JTIF | Jurnal INOVATIF WIRA WACANA Vol. 04 No. 02 September 2025, hlm 233-240 **pISSN (2961-7316)**
eISSN (2962-5998)

IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK KLASIFIKASI STATUS GIZI BALITA

(Studi kasus: Puskesmas Kambaniru)

(IMPLEMENTATION OF RANDOM FOREST ALGORITHM FOR TODDLER NUTRITIONAL STATUS CLASSIFICATION, A CASE STUDY AT KAMBANIRU HEALTH CENTER)

Jeaneth Frisilia Konda ngguna[1], Arini Aha Pekuwali[2], Leonard Marthen Doni Ratu[3]

[1,2,3]Program Studi Teknik Informatika, Universitas Kristen Wira Wacana Sumba

E-mail: [1]jeanethfrisiliakondangguna@gmail.com, [2]arini.pekuwali@unkriswina.ac.id, [3]leonard.ratu@unkriswina.ac.id

KEYWORDS:
Classification, Nutritional Status of Toddlers, Imbalanced Data, Random Forest Method

ABSTRACT

East Sumba faces significant toddler malnutrition challenges, with 14.9% stunting prevalence (SSGI 2022). Poor nutritional intake can have a negative impact on child growth and development. Therefore, a method is needed that can accurately and efficiently classify the nutritional status of infants as a basis for decision-making in efforts to prevent and combat malnutrition in children. This study is useful for implementing a model for classifying the nutritional status of infants using the Random Forest algorithm method to aid in rapid and accurate identification. The data used was obtained from the Kambaniru Community Health Center and Village Office, and the children's data was divided into three categories: good nutrition, poor nutrition, and malnutrition. The variables used were age at measurement, weight, height, gender, arm circumference, measurement method, parents' occupation, and number of family members. The model was trained by dividing the data into 80% training data and 20% test data. Model performance was evaluated using a confusion matrix, accuracy, precision, recall, and F1-score.

KATA KUNCI:
Klasifikasi*, *Status Gizi Balita*, Data Yang Tidak Seimbang, *Metode Random Forest

ABSTRAK

Sumba Timur menghadapi tantangan malnutrisi balita yang signifikan, dengan prevalensi stunting sebesar 14,9% (SSGI 2022). Asupan gizi yang kurang baik dapat berdampak buruk terhadap tumbuh kembang anak. Oleh karena itu, diperlukan metode yang mampu mengklasifikasikan kondisi gizi balita secara akurat dan efisien sebagai dasar pengambilan keputusan dalam upaya pencegahan dan penanggulangan masalah kurang gizi pada anak. Penelitian ini berguna untuk mengimplementasikan model klasifikasi status gizi balita menggunakan metode algoritma Random Forest guna membantu proses identifikasi secara cepat dan akurat. Data yang digunakan diperoleh dari Puskesmas dan Kelurahan Kambaniru, kemudian data anak terbagi dalam tiga kategori: gizi baik, gizi kurang, dan gizi buruk. Variabel yang digunakan yaitu usia saat ukur, berat badan, tinggi badan, jenis kelamin, lingkar lengan, cara ukur, pekerjaan orang tua, dan jumlah keluarga. Model dilatih dengan membagi data menjadi 80% data latih dan 20% data uji. Evaluasi performa model menggunakan confusion matrix, akurasi, precision, recall, dan F1-score.

PENDAHULUAN

Kondisi Gizi balita merupakan indikator penting dalam menentukan kualitas kesehatan generasi masa depan [1]. Gizi yang tidak seimbang pada masa awal kehidupan dapat menimbulkan dampak jangka pendek seperti gangguan pertumbuhan fisik, penurunan daya tahan tubuh, dan keterlambatan perkembangan kognitif [2]. Dalam jangka panjang, kondisi ini dapat berujung pada rendahnya produktivitas, peningkatan risiko penyakit kronis, serta penurunan kualitas sumber daya manusia secara keseluruhan [3].

