

Implementasi Algoritma Random Forest untuk Klasifikasi Status Gizi Balita Berbasis Web

Implementation of Random Forest Algorithm for Web-Based Toddler Nutritional Status Classification

Muhammad Raddhyan Rashif^{*1}, Deanna Durbin Hutagalung²
Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang
¹rashif2410@gmail.com, ²deanna.upn91@gmail.com

Received: July 21, 2026 | Revised: July 28, 2026 | Accepted: August 28, 2026 | Published: September 13, 2026



COSIE is licensed a Creative Commons Attribution-Share Alike 4.0 International License

Abstrak

Stunting adalah masalah gizi kronis yang masih menjadi tantangan utama kesehatan masyarakat Indonesia, dengan prevalensi nasional 21,6% menurut Survei Status Gizi Indonesia (SSGI) tahun 2022 [1], sementara klasifikasi status gizi balita di Posyandu Makmur 1 masih dilakukan secara manual melalui tabel z-score WHO yang rentan terhadap human error. Penelitian ini mengintegrasikan algoritma Random Forest dengan optimasi hyperparameter melalui Grid Search Cross-Validation 5-fold untuk mengklasifikasikan status gizi balita ke dalam empat kategori (Gizi Buruk, Gizi Kurang, Gizi Baik, dan Gizi Lebih) berdasarkan data antropometri, serta mengimplementasikan model tersebut ke dalam aplikasi web berbasis Flask melalui metode Rapid Application Development (RAD). Model dilatih pada 263 data balita usia 0-59 bulan (rasio latih-uji 80:20) dan diuji pada 53 data uji, menghasilkan akurasi 83,02% dengan kombinasi parameter $n_estimators=200$ dan $max_depth=15$. Namun demikian, evaluasi tambahan berbasis macro F1-Score (59,90%) dan balanced accuracy (61,64%) mengungkap penurunan performa yang signifikan pada kelas minoritas, terutama Gizi Kurang, akibat ketidakseimbangan distribusi kelas dalam dataset. Analisis feature importance menunjukkan Tinggi Badan (35,5%) dan Usia (30,1%) sebagai fitur paling berpengaruh, sejalan dengan formula Z-Score TB/U yang menjadi dasar pelabelan data. Sistem yang dikembangkan lulus pengujian White Box dan Black Box, menunjukkan kelayakan fungsionalnya sebagai alat bantu deteksi dini status gizi di layanan kesehatan primer, sekaligus menegaskan perlunya strategi penanganan class imbalance pada penelitian klasifikasi status gizi berbasis machine learning selanjutnya.

Kata kunci: Random Forest, Klasifikasi Status Gizi, Balita, Stunting, Aplikasi Web

Abstract

Stunting remains a persistent public health challenge in Indonesia, affecting 21.6% of children under five according to the 2022 Indonesian Nutrition Status Survey (SSGI), while nutritional status classification at Posyandu Makmur 1 is still performed manually using the WHO z-score table, a process prone to human error. This study integrates a Random Forest algorithm, optimized through 5-fold Grid Search Cross-Validation, to classify toddler nutritional status into four categories (severe malnutrition, undernutrition, normal nutrition, and overnutrition) based on anthropometric data, and implements the resulting model into a Flask-based web application developed using the Rapid Application Development (RAD) method. The model was trained on 263 records of toddlers aged 0-59 months (80:20 train-test split) and evaluated on 53 test samples, achieving 83.02% accuracy with $n_estimators=200$ and $max_depth=15$. However, additional evaluation using macro F1-score (59.90%) and balanced accuracy (61.64%) reveals a substantial performance drop for minority classes, particularly undernutrition, due to class imbalance in the dataset. Feature importance analysis identified height (35.5%) and age (30.1%) as the most influential predictors, consistent with the height-for-age Z-score formula used for data labeling. The resulting system passed both White Box and Black Box testing, confirming its functional readiness as an early screening

tool at primary health services, and underscoring the need for class imbalance mitigation strategies in future machine learning-based nutritional classification research.

