

**KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN
ALGORITMA RANDOM FOREST DI DESA MLESE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ALIVIA DYATAMA ARISUSANTO

21.11.4123

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN ALGORITMA RANDOM FOREST DI DESA MLESE

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



disusun oleh

ALIVIA DYATAMA ARISUSANTO

21.11.4123

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN ALGORITMA
RANDOM FOREST DI DESA MLESE**

yang disusun dan diajukan oleh

Alivia Dyatama Arisusanto

21.11.4123

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Februari 2026

Dosen Pembimbing,



Anna Baita, S.Kom., M.Kom.
NIK. 190302290

HALAMAN PENGESAHAN
SKRIPSI
KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN ALGORITMA
RANDOM FOREST DI DESA MLESE

yang disusun dan diajukan oleh

Alivia Dyatama Arisusanto

21.11.4123

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Februari 2026

Susunan Dewan Penguji

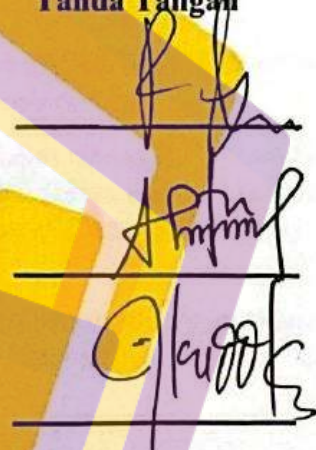
Nama Penguji

Tanda Tangan

Theopilus Bavu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Acihmah Sidauruk, S.Kom., M.Kom.
NIK. 190302238

Anna Baita, S.Kom., M.Kom.
NIK. 190302290



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Februari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Alivia Dyatama Arisusanto
NIM : 21.11.4123

Menyatakan bahwa Skripsi dengan judul berikut:

Klasifikasi Status Gizi Balita Menggunakan Algoritma Random Forest di Desa Mlese

Dosen Pembimbing : Anna Baita, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Februari 2026

Yang Menyatakan,



Alivia Dyatama Arisusanto

HALAMAN PERSEMBAHAN

Lembar persembahan ini merupakan ungkapan penghargaan dan rasa syukur penulis atas terselesaikannya skripsi ini. Dengan mengucap *Alhamdulillah Rabbil 'Alamin*, skripsi ini penulis persembahkan kepada :

Allah SWT yang telah memberikan rahmat, kesehatan, kemudahan, dan kekuatan sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

Ayah tercinta, almarhum Agus Susanto, yang menjadi panutan dan teladan dalam kehidupan penulis. Meskipun Ayah telah lebih dahulu berpulang, kasih sayang, doa, dan nilai-nilai kehidupan yang ditanamkan senantiasa menjadi pegangan dalam setiap langkah perjuangan penulis.

Ibunda tercinta, Ari Handayani, yang selalu melangitkan doa, memberikan dukungan, serta pengorbanan yang tiada henti hingga penulis mampu menyelesaikan pendidikan ini. Terima kasih atas ketulusan dan kekuatan yang selalu diberikan.

Adik tersayang, Anatasya Mischa Arisusanto, serta seluruh keluarga besar yang senantiasa memberikan doa, semangat, dan dukungan dalam setiap proses yang dilalui.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Status Gizi Balita Menggunakan Algoritma Random Forest di Desa Mlese” dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Dalam proses penyusunan skripsi ini, penulis memperoleh banyak bantuan, bimbingan, arahan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT, atas segala rahmat, karunia, kesehatan, dan kekuatan yang diberikan sehingga penulis dapat melalui setiap proses dan menyelesaikan skripsi ini dengan baik.
2. Prof. Dr. M. Suyanto, M.M. Selaku rektor Universitas Amikom Yogyakarta yang telah memberikan kesempatan kepada penulis untuk melaksanakan studi di Universitas Amikom Yogyakarta.
3. Prof. Dr. Kusrini, S.Kom., M.Kom. Selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Ibu Anna Baita, S.Kom, M.Kom. Selaku Dosen Pembimbing yang tidak hanya memberikan bimbingan secara akademik, tetapi juga menjadi pengarah, penyemangat, dan motivator bagi penulis selama proses penyusunan skripsi ini. Penulis menyampaikan terima kasih yang sebesar-besarnya atas waktu, kesabaran, perhatian, serta arahan yang telah diberikan. Terima kasih atas kemudahan dalam setiap proses, keterbukaan dalam konsultasi, serta pemahaman terhadap berbagai kendala yang penulis hadapi. Semoga Allah SWT senantiasa melimpahkan kesehatan, keberkahan, dan balasan terbaik atas segala kebaikan yang telah diberikan. Tim Dosen Penguji yang telah

memberikan saran dan masukan yang membangun demi penyempurnaan skripsi ini.