Di Indonesia, masalah gizi pada balita masih menjadi tantangan serius. Tantangan utama dalam klasifikasi data kesehatan adalah data yang tidak lengkap, tidak seimbang, dan berisik, sehingga perlu proses pembersihan dan pemilihan fitur yang tepat agar hasilnya akurat. Data dari Survei Status Gizi Indonesia (SSGI) tahun 2022 menunjukkan bahwa prevalensi stunting nasional mencapai 21,6% [4]. Provinsi Nusa Tenggara Timur (NTT) mencatat angka yang lebih tinggi, yaitu 35,3% [5]. Secara khusus, Kabupaten Sumba Timur memiliki prevalensi stunting sebesar 14,9% pada tahun 2022. Di wilayah kerja Puskesmas Kambaniru, Kecamatan Kambera, terdapat 590 balita yang ditimbang pada tahun 2021, dengan 472 diantaranya memiliki status gizi baik, sementara sisanya mengalami gizi kurang atau buruk.

Seiring berkembangnya teknologi salah satu pendekatan yang menjanjikan adalah penerapan teknologi data *science*, khususnya algoritma *Random Forest* yang dimana metode yang digunakan pada penelitian ini menjadi perbedaan antara penelitian terdahulu seperti Azzahra et al. (2024). [6]. *Random Forest* merupakan metode *ensemble learning* yang menggabungkan banyak pohon keputusan *(decision tree)* untuk membentuk model klasifikasi yang kuat [7]. Setiap pohon dilatih menggunakan data acak dan subset fitur yang berbeda, sehingga hasilkan model yang keluar tidak hanya akurat, tetapi juga tahan terhadap *overfitting* [8].

Dalam konteks klasifikasi status gizi balita, *Random Forest* bekerja dengan menganalisis berbagai fitur yang terkait data yang diambil di puskesmas untuk menentukan kategori gizi: baik, kurang, atau buruk. Proses klasifikasi dilakukan melalui voting mayoritas, di mana setiap pohon memberikan prediksi, dan hasil akhir ditentukan oleh prediksi terbanyak [9].

Keunggulan lainnya dari *Random Forest* adalah kemampuannya untuk melakukan evaluasi performa model secara menyeluruh. Pada tahap evaluasi ini dilakukan dengan metrik seperti akurasi, *precision*, *recall*, dan F1-*score*, yang dapat memberikan gambaran sejauh mana model berhasil mengklasifikasikan data dengan tepat. Selain itu, *Random Forest* juga dapat menghitung *importance* dari masing-masing fitur, sehingga dapat diketahui variabel mana yang paling berpengaruh dalam menentukan masing-masing status gizi yang ada [10].

Dengan kemampuan tersebut, penerapan *Random Forest* dalam klasifikasi status gizi balita, khususnya di wilayah kerja Puskesmas Kambaniru, diharapkan dapat menjadi solusi yang inovatif dan berbasis data dalam membantu pengambilan keputusan dan intervensi kesehatan masyarakat.

METODE PENELITIAN

A. Alur Penelitian

Gambar 1. Alur Penelitian

1. Studi Literatur

Pada tahap ini, penelitian melakukan kajian terhadap Teori status gizi balita dan indikator pengukuran BB/TB), Metode *Machine Learning*, khususnya *Random Forest*. Penelitian-penelitian

terdahulu terkait klasifikasi status gizi *menggunakan Machine Learning*. Tujuan dari studi literatur adalah memberikan dasar teori yang kuat untuk mendukung pelaksanaan penelitian

2. Pengumpulan Data

   Data yang dikumpulkan dari Puskesmas Kambaniru dan Kelurahan Kambaniru. Data yang diambil meliputi umur, bb, tb, jk, serta status gizi balita lainnya dan Data sosial ekonomi. Data ini menjadi dataset utama untuk proses klasifikasi.