Keywords: Random Forest, Nutritional Status Classification, Toddler, Stunting, Web Application

1. PENDAHULUAN

Stunting tetap menjadi salah satu isu gizi yang berkelanjutan dan sangat penting dalam sektor kesehatan di Indonesia. Kondisi ini terjadi karena anak mendapatkan asupan gizi yang kurang selama waktu yang lama, sehingga tinggi badannya tidak mencapai standar yang ditetapkan oleh WHO. Selain menghambat pertumbuhan fisik, stunting juga berisiko menimbulkan gangguan kemampuan berpikir, kesulitan belajar, dan kerentanan terhadap penyakit infeksi, yang pada akhirnya bisa berdampak pada mutu tenaga kerja dalam jangka panjang. Menurut Survei Status Gizi Indonesia (SSGI) pada tahun 2022, prevalensi stunting di Indonesia masih berada di angka 21,6% dan tetap menjadi masalah kesehatan masyarakat yang memerlukan perhatian serius, khususnya pada layanan kesehatan primer [1]. Proses deteksi dan klasifikasi status gizi balita di Posyandu Makmur 1 saat ini cara pemeriksaan masih dilakukan dengan cara konvensional, yaitu dengan mengukur tinggi dan berat badan anak balita kemudian membandingkannya dengan tabel z-score standar WHO secara manual. Metode konvensional yang bergantung pada perhitungan manual dan pembacaan tabel memiliki kelemahan, antara lain rentan terhadap human error dalam pengukuran, pencatatan, dan interpretasi hasil, serta kurang efektif dalam mengolah data berjumlah besar sehingga berpotensi menghasilkan klasifikasi yang kurang akurat dan konsisten [2]. Penerapan machine learning menjadi solusi yang efektif untuk melakukan klasifikasi status gizi secara lebih objektif dan akurat. Beberapa algoritma yang umum digunakan pada data kesehatan antara lain Naive Bayes, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), dan Random Forest. Naive Bayes unggul dalam kesederhanaan komputasi namun memiliki keterbatasan pada asumsi tidak adanya ketergantungan antara fitur yang sering kali tidak dipenuhi dalam data kesehatan; Decision Tree mampu memodelkan hubungan non-linear tetapi cenderung overfitting pada data kompleks; sedangkan SVM memiliki performa baik pada data berdimensi tinggi namun membutuhkan proses tuning parameter dan waktu komputasi yang lebih besar [3]. Random Forest, sebagai teknik ensemble learning yang mengkombinasikan berbagai pohon keputusan, terbukti memiliki tingkat ketepatan yang lebih baik dibandingkan teknik lainnya dibandingkan metode lain pada kasus prediksi stunting karena mampu menangani data dengan banyak variabel dan pola kompleks serta mengurangi risiko overfitting [4].

Beberapa penelitian terdahulu telah menerapkan algoritma berbasis penerapan machine learning untuk mengidentifikasi status gizi atau stunting pada anak-anak kecil. Mundirin dkk, membuat sistem klasifikasi gizi anak balita dengan menggunakan data ukuran tubuh dan metode Random Forest, namun belum diintegrasikan pada aplikasi berbasis web yang mendukung penggunaan langsung oleh tenaga kesehatan [5]. Fatmawati dkk. mengimplementasikan algoritma Random Forest pada website klasifikasi status gizi balita, akan tetapi tidak membahas proses hyperparameter tuning maupun analisis feature importance secara mendalam [6]. Penelitian lain berfokus pada penanganan ketidakseimbangan kelas menggunakan teknik SMOTE yang dikombinasikan dengan Random Forest [7], perbandingan performa Random Forest terhadap KNN, Decision Tree, dan XGBoost pada deteksi stunting [8], serta pengujian parameter Random Forest untuk prediksi stunting [9]. Studi perbandingan yang lebih luas antara XGBoost, Random Forest, SVM, dan k-NN dengan SMOTE juga menunjukkan bahwa Random Forest tetap menjadi salah satu algoritma dengan keseimbangan optimal antara akurasi, recall, dan F1-score meskipun bukan selalu yang tertinggi [10].

Dibandingkan dengan penelitian-penelitian tersebut, penelitian ini menawarkan perbedaan konseptual dan praktis yang lebih eksplisit. Berbeda dari Mundirin dkk. yang membangun model klasifikasi berbasis Random Forest tanpa integrasi ke aplikasi berbasis web [5], dan Fatmawati dkk. yang mengimplementasikan Random Forest pada website klasifikasi status gizi namun tanpa proses hyperparameter tuning maupun analisis feature importance secara mendalam [6], penelitian ini menggabungkan ketiga aspek tersebut secara utuh, yaitu optimasi hyperparameter melalui Grid Search Cross-Validation, analisis feature importance berbasis Gini Importance untuk interpretabilitas model, serta