5. Tim Dosen Penguji yang telah memberikan saran dan masukan yang konstruktif demi penyempurnaan skripsi ini.
6. Seluruh Bapak dan Ibu Dosen Program Studi Informatika Universitas Amikom Yogyakarta yang telah memberikan ilmu pengetahuan, wawasan, pengalaman, serta bimbingan akademik selama masa perkuliahan sehingga penulis memperoleh bekal yang memadai dalam menyusun dan menyelesaikan skripsi ini.
7. Pihak Desa Mlese, tenaga kesehatan, serta instansi terkait yang telah memberikan izin, dukungan, dan kerja sama dalam pelaksanaan penelitian serta proses pengumpulan data sehingga penelitian ini dapat terlaksana dengan baik.
8. Ayah tercinta, almarhum Agus Susanto, yang menjadi panutan dan sosok istimewa dalam kehidupan penulis. Walaupun telah lebih dahulu berpulang dan tidak dapat menemani setiap proses yang penulis jalani dalam penyelesaian pendidikan ini, kasih sayang, didikan, serta doa yang beliau berikan semasa hidupnya senantiasa hidup dan menjadi sumber kekuatan bagi penulis dalam menyelesaikan pendidikan ini. Tanggung jawab dan semangat yang beliau tanamkan terus menjadi pedoman dalam setiap langkah penulis. Semoga Allah SWT melimpahkan rahmat dan menempatkan beliau di tempat terbaik di sisi-Nya.
9. Ibunda tercinta, Ari Handayani, sosok istimewa dalam kehidupan penulis, satu-satunya orang tua yang penulis miliki saat ini. Beliau senantiasa mendampingi penulis dalam setiap kondisi, serta menjadi sumber kekuatan yang tidak pernah padam. Beliau merupakan perempuan tangguh, yang tidak hanya melahirkan penulis, tetapi juga menjadi tulang punggung keluarga dan menjalankan dua peran sekaligus dengan penuh keikhlasan. Terima kasih atas setiap pengorbanan, semangat, dan ridho yang tulus. Doa-doa beliau menjadi nafas

perjuangan penulis dalam menyelesaikan skripsi ini. Semoga Allah SWT senantiasa melimpahkan kesehatan dan umur panjang.

10. Adik tersayang, Anatasya Mischa Arisusanto, yang senantiasa memberikan semangat, doa, dan dukungan kepada penulis dalam setiap proses penyelesaian skripsi ini. Kehadiran dan perhatian yang diberikan menjadi motivasi tersendiri bagi penulis untuk terus berusaha dan menyelesaikan pendidikan dengan baik.
11. Seluruh keluarga tercinta kakung, uti, bulik, om, dan keponakan penulis yang selalu memberikan doa dan menyayangi penulis dengan tulus.
12. Sahabat yang penulis anggap sebagai saudara, yaitu Gigih, Marthatia, dan Riko, yang senantiasa memberikan dukungan, motivasi, serta semangat kepada penulis dalam setiap proses penyelesaian skripsi ini. Kebersamaan, perhatian, dan dukungan yang diberikan menjadi penguat bagi penulis dalam menghadapi berbagai tantangan selama masa perkuliahan. Penulis mengucapkan terima kasih atas segala doa, motivasi, dan kebersamaan yang telah diberikan.
13. Teruntuk sahabat seperjuangan, Luthfia Ridho Damayanti, yang telah mendampingi penulis sejak awal perkuliahan hingga tahap akhir penyusunan skripsi ini. Terima kasih atas kebersamaan, kesabaran, dukungan, dan bantuan yang senantiasa hadir dalam setiap situasi, baik suka maupun duka. Pertemanan ini merupakan salah satu bagian terindah yang pernah penulis rasakan selama masa perkuliahan. Setiap langkah yang dilalui bersama menjadi bagian berharga dalam perjalanan akademik penulis dan memberikan kekuatan dalam menyelesaikan pendidikan ini.
14. Sahabat sejak bangku SMK, Anisa Ayu Ikhsani, yang telah kebersamai penulis selama delapan tahun perjalanan kehidupan dan pendidikan. Terima kasih atas kesetiaan, doa, serta motivasi yang senantiasa diberikan dalam setiap fase, baik dalam kebahagiaan maupun tantangan. Kehadiran dan dukungan yang tulus menjadi salah satu

penguat bagi penulis untuk terus bertahan dan menyelesaikan pendidikan ini.