   Pengumpulan data dilakukan melalui dua cara:

   - Data Primer: Data tentang status gizi balita diperoleh langsung dari puskesmas kambaniru dan kelurahan kambaniru dengan pengukuran BB dan TB anak di posyandu.
   - Data Sekunder: Data tambahan seperti umur dan jenis kelamin diambil dari catat kesehatan anak di puskesmas dan posyandu.

3. Preprocessing Data

   Tahap ini bertujuan untuk membersihkan, mengubah, dan menyiapkan data mentah sebelum dapat digunakan untuk analisis atau pemodelan data yang dikumpulkan di proses agar siap digunakan dalam pemodelan *Machine Learning.* Langkah-langkah preprocessing meliputi:

   - Pembersihan data: menghilangkan data duplikat, kosong, atau eror
   - Transformasi data: meliputi encoding untuk variabel kategorikal (jenis kelamin) dan normalisasi untuk variabel numerik (berat badan, tinggi badan, dan umur) jika diperlukan
   - Normalisasi data numerik: Fitur numerik seperti berat, tinggi, dan umur bisa memiliki rentang nilai yang berbeda-beda.
   - Splitting Data: bagi data menjadi data latih *(training* set) dan data uji (*testing*
   - set), biasanya dengan rasio 80:20.

4. Penerapan Metode

   Implementasi algoritma *Random Forest* disini untuk membangun model klasifikasi yang dapat menentukan status gizi berdasarkan beberapa fitur seperti umur, bb, tb, jk, dan variabel relevan lainnya.

5. Evaluasi Model

   Evaluasi model dilakukan dengan membandingkan prediksi model dengan nilai sebenarnya dari data pengujian yang sebelumnya tidak digunakan dalam proses pelatihan model. Model dievaluasi dengan beberapa metrik yaitu confusion matrix, akurasi, precision, recall, dan F1-score. Selain itu juga ditampilkan grafik feature importance untuk menunjukkan pengaruh masing-masing fitur terhadap keputusan model.

6. Pengujian dan Validasi Model

   Tahap ini bertujuan untuk menguji kinerja dan keandalan model dalam mengklasifikasikan data yang belum pernah dilihat sebelumnya. Dengan demikian, pengujian dan validasi menjadi penting untuk mengetahui apakah model hanya bekerja baik pada data pelatihan atau juga dapat bekerja baik pada data baru.

HASIL DAN PEMBAHASAN

PENGUMPULAN DATA

Dalam penelitian ini, data yang dikumpulkan dari Puskesmas Kambaniru dan Kelurahan Kambaniru. Data ini mencangkup atribut seperti usia saat ukur, berat badan, tinggi badan, cara ukur, lingkar lengan jenis

kelamin, serta data sosial ekonomi. Data ini menjadi dataset utama untuk proses klasifikasi. Jumlah total data yang dikumpulkan adalah sebanyak 177 data. Berikut adalah data pada bulan Februari 2025:

Tabel 1. Data balita bulan Februari 2025

No	Usia Saat Ukur	BB	TB	Cara Ukur	Lingkar Lengan	Jenis Kelamin
1	2 Tahun - 11 Bulan - 14 Hari	3	49	Berdiri	14.8	P
2	4 Tahun - 0 Bulan - 27 Hari	2.9	49	Berdiri	13.5	P
3	2 Tahun - 7 Bulan - 30 Hari	3.1	49	Berdiri	15.3	L
4	1 Tahun - 11 Bulan - 23 Hari	3.6	48	Talentang	13	L
5	1 Tahun - 11 Bulan - 23 Hari	3.5	50	Talentang	14	P
.....	.....	.....	.....	.....	.....	.....
177	1 Tahun - 11 Bulan - 23 Hari	3.9	49	Talentang	13	L

PREPROCESSING DATA

Setelah data terkumpul, tahap selanjutnya adalah *preprocessing* data. Istilah preprocessing data, yang merupakan langkah krusial untuk memastikan kualitas dan konsistensi data yang akan digunakan dalam model pembelajaran mesin. Langkah-langkah yang dilakukan pada tahap preprocessing mencakup beberapa tahapan penting, yaitu:

1. *Cleaning*

   Langkah pertama dalam pembersihan data adalah memeriksa dan menghapus data duplikat. Data duplikat dapat terjadi karena kesalahan dalam pengumpulan atau entri data, yang dapat menyebabkan model bekerja dengan data yang tidak representatif, sehingga menghasilkan prediksi yang bias. Dan pada tahap ini tidak ada data yang tidak relevan sehingga pada saat melakukan tahap cleaning tidak ada perubahan apapun pada data.
2. Transformasi Data

   Selanjutnya pada tahap ini, data yang telah dikumpulkan dan dibersihkan akan dipersiapkan untuk digunakan dalam model pembelajaran mesin. Salah satu teknik yang digunakan adalah Label *Encoding* dan *One-Hot Encoding*, yang bertujuan untuk mengubah variabel kategorikal menjadi format numerik agar dapat diproses oleh algoritma *machine learning.*

Tabel 2. Hasil One-Hot Encoding

Sebelum proses	Sesudah proses
Perempuan	[0,1]
Laki-laki	[1,0]

3. Normalisasi Data Numerik

Pada tahap ini, proses untuk skala variabel numerik agar berada dalam rentang yang seragam. Tujuannya adalah untuk memastikan bahwa semua fitur numerik memiliki skala yang sebanding, sehingga tidak ada satu variabel yang mendominasi model machine learning hanya karena memiliki skala yang lebih besar. Pada tahap ini, digunakan metode *Min-Max Scaling* untuk melakukan normalisasi pada data numerik. *Min-Max Scaling* mengubah data sehingga semua nilai berada dalam rentang [0, 1].

Tabel 3. Hasil Normalisasi Data Numerik

Nama	Berat (kg)	Tinggi (cm)	Normalisasi (Min-Max)
Haikal	3.0	49.0	Berat:0.6 Tinggi:0.9
Nadia	2.9	49.0	Berat:0.5 Tinggi:0.9
Indri	3.1	49.0	Berat:0.6 Tinggi:0.9

4. Splitting Data

Pada tahap ini dilakukan pembagian dataset menjadi dua bagian utama, yaitu data latih (*training* data) dan data uji (*testing* data). Proses ini penting untuk menguji performa model yang dibangun. Pembagian data dilakukan dengan menggunakan fungsi *train_test_split* dari pustaka *sklearn*. Model_*selection*, yang membagi data secara acak menjadi dua subset: satu untuk melatih model dan satu lagi untuk menguji model yang telah dilatih.

PENERAPAN METODE

Prosesnya dimulai dengan membagi dataset menjadi dua bagian: 80% untuk pelatihan dan 20% untuk pengujian. Hal ini dicapai dengan menggunakan fungsi *train_test*_split, yang memastikan bahwa data dibagi secara acak sambil mempertahankan distribusi kelas. Selanjutnya, Pengklasifikasi Hutan Acak diinisialisasi dengan 100 pohon keputusan, menyediakan model ansambel yang kuat yang memanfaatkan pengambilan keputusan kolektif dari beberapa pohon. Untuk memastikan bahwa hasilnya konsisten dan dapat direproduksi di berbagai proses, benih acak tetap ditetapkan selama inisialisasi. Setelah model siap, model tersebut dilatih pada data pelatihan menggunakan metode fit(). Selama fase pelatihan ini, setiap pohon keputusan di hutan belajar mengklasifikasikan data dengan mempertimbangkan berbagai subset dan fitur, yang berkontribusi pada stabilitas dan akurasi keseluruhan ansambel. Pendekatan ini menggabungkan beberapa pohon keputusan untuk menghasilkan prediksi yang andal, mengurangi *overfitting* dan memperbaiki performa generalisasi.