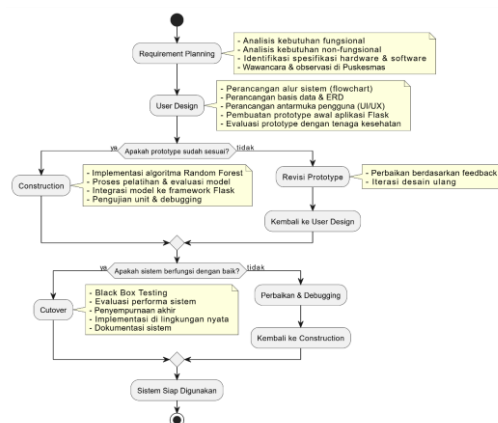
implementasi penuh ke dalam aplikasi web siap pakai bagi tenaga kesehatan di layanan primer. Penelitian yang berfokus pada penanganan class imbalance melalui SMOTE maupun perbandingan algoritma pada umumnya dilakukan pada level eksperimen model tanpa disertai implementasi sistem, sehingga kontribusi praktisnya terhadap tenaga kesehatan di lapangan masih terbatas [7], [8], [9], [10]. Dengan demikian, kebaruan penelitian ini terletak pada integrasi menyeluruh antara optimasi model, interpretabilitas melalui feature importance, dan penerapan langsung pada studi kasus nyata di Posyandu Makmur 1, yang belum ditemukan secara bersamaan pada penelitian-penelitian terdahulu.

Berdasarkan tinjauan tersebut, terdapat scientific gap berupa terbatasnya penelitian yang mengintegrasikan algoritma Random Forest dengan proses hyperparameter tuning berbasis Grid Search Cross-Validation, analisis feature importance, dan implementasi penuh ke dalam aplikasi web yang siap digunakan oleh tenaga kesehatan di layanan primer seperti posyandu. Penelitian ini menutup gap tersebut dengan Menggunakan algoritma Random Forest untuk mengklasifikasikan kondisi gizi anak balita di Posyandu Makmur 1 berdasarkan data antropometri, mengoptimalkan model melalui Grid Search Cross-Validation, serta mengimplementasikannya Dibuat dalam aplikasi berbasis web menggunakan framework Flask dengan cara pengembangan perangkat lunak yang cepat adalah Rapid Application Development (RAD). Flask dipilih karena bersifat ringan dan mudah diintegrasikan dengan model machine learning sehingga memungkinkan integrasi model tersebut dimasukkan ke dalam sistem yang dapat diakses secara online oleh para pengguna [11], sedangkan metode RAD dipilih karena penekanannya pada cepatnya proses pengembangan melalui serangkaian iterasi dan pembuatan prototype yang melibatkan umpan balik dari pengguna secara aktif [12].

Tujuan dari penelitian ini adalah: (1) menerapkan algoritma Random Forest untuk mengelompokkan status gizi anak-anak di Posyandu Makmur 1 ke dalam empat kelompok yang meliputi Gizi Buruk, Gizi Kurang, Gizi Baik, dan Gizi Lebih; (2) mengembangkan model klasifikasi yang mencapai tingkat akurasi tertinggi dengan menggunakan variabel usia, jenis kelamin, berat badan, tinggi badan, serta lingkaran lengan atas; serta (3) menerapkan model tersebut ke dalam aplikasi berbasis web agar bisa digunakan dengan mudah, cepat, dan efisien oleh para tenaga kesehatan dalam membantu deteksi dini kondisi gizi balita.

2. METODE PENELITIAN

Penelitian ini menerapkan metode pengembangan sistem dengan pendekatan Rapid Application Development (RAD) yang terbagi menjadi empat tahap yang berulang, yaitu Perencanaan Kebutuhan, Desain Pengguna, Konstruksi, dan Pemindahan. Gambar 1 menampilkan langkah-langkah pengembangan sistem yang digunakan dalam penelitian ini.



Gambar 1. Langkah-langkah Pengembangan Sistem dengan Metode RAD

2.1 Pengumpulan Data

Data diperoleh melalui wawancara dan pengamatan terhadap tenaga kesehatan di Posyandu Makmur 1, serta dokumentasi rekam medis balita usia 0-59 bulan. Dataset yang digunakan berjumlah 263 record dengan tujuh variabel input, yaitu usia, jenis kelamin, berat badan, tinggi badan, lingkaran lengan atas

(LILA), dan indeks massa tubuh (IMT), dan riwayat penyakit infeksi, serta satu variabel target berupa status gizi. Label status gizi ditentukan berdasarkan perhitungan Z-Score Tinggi Badan menurut Usia (TB/U) mengikuti standar WHO Child Growth Standards 2006, dengan kategori: Gizi Buruk (Z-Score < -3 SD), Gizi Kurang (-3 SD s.d. < -2 SD), Gizi Baik (-2 SD s.d. +2 SD), dan Gizi Lebih (> +2 SD). Distribusi kelas pada dataset ditunjukkan pada Tabel 1.