15. Kepada penulis, Alivia Dyatama Arisusanto, yang telah bertahan hingga saat ini meskipun sempat merasa ragu terhadap kemampuan diri. Penulis menyadari bahwa setiap langkah kecil merupakan bagian dari perjalanan yang memerlukan ketekunan, kesabaran, dan tekad. Berbagai kendala, keraguan, dan kelelahan yang sempat menghambat proses penyelesaian skripsi ini justru menjadi dorongan untuk terus berusaha hingga studi ini dapat diselesaikan. Terima kasih telah berjuang sejauh ini dan tetap memilih untuk tidak menyerah. Pencapaian ini patut disyukuri sebagai bagian dari proses menuju langkah yang lebih baik ke depan. Berbahagialah selalu apapun kekurangan dan kelebihanmu mari tetap berjuang untuk kedepan.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan dan jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya di bidang informatika dan kesehatan masyarakat.

Yogyakarta, 23 Februari 2026

Penulis



(Alivia Dyatama Arisusanto)

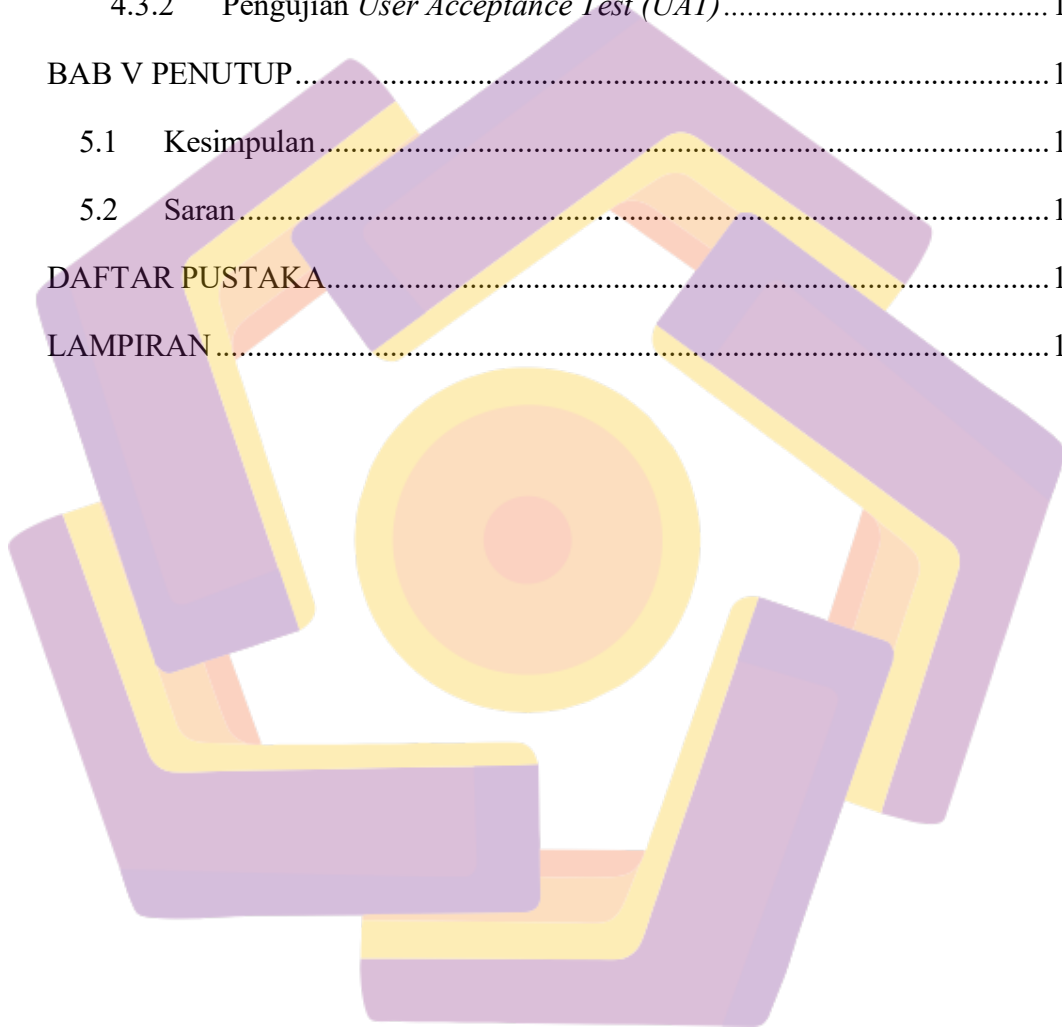
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xvii
DAFTAR LAMBANG DAN SINGKATAN.....	xviii
DAFTAR ISTILAH	xx
INTISARI.....	xxvi
ABSTRACT	xxvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.5.1 Manfaat Praktis	5
1.5.2 Manfaat Teoritis	5