EVALUASI MODEL

Berikut ini melatih model pada data pelatihan yang ditentukan, performanya kemudian dievaluasi menggunakan data uji terpisah. Proses ini bertujuan untuk menilai akurasi, keandalan, dan kemampuan model untuk melakukan generalisasi secara efektif ke data yang belum pernah dilihat sebelumnya, sehingga memastikan performa optimal dalam aplikasi dunia nyata. Beberapa metrik yang umum digunakan untuk evaluasi model klasifikasi adalah *Confusion Matrix*, *Classification Report, dan Accuracy*. Setelah proses evaluasi dilakukan, berikut adalah hasil yang diperoleh dari algoritma tersebut.

Classification Report:

	precision	recall	f1-score	support
0	0.43	0.27	0.33	11
1	0.00	0.00	0.00	2
2	0.66	0.86	0.75	22
3	0.00	0.00	0.00	1
accuracy			0.61	36
macro avg	0.27	0.28	0.27	36
weighted avg	0.53	0.61	0.56	36

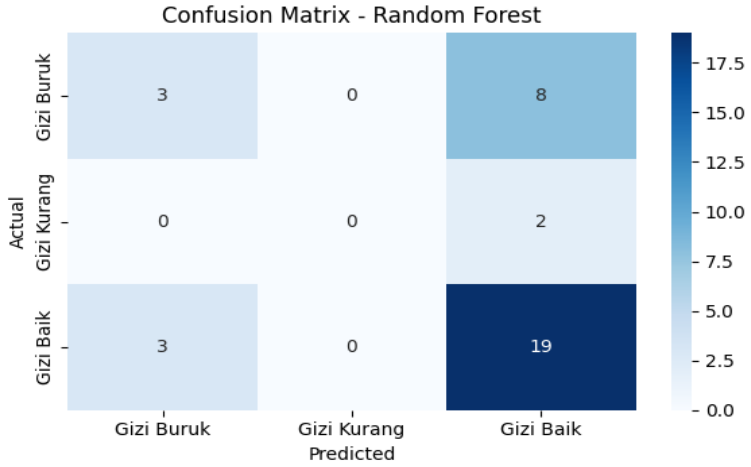
Gambar 2. Hasil Klasifikasi

Hasil evaluasi model menunjukkan metrik-metrik yang menggambarkan performa model dalam mengklasifikasikan status gizi balita. Precision mengukur seberapa banyak prediksi positif yang benar dibandingkan dengan total prediksi positif. Untuk kelas 0 (misalnya "Gizi Buruk"), precision sebesar 0.43 menunjukkan bahwa 43% dari semua prediksi "Gizi Buruk" adalah benar. Kelas 1 ("Gizi Kurang") memiliki precision 0.00, yang berarti model tidak dapat memprediksi dengan benar untuk kelas ini. Kelas 2 ("Gizi Baik") menunjukkan precision 0.66, yang berarti 66% dari semua prediksi "Gizi Baik" benar. Sementara itu, kelas 3 memiliki precision 0.00, menunjukkan bahwa tidak ada prediksi yang benar untuk kelas ini.

Recall mengukur seberapa banyak dari data yang benar-benar positif yang berhasil diprediksi oleh model. Kelas 0 memiliki *recall* 0.27, yang berarti model hanya berhasil mengidentifikasi 27% dari data "Gizi Buruk". Kelas 1 memiliki recall 0.00, menunjukkan bahwa model tidak dapat mengidentifikasi data dari kelas ini. Kelas 2 memiliki *recall* 0.86, yang menunjukkan bahwa model berhasil mengidentifikasi 86% data "Gizi Baik". Kelas 3 memiliki recall 0.00, menunjukkan model gagal mendeteksi data kelas ini.

F1-Score, yang merupakan rata-rata harmonik antara *precision* dan *recall*, memberikan gambaran seimbang tentang kinerja model. Kelas 0 memiliki f1-score 0.33, yang berarti model kurang baik dalam mengklasifikasikan kelas ini. Kelas 1 memiliki *f1-score* 0.00, menunjukkan kinerja yang sangat buruk untuk kelas ini. Kelas 2 memiliki *f1-score* 0.75, yang menghasilkan performa yang cukup baik dalam mengklasifikasikan kelas ini. Kelas 3 memiliki *f1-score* 0.00, menunjukkan tidak adanya kemampuan model dalam mengklasifikasikan kelas ini.

