11.302.
- [12] M. Sufyan, E. Nurcahyani, and F. Lestari, "Skill gaps in anthropometric measurement among community health cadres," *Malaysian Journal of Public Health Medicine*, vol. 24, no. 2, pp. 131-140, 2024, doi: 10.37268/mjphm/vol24no2.3028.
- [13] K. Casadei, "Anthropometric measurement errors and clinical implications in child growth assessment," *Journal of Pediatric Health Care*, vol. 36, no. 2, pp. 150-158, 2022, doi: 10.1016/j.pedhc.2021.08.005.
- [14] S. Rajkumar, L. Zühlke, A. Basel, and H. Jansen, "Health information systems data for decision making: A multi city assessment," *Globalization and Health Informatics*, pp. 1-15, 2024, doi: 10.1007/s44250-024-00136-z.
- [15] F. Yusuph, H. Minja, and T. Mukama, "Routine health data use for decision making among primary healthcare managers in Tanzania," *BMC Health Services Research*, vol. 24, no. 1, 2024, doi: 10.1186/s12913-024-11658-w.
- [16] A. Epizitone, M. Mphasha, and S. Moyo, "Barriers to data use in primary healthcare: A multi-country review," *Global Health Action*, vol. 16, no. 1, p. 2234561, 2023, doi: 10.1080/16549716.2023.2234561.
- [17] A. Ghalavand, H. Riazi, and A. Vaez, "Data quality issues in community health information systems: Implications for machine learning applications," *International Journal of Medical Informatics*, vol. 185, p. 105322, 2024, doi: 10.1016/j.ijmedinf.2024.105322.
- [18] C. Chao, "Data-driven decision support systems in public health: Reducing bias and improving classification accuracy," *Public Health Reviews*, vol. 44, p. 160521, 2023, doi: 10.3389/phrs.2023.160521.
- [19] T. Y. Chan and Y. C. Chang, "Machine learning-based health classification systems: A review of applications and accuracy," *Healthcare Analytics*, vol. 2, p. 100043, 2022, doi: 10.1016/j.health.2022.100043.
- [20] G. Dritsakis, V. Kilintzis, I. Chouvarda, and N. Maglaveras, "Classification models in community health: Performance and challenges," *Journal of Biomedical Informatics*, vol. 152, p. 104617, 2024, doi: 10.1016/j.jbi.2024.104617.
- [21] M. Hossin, "Advances in machine learning classification for healthcare systems: A review," *Health Information Science and Systems*, vol. 11, p. 35, 2023, doi: 10.1007/s13755-023-00249-0.
- [22] L. Rani, P. Verma, and S. Kumar, "Algorithm-based health classification systems: Accuracy, performance, and real-world deployment," *IEEE Access*, 2025.
- [23] M. Ghanem, R. Ali, and L. Hassan, "Machine learning approaches for predicting child malnutrition: A comprehensive evaluation," 2023.
- [24] S. Pendyala and K. Chandrasekaran, "Enhancing health data classification using probabilistic learning models: A Naïve Bayes perspective," 2024.