1.6	Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA		7
2.1	Studi Literatur	7
2.2	Dasar Teori	12
2.2.1	Gizi dan Status Gizi Balita	12
2.2.2	<i>Data Mining</i>	14
2.2.3	<i>Machine Learning</i>	15
2.2.4	Klasifikasi	16
2.2.5	<i>Exploratory Data Analysis (EDA)</i>	16
2.2.6	<i>Pre-Processing Data</i>	16
2.2.7	<i>Feature Selection</i>	18
2.2.8	<i>Split Data</i>	19
2.2.9	Penanganan Ketidakseimbangan Kelas	20
2.2.10	Algoritma <i>Random Forest</i>	23
2.2.11	Evaluasi Model	26
2.2.12	Evaluasi <i>User Acceptance Test (UAT)</i>	29
2.2.13	Optimasi <i>Hyperparameter Tuning</i>	31
2.2.14	<i>Confusion Matrix</i>	31
2.2.15	<i>Deployment</i>	32
BAB III METODE PENELITIAN		33
3.1	Objek Penelitian	33
3.2	Alur Penelitian	34
3.2.1	Pengumpulan Dataset	35
3.2.2	Loading Dataset	35
3.2.3	Exploratory Data Analysis (EDA)	38

3.2.4	<i>Pre-Processing Data</i>	40
3.2.5	<i>Feature Selection</i>	47
3.2.6	<i>Split Data</i>	50
3.2.7	Penanganan Ketidakseimbangan Kelas	51
3.2.8	<i>Modeling</i>	53
3.2.9	<i>Optimasi Hyperparameter Tuning</i>	54
3.2.10	<i>Deployment</i>	55
3.3	Alat dan Bahan	56
3.3.1	Data Penelitian	56
3.3.2	Alat/instrumen	56
BAB IV HASIL DAN PEMBAHASAN		58
4.1	Hasil	58
4.1.1	<i>Profile Dataset</i>	58
4.1.2	<i>Karakteristik Dataset</i>	59
4.1.3	<i>Exploratory Data Analysis (EDA)</i>	60
4.1.4	<i>Pre-Processing Data</i>	71
4.1.5	<i>Feature Selection</i>	81
4.1.6	<i>Split Data</i>	83
4.1.7	Penanganan Ketidakseimbangan Kelas	84
4.1.8	<i>Modeling</i>	85
4.1.9	<i>Optimasi Hyperparameter Tuning</i>	90
4.1.10	<i>Confusion Matrix</i>	92
4.2	<i>Deployment</i>	95
4.2.1	<i>Visual Studio Code (VS Code)</i>	95
4.2.2	<i>GitHub</i>	95