Akurasi model, yaitu persentase prediksi yang benar dari total data, adalah 0.61 (61%), yang dapat dilihat bahwa model berhasil memprediksi dengan benar dengan hasil 61% dari keseluruhan data. Untuk *Macro Average*, yang menunjukkan rata-rata metrik tanpa memperhitungkan distribusi jumlah data antar kelas, diperoleh nilai 0.27 untuk *precision*, 0.28 untuk recall, dan 0.27 untuk f1-score. Sedangkan *Weighted Average*, yang memperhitungkan jumlah data setiap kelas, menghasilkan nilai 0.53 untuk precision, 0.61 untuk *recall*, dan 0.56 untuk *f1-score*.



Gambar 3. *Cunfusion Matriks*

Berdasarkan hasil evaluasi menggunakan *Confusion matrix*, evaluasi model *Random Forest* dalam mengklasifikasikan status gizi balita berdasarkan data uji. Pada kelas "Gizi Buruk", terdapat 3 data yang diprediksi dengan benar sebagai "Gizi Buruk" (*True Positive*/TP), namun 8 data lainnya salah diprediksi sebagai "Gizi Baik" (*False Positive*/FP). Untuk kelas "Gizi Kurang", model gagal memprediksi dengan benar, karena tidak ada data yang diprediksi sebagai "Gizi Kurang" (*True Negative*/TN), dan 2 data lainnya diprediksi sebagai "Gizi Baik" (*False Negative*/FN). Sedangkan pada kelas "Gizi Baik", 19 data diprediksi dengan benar sebagai "Gizi Baik" (TP), tetapi 3 data lainnya diprediksi sebagai "Gizi Buruk" (FP), dan tidak ada data yang diprediksi sebagai "Gizi Kurang".

KESIMPULAN DAN SARAN

Berdasarkan hasil evaluasi model Random Forest yang diterapkan pada klasifikasi status gizi balita, diperoleh beberapa temuan penting. Model menunjukkan akurasi sebesar 61%, yang berarti model berhasil memprediksi 61% dari total data uji dengan benar. Meskipun akurasi ini tergolong cukup baik, model menunjukkan kesulitan dalam mengklasifikasikan kelas Gizi Kurang dan Gizi Buruk, yang tercermin dari nilai presisi dan *recall* yang sangat rendah pada kedua kelas tersebut. Untuk kelas Gizi Baik, presisi mencapai 0.66 dan recall mencapai 0.86, menunjukkan bahwa model berhasil dengan baik pada kelas yang lebih sering muncul. Namun, kelas Gizi Kurang dan Gizi Buruk memiliki nilai presisi dan recall mendekati 0, yang mengindikasikan bahwa model kesulitan mendeteksi data dari kedua kelas ini. F1-Score, yang menggabungkan presisi dan recall, menunjukkan nilai yang seimbang untuk Gizi Baik (0.75), tetapi sangat rendah untuk Gizi Kurang dan Gizi Buruk. Hasil macro average juga mencerminkan ketidakseimbangan data, dengan nilai presisi, recall, dan *F1-Score* yang rendah (masing-masing 0.27, 0.28, dan 0.27). Confusion matrix mengungkapkan bahwa banyak data dari kelas Gizi Kurang dan Gizi Buruk salah diprediksi sebagai Gizi Baik, menandakan perlunya perbaikan dalam mengatasi ketidakseimbangan kelas.