Tabel 1. Distribusi Status Gizi Balita pada Dataset

Status Gizi	Jumlah	Persentase
Gizi Lebih	145	55,1%
Gizi Baik	58	22,1%
Gizi Buruk	43	16,3%
Gizi Kurang	17	6,5%
Total	263	100%

2.2 Praproses Data

Tahap praproses data meliputi empat langkah utama. Pertama, usia balita dalam satuan bulan dihitung dari selisih tanggal pemeriksaan dan tanggal lahir dibagi rata untuk jumlah hari dalam sebulan (30,44 hari). Kedua, variabel jenis kelamin dikonversi menjadi numerik melalui label encoding (Laki-laki = 1, Perempuan = 0). Ketiga, Seluruh fitur numerik diatur ulang menggunakan StandardScaler agar memiliki rata-rata 0 dan standar deviasi 1, dengan rumus:

$$Z = (X - \mu) / \sigma$$

dengan μ adalah nilai rata-rata dan σ adalah simpangan baku dari fitur pada data latih. Keempat, dataset data dipisah menjadi set data latih dan set data uji dengan proporsi 80:20 menggunakan metode stratified split agar proporsi setiap kelas tetap terjaga, sehingga menghasilkan 210 data latih dan 53 data uji.

2.3 Algoritma Random Forest

Random Forest merupakan teknik pembelajaran yang berbasis ensemble yang menghasilkan banyak pohon keputusan secara bersamaan secara bersamaan dan menggabungkan hasil dari setiap pohon dengan cara memilih jawaban yang paling banyak didukung. Algoritma ini mengatasi masalah overfitting yang umum terjadi pada single decision tree melalui dua prinsip utama, yaitu Bootstrap Aggregating (Bagging) dan Random Feature Selection [13]. Prediksi akhir Random Forest dirumuskan sebagai:

$$y = \operatorname{argmax} \sum I(h_b(x) = c), c \in C$$

dengan B adalah jumlah pohon keputusan, $h_b(x)$ adalah prediksi pohon ke-b untuk input x, C adalah himpunan kelas kondisi Gizi Baik, Gizi Buruk, Gizi Kurang, Gizi Lebih, dan I adalah fungsi indikator. Pemilihan fitur terbaik pada setiap node pohon menggunakan kriteria Gini Index, yang mengukur tingkat impurity suatu node:

$$Gini = 1 - \sum p_i^2, i = 1 \text{ sampai } 4$$

Model diimplementasikan menggunakan pustaka Scikit-learn pada Python dengan random_state = 42 digunakan untuk memastikan konsistensi hasil. Hyperparameter pada model dioptimalisasi dengan Grid Search Cross-Validation 5-lipat, dimana kombinasi parameter terbaik tercantum di Tabel 2.

Tabel 2. Hasil Grid Search dan Parameter Model Random Forest

Parameter	Nilai yang Diuji	Hasil Terbaik
n_estimators	50, 100, 200	200
max_depth	5, 10, 15	15
min_samples_split	2, 5, 10	5
min_samples_leaf	1, 2, 4	2

2.4 Evaluasi Model dan Implementasi Sistem

Evaluasi model dilakukan pada 53 data uji menggunakan empat metrik, yaitu akurasi, presisi, recall, dan F1-Score, yang dihitung sebagai berikut:

$$Accuracy = (\Sigma \text{Prediksi Benar}) / (\text{Total Data Uji})$$

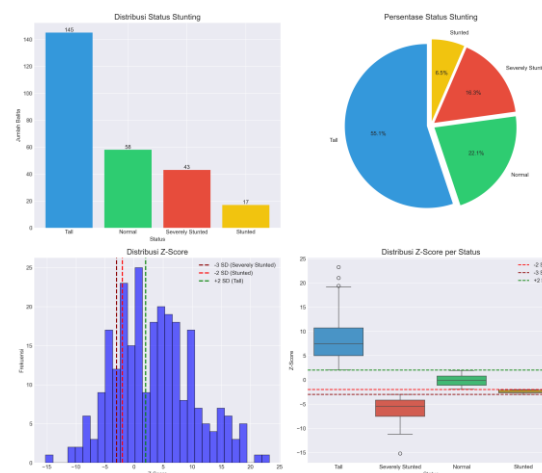
$$Precision = TP / (TP + FP) \quad Recall = TP / (TP + FN)$$