Jadi saran, Untuk meningkatkan kinerja model di masa mendatang, beberapa langkah dapat dilakukan. Pertama, peningkatan kualitas data sangat diperlukan, terutama dengan menambah jumlah data untuk kelas Gizi Kurang dan Gizi Buruk, sehingga model dapat belajar lebih banyak pola dari kelas-kelas minoritas ini. Selain itu, teknik penyeimbangan data seperti SMOTE (*Synthetic Minority Over-sampling Technique*) atau undersampling untuk mengurangi dominasi kelas mayoritas, dapat diterapkan untuk mengoptimalkan prediksi pada kelas yang lebih sedikit. Pengoptimalan hyperparameter, seperti jumlah pohon dan kedalaman pohon pada algoritma Random Forest, juga perlu dilakukan untuk meningkatkan akurasi model. Di samping itu, penelitian selanjutnya dapat mengeksplorasi penggunaan algoritma lain seperti XGBoost atau *Support*

Vector Machine (SVM) yang mungkin lebih efektif dalam menangani ketidakseimbangan kelas. Terakhir, melakukan validasi eksternal dengan data dari lokasi atau periode yang berbeda dapat memberikan gambaran yang lebih baik mengenai keandalan model dalam klasifikasi status gizi balita di wilayah lain. Dengan langkah-langkah perbaikan tersebut, penelitian selanjutnya dapat menghasilkan model yang lebih akurat dan andal dalam klasifikasi status gizi balita.

DAFTAR PUSTAKA

[1] Khumaeroh NF, Wahyani AD, Ratnasari D. 2022. Analisis Faktor-Faktor yang Mempengaruhi Status Gizi Kurang pada Balita Usia 3-5 Tahun di Wilayah Kerja Puskesmas Kersana. *Jurnal Ilmu Gizi dan Kesehatan*. 3(02): 71–75.

[2] Timur CJ, Irianto SE, Rahayu D. 2023. Analisis Faktor yang Mempengaruhi Status Gizi pada Balita di Kabupaten Lampung Utara. *JPKM (Jurnal Profesi Kesehatan Masyarakat)*. 4(2): 85–93.

[3] Azzahra F, Suarna N, Wijaya YA. 2024. Penerapan Algoritma Random Forest dan Cross Validation untuk Prediksi Data Stunting. *Jurnal Ilmiah Manajemen Informatika dan Komputer*. 8(01): 1–6.

[4] Aprilia YN, Sani DA, Anggadimas NM. 2024. Klasifikasi Status Penderita Gizi Stunting Pada Balita Menggunakan Metode Random Forest (Studi Kasus di Kelurahan Petamanan Kota Pasuruan). *Journal of Information Technology*. 9(2): 143–154.

[5] Perdana AY, Latuconsina R, Dinimaharawati A. 2021. Prediksi Stunting pada Balita dengan Algoritma Random Forest. *e-Proceeding of Engineering*. 8(5): 6650-6655.

[6] Dede Y, Manongga SP, Romeo P. 2023. Faktor yang Mempengaruhi Kejadian Gizi Kurang pada Anak Balita di Wilayah Kerja Puskesmas Kanatang Kabupaten Sumba Timur. *Jurnal Kesehatan Tambusai*. 4(3): 1998–3010.

[7] Nissa II. 2024. Klasifikasi Status Gizi Balita Menggunakan Algoritma Random Forest dan Feature Selection Relief-F. *Journal of Information Technology*. 2(3): 66–74.

[8] Lewoema SLZ, Prasetyaningrum PT. 2024. Implementasi Data Mining Pada Klasifikasi Status Gizi Bayi Dengan Metode Decision Tree CHAID (Studi Kasus: Puskesmas Godean 1 Yogyakarta). *Journal of Information Technology Ampera*. 5(1): 66–74, doi: 10.51519/journalita.v5i1.538.

[9] Alpiansah AB, Ramdhani Y. 2023. Optimasi Fitur dengan Forward Selection pada Estimasi Tingkat Obesitas menggunakan Random Forest. *Jurnal Sistem Informasi*. 12(3): 860–873.

[10] Handayani P, Fauzan AC, Harliana. 2024. Machine Learning Klasifikasi Status Gizi Balita Menggunakan Algoritma Random Forest. *Kajian Ilmiah Informatika dan Komputer*. 4(6): 3064–3072, doi: 10.30865/klik.v4i6.1909.