$$F1\text{-Score} = 2 \times (Precision \times Recall) / (Precision + Recall)$$

Dengan TP (True Positive) adalah jumlah data dari suatu kelas yang berhasil diprediksi dengan benar, FP (False Positive) adalah data dari kelas yang berbeda yang salah diprediksi sebagai kelas tersebut, dan FN (False Negative) adalah data dari kelas tersebut yang keliru diprediksi sebagai kelas lain. Model yang sudah dievaluasi kemudian diintegrasikan ke dalam aplikasi berbasis web dengan menggunakan framework Flask, dengan basis data MySQL untuk menyimpan riwayat pemeriksaan. Sistem menyediakan fitur input data antropometri, klasifikasi status gizi secara real-time, manajemen data riwayat, serta autentikasi pengguna. Pengujian sistem dilakukan melalui White Box Testing pada logika internal (fungsi preproses, scoring model, dan endpoint API) dan Uji Black Box menggunakan teknik Equivalence Partitioning terhadap semua fitur dalam sistem.

3. HASIL DAN PEMBAHASAN

Hasil analisis deskriptif mengungkapkan bahwa kategori Gizi Lebih paling banyak terdapat dalam dataset, yaitu (145 balita, 55,1%), diikuti Gizi Baik (58 balita, 22,1%), Gizi Buruk (43 balita, 16,3%), dan Gizi Kurang (17 balita, 6,5%), sebagaimana ditunjukkan pada Gambar 2. Ketidakseimbangan ini berpengaruh terhadap performa model pada kelas minoritas, sebagaimana dibahas lebih lanjut pada bagian evaluasi model.



Gambar 2. Distribusi Status Gizi Balita pada Dataset

3.1 Evaluasi Model Random Forest

Kombinasi parameter terbaik adalah $n_estimators = 200$ dan $max_depth = 15$, dengan akurasi rata-rata dari validasi silang mencapai $0,821 \pm 0,024$. Model dengan parameter tersebut kemudian diuji pada 53 data uji, menghasilkan confusion matrix seperti yang tercantum pada Tabel 3.

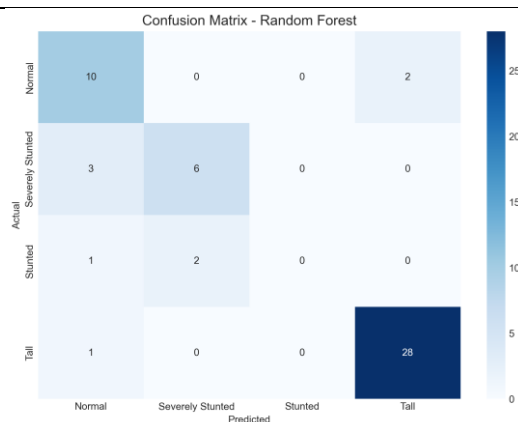
Tabel 3. Confusion Matrix Hasil Pengujian Model

Aktual \ Prediksi	Gizi Baik	Gizi Buruk	Gizi Kurang	Gizi Lebih	Total
Gizi Baik	10	0	0	2	12
Gizi Buruk	3	6	0	0	9
Gizi Kurang	1	2	0	0	3
Gizi Lebih	1	0	0	28	29
Total	15	8	0	30	53

Berdasarkan Tabel 3, model berhasil mengklasifikasikan 44 dari 53 data uji secara tepat, sehingga diperoleh akurasi keseluruhan sebesar 83,02%. Rekapitulasi precision, recall, dan F1-Score per kelas yang dicantumkan pada Tabel 4.

Tabel 4. Rekapitulasi Metrik Evaluasi per Kelas

Kelas	Precision	Recall	F1-Score	Support
Gizi Baik	66,67%	83,33%	74,07%	12
Gizi Buruk	75,00%	66,67%	70,59%	9
Gizi Kurang	0,00%	0,00%	0,00%	3
Gizi Lebih	93,33%	96,55%	94,92%	29
Rata-rata	83,02%	83,02%	83,02%	53



Gambar 3. Visualisasi Confusion Matrix Model Random Forest

Model menunjukkan performa terbaik pada kelas Gizi Lebih, yang memiliki jumlah data uji paling banyak (29 data), dengan precision 93,33%, recall 96,55%, dan F1-Score 94,92%. Sebaliknya, model mengalami kesulitan mengenali kelas Gizi Kurang, di mana seluruh 3 data uji pada kelas tersebut salah diklasifikasikan sebagai Gizi Baik atau Gizi Buruk. Kondisi ini disebabkan oleh Kelas gizi kurang hanya ada 17 dari total 263 data, yang artinya hanya sekitar 6,5% dari seluruh dataset, sehingga model kurang memiliki cukup contoh untuk mempelajari pola pada kelas tersebut secara memadai.

Mengingat akurasi keseluruhan (83,02%) berpotensi menyamarkan performa model pada kelas minoritas akibat distribusi data yang tidak seimbang, evaluasi tambahan dilakukan menggunakan macro F1-Score dan balanced accuracy, yang memberi bobot setara pada setiap kelas terlepas dari jumlah datanya. Hasil perhitungan menunjukkan macro F1-Score sebesar 59,90% dan balanced accuracy sebesar 61,64%, jauh lebih rendah dibandingkan akurasi keseluruhan. Kesenjangan ini mengonfirmasi bahwa performa model secara aktual sangat dipengaruhi oleh dominasi kelas Gizi Lebih dan Gizi Baik, sementara kelas Gizi

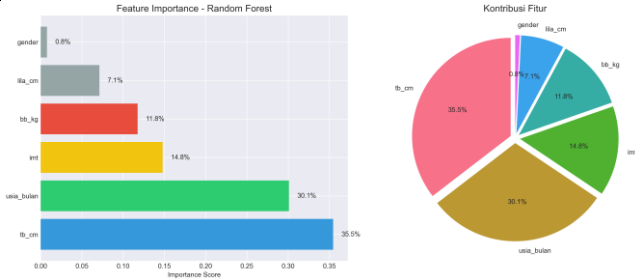
Kurang sama sekali tidak berhasil dikenali (precision, recall, dan F1-Score sebesar 0%). Temuan ini menegaskan perlunya penanganan class imbalance yang lebih eksplisit pada penelitian selanjutnya, misalnya melalui teknik oversampling seperti SMOTE atau pembobotan kelas (class weighting), agar performa model pada kelas minoritas dapat ditingkatkan tanpa mengorbankan akurasi keseluruhan. Perlu diakui pula bahwa model dalam penelitian ini hanya dilatih dan diuji pada data dari satu lokasi, yaitu Posyandu Makmur 1, sehingga kemampuan generalisasinya pada populasi balita di wilayah lain dengan karakteristik demografis berbeda masih belum teruji; validasi eksternal menggunakan dataset dari posyandu atau puskesmas lain diperlukan sebelum model ini diterapkan secara lebih luas.

3.2 Feature Importance

Analisis feature importance menggunakan metode Gini Importance menunjukkan kontribusi setiap variabel input terhadap keputusan klasifikasi model, sebagaimana ditunjukkan pada Tabel 5 dan Gambar 4.

Tabel 5. Hasil Feature Importance

Fitur	Importance	Persentase
Tinggi Badan (TB)	0,3550	35,5%
Usia	0,3010	30,1%
IMT	0,1480	14,8%
Berat Badan (BB)	0,1180	11,8%
LILA	0,0710	7,1%
Jenis Kelamin	0,0080	0,8%



Gambar 4. Feature Importance Model Random Forest

Variabel Tinggi Badan memberikan kontribusi paling dominan (35,5%), diikuti oleh Usia (30,1%); keduanya menyumbang lebih dari 65% dari total importance. Temuan ini sejalan dengan prinsip perhitungan Z-Score Tinggi Badan menurut Usia (TB/U) yang menjadi dasar pelabelan data pada penelitian ini, serta konsisten dengan hasil analisis korelasi yang menunjukkan Tinggi Badan memiliki korelasi terkuat terhadap Z-Score (0,76). Variabel IMT dan Berat Badan berkontribusi masing-masing 14,8% dan 11,8%, sedangkan LILA dan Jenis Kelamin memberikan kontribusi relatif kecil (7,1% dan 0,8%), yang mengindikasikan bahwa kedua variabel tersebut kurang diskriminatif dalam membedakan kelas status gizi pada dataset ini.

3.3 Pengujian Sistem

Pengujian performa model pada tahap Cutover menunjukkan bahwa seluruh kriteria kelayakan minimum terpenuhi, seperti yang ditunjukkan dalam Tabel 6.

Tabel 6. Hasil Pengujian Performa Model dan Sistem

Aspek Pengujian	Kriteria	Hasil	Status
Akurasi model	$\geq 80\%$	83,02%	Lulus
Precision kelas mayoritas	$\geq 80\%$	93,33%	Lulus
Recall kelas mayoritas	$\geq 80\%$	96,55%	Lulus
F1-Score kelas mayoritas	$\geq 80\%$	94,92%	Lulus
Waktu pemuatan model	≤ 3 detik	$\approx 1,2$ detik	Lulus
Waktu respons endpoint /predict	≤ 2 detik	$\approx 0,4$ detik	Lulus

Selain pengujian performa model, pengujian fungsional sistem dilakukan melalui dua pendekatan yang saling melengkapi. Black Box Testing menggunakan teknik Equivalence Partitioning pada 13 skenario uji mengevaluasi perilaku sistem dari sisi pengguna tanpa memeriksa kode program, dengan membagi setiap input menjadi kelas data valid dan tidak valid pada empat modul utama, yaitu Login, Dashboard, Klasifikasi, dan Riwayat; seluruh skenario menunjukkan bahwa keluaran sistem sesuai dengan spesifikasi kebutuhan yang ditetapkan. White Box Testing pada 10 skenario uji menelusuri struktur kode dan alur logika internal sistem secara langsung, mencakup fungsi pra-proses data (penanganan nilai kosong dan encoding fitur), proses scoring model klasifikasi, serta endpoint API Flask (/predict) yang menjadi penghubung antara model dan antarmuka pengguna; seluruh jalur eksekusi (path) yang diuji berjalan sesuai dengan logika yang dirancang tanpa ditemukan error. Dengan demikian, Sistem klasifikasi status gizi untuk balita yang berbasis Random Forest yang dikembangkan pada penelitian ini terbukti layak digunakan sebagai alat bantu deteksi dini di Posyandu Makmur 1, meskipun performanya pada kelas minoritas masih perlu ditingkatkan pada penelitian selanjutnya, misalnya melalui teknik oversampling seperti SMOTE [14],[15].

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma Random Forest untuk mengklasifikasikan status gizi anak-anak balita. di Posyandu Makmur 1 ke dalam empat kategori berdasarkan data antropometri, serta mengimplementasikannya dibuat dalam aplikasi berbasis web dengan menggunakan framework Flask melalui metode pengembangan Rapid Application Development (RAD). Model dengan kombinasi hyperparameter optimal $n_estimators = 200$ dan $max_depth = 15$ hasil Grid Search Cross-Validation mencapai akurasi sebesar 83,02% pada data uji, dengan performa terbaik pada kelas Gizi Lebih (precision 93,33%, recall 96,55%, F1-Score 94,92%). Analisis feature importance menunjukkan bahwa variabel Tinggi Badan (35,5%) dan Usia (30,1%) merupakan fitur paling berpengaruh terhadap keputusan klasifikasi. Seluruh pengujian White Box dan Black Box menunjukkan bahwa sistem bekerja sesuai dengan kebutuhan pengguna, serta seluruh kriteria kelayakan performa model, meliputi akurasi, precision, recall, F1-Score, waktu pemuatan model, dan waktu respons, dinyatakan lulus. Model masih memiliki keterbatasan dalam mengenali kelas Gizi Kurang akibat jumlah data yang sangat sedikit pada kelas tersebut. Penelitian selanjutnya disarankan untuk menambah data latih pada kelas minoritas menggunakan teknik oversampling seperti SMOTE, Membandingkan hasil kerja Random Forest dengan algoritma lain seperti XGBoost atau SVM, serta mengembangkan fitur tambahan seperti notifikasi otomatis dan integrasi dengan sistem informasi posyandu yang telah berjalan.

UCAPAN TERIMA KASIH

Penulis ingin menyampaikan apresiasi kepada tenaga kesehatan di Posyandu Makmur 1 yang telah memberikan bantuan berupa data dan informasi sepanjang berlangsungnya penelitian ini, serta kepada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang yang telah memberikan arahan akademik dalam menyelesaikan penelitian ini. Penelitian ini tidak menerima pendanaan khusus dari lembaga manapun.

PERNYATAAN KONFLIK KEPENTINGAN

Para penulis menyatakan tidak ada konflik kepentingan.

PENDANAAN
penelitian ini tidak menerima pendanaan eksternal.

DAFTAR PUSTAKA

- [1] R. Sentika *et al.*, “Expert Consensus on Interprofessional Collaboration (IPC) Guidelines on Stunting Management in Indonesian Primary Healthcare (Puskesmas),” *Open Public Health J.*, pp. 1–12, 2024, doi: 10.2174/0118749445352608241119164446.
- [2] Muhammad Davin Diza Ghifary, Chairunisah, Said Iskandar Al Idrus, Faridawaty Marpaung, and Kana Saputra, “Classification of Stunting using a Website-Based Support Vector Machine (SVM) Algorithm (Case Study: Pagar Merbau Community Health Center),” *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, vol. 5, no. 1, pp. 2001–2006, 2025, doi: 10.59934/jaiea.v5i1.1774.
- [3] I. Ozcan, H. Aydin, and A. Cetinkaya, “Comparison of Classification Success Rates of Different Machine Learning Algorithms in the Diagnosis of Breast Cancer,” *Asian Pacific Journal of Cancer Prevention*, vol. 23, no. 10, pp. 3287–3297, 2022, doi: 10.31557/APJCP.2022.23.10.3287.
- [4] W. S. Lestari, Y. M. Saragih, and Caroline, “Comparison of Deep Neural Networks and Random Forest Algorithms for Multiclass Stunting Prediction in Toddlers,” *Teknika*, vol. 13, no. 3, pp. 412–417, 2024, doi: 10.34148/teknika.v13i3.1063.
- [5] Mundirin, Idawati, and I. Latief, “Klasifikasi Status Gizi Balita Berbasis Data Antropometri menggunakan Random Forest,” *Journal of Computer Science and Informatics Engineering*, vol. 4, no. 4, pp. 324–333, 2025, [Online]. Available: <https://journal.aira.or.id/index.php/cosie/article/view/1175>
- [6] M. Fatmawati, B. A. Herlambang, and N. Q. Nada, “Random Forest Algorithm for Toddler Nutritional Status Classification Website,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 8, no. 2, pp. 428–433, 2024.
- [7] R. Belferik, F. M. Sinaga, M. A. S. Manullang, and T. Sinaga, “Addressing Class Imbalance in Stunting Classification Using SMOTE Enhanced Random Forest,” *Sinkron : Jurnal dan Penelitian Teknik Informatika*, vol. 9, no. 4, pp. 2108–2116, 2025.
- [8] Z. I. Bimawan, T. Astuti, and P. Arsi, “Comparison of Random Forest, K-Nearest Neighbor, Decision Tree, and Xgboost Algorithms for Detecting Stunting in Toddlers,” *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 6, pp. 1599–1607, 2024, doi: 10.52436/1.jutif.2024.5.6.2629.
- [9] A. H. Mubarak, P. Pujiono, D. Setiawan, D. F. Wicaksono, and E. Rimawati, “Parameter Testing on Random Forest Algorithm for Stunting Prediction,” *Sinkron*, vol. 9, no. 1, pp. 107–116, 2025, doi: 10.33395/sinkron.v9i1.14264.
- [10] T. Sugihartono, B. Wijaya, Marini, A. F. Alkayes, and H. A. Anugrah, “Optimizing Stunting Detection through SMOTE and Machine Learning: a Comparative Study of XGBoost, Random Forest, SVM, and k-NN,” *Journal of Applied Data Sciences*, vol. 6, no. 1, pp. 667–682, 2025, doi: 10.47738/jads.v6i1.494.
- [11] Elan Suherlan, Shindy Arti, Siti Nabilah, and Zahwah Hazimah, “Penerapan Flask Framework Untuk Deployment Model Machine Learning Dalam Mendukung Analisis Adaptasi Mahasiswa Pada Pembelajaran Daring,” *PINTER : Jurnal Pendidikan Teknik Informatika dan Komputer*, vol. 9, no. 1, pp. 110–117, 2025, doi: 10.21009/pinter.9.1.15.
- [12] A. Meyliana, P. T. Rapiyanta, and A. Andriani, “Application of the Rapid Application Development (RAD) Method for Web-Based Financial Management and Wood Inventory Using CodeIgniter,” *ARRUS Journal of Engineering and Technology*, vol. 4, no. 1, pp. 81–89, 2024, doi: 10.35877/jetech2722.
- [13] M. Ibrahim, “Evolution of Random Forest from Decision Tree and Bagging : A Bias-Variance Perspective,” vol. 7, no. 1, pp. 66–71, 2022.
- [14] I. Fadil, R. S. Manggala, E. Firmansyah, and M. A. Helmiawan, “Performance Optimization of Support Vector Machine with SMOTE for Multiclass Stunting Prediction in Sumedang District , Indonesia,” vol. 6, no. 4, pp. 2917–2928, 2025.
- [15] S. Lestari, “Prediction of Stunting in Toddlers Using Bagging and Random Forest Algorithms,” *Sinkron : Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 2, pp. 947–955, 2024.