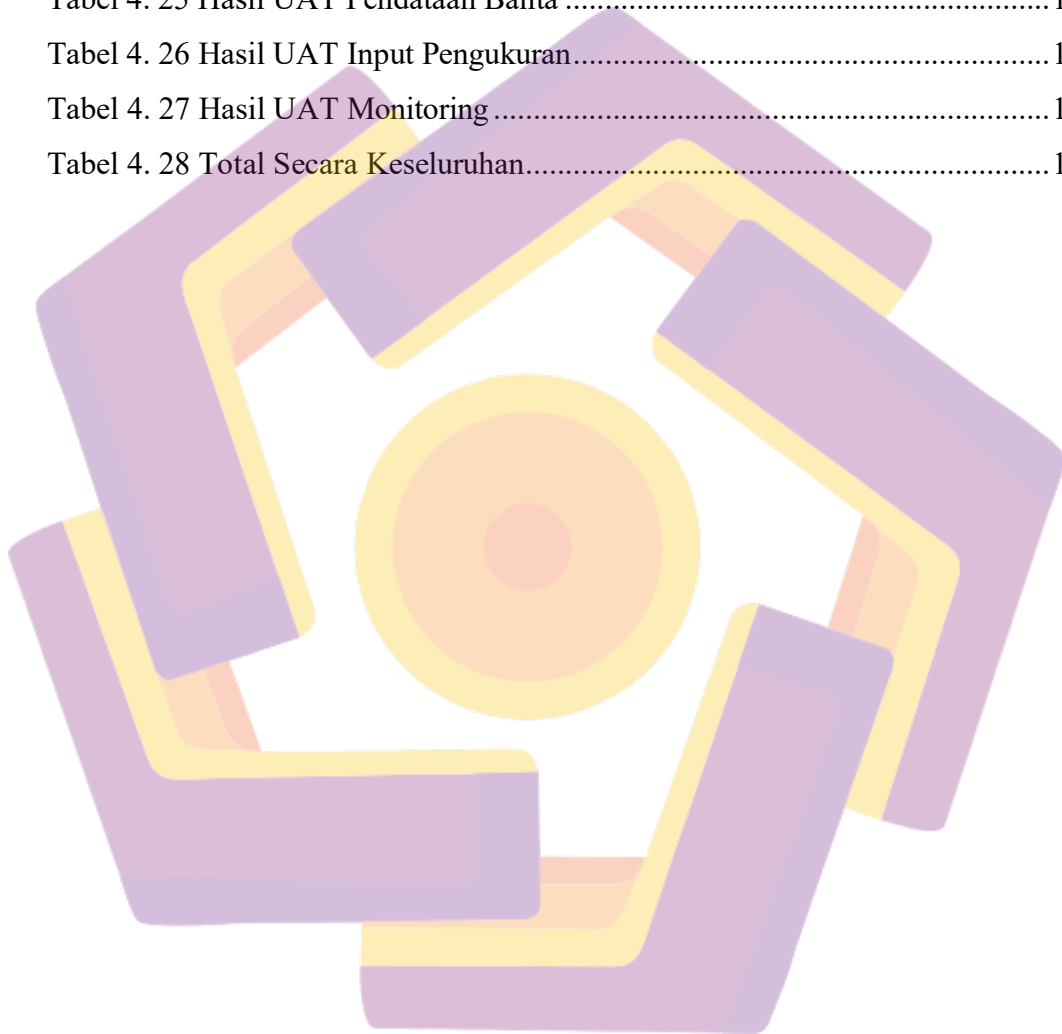
4.2.3	<i>Google Sheets dan Google Cloud Console</i>	96
4.2.4	<i>Streamlit Cloud</i>	96
4.2.5	Implementasi Tampilan Aplikasi	96
4.3	Pengujian Aplikasi.....	100
4.3.1	<i>Black Box Testing</i>	100
4.3.2	Pengujian <i>User Acceptance Test (UAT)</i>	108
BAB V PENUTUP.....		116
5.1	Kesimpulan.....	116
5.2	Saran.....	117
DAFTAR PUSTAKA.....		119
LAMPIRAN.....		126



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	8
Tabel 2. 2 Indeks Status Gizi Berdasarkan Z-Score	13
Tabel 2. 3 Kriteria Penilaian AUC	29
Tabel 3. 1 Variabel Dataset	37
Tabel 3. 2 Analisis Nilai Hilang	40
Tabel 3. 3 Encoding Data	46
Tabel 3. 4 Distribusi Proporsi Kelas	51
Tabel 3. 5 Distribusi Kelas Data Latih	51
Tabel 4. 1 Profile Dataset	58
Tabel 4. 2 Karakteristik Dataset	60
Tabel 4. 3 Statistik Deskriptif	60
Tabel 4. 4 Distribusi Jenis Kelamin	63
Tabel 4. 5 Distribusi Desa	65
Tabel 4. 6 Distribusi Dusun	66
Tabel 4. 7 Status Gizi BB/U	67
Tabel 4. 8 Status Gizi TB/U	68
Tabel 4. 9 Status Gizi BB/TB	69
Tabel 4. 10 Penanganan Missing Value	72
Tabel 4. 11 Feature Engineering	75
Tabel 4. 12 Konversi Tipe Data	76
Tabel 4. 13 Encoding Data	80
Tabel 4. 14 Metode dan Kriteria Feature Selection	82
Tabel 4. 15 Feature Final	83
Tabel 4. 16 Proporsi Data Latih dan Data Uji	84
Tabel 4. 17 Parameter Random Forest	86
Tabel 4. 18 Skenario Pemodelan Random Forest	87
Tabel 4. 19 Hasil Evaluasi Model Random Forest	88

Tabel 4. 20 Rentang Nilai Parameter Random Forest.....	90
Tabel 4. 21 Perbandingan Kinerja Model Random Forest.....	91
Tabel 4. 22 Hasil Black Box Testing	101
Tabel 4. 23 Profile Data Responden Pengujian.....	108
Tabel 4. 24 Hasil UAT Dashboard.....	109
Tabel 4. 25 Hasil UAT Pendataan Balita	110
Tabel 4. 26 Hasil UAT Input Pengukuran.....	111
Tabel 4. 27 Hasil UAT Monitoring	113
Tabel 4. 28 Total Secara Keseluruhan.....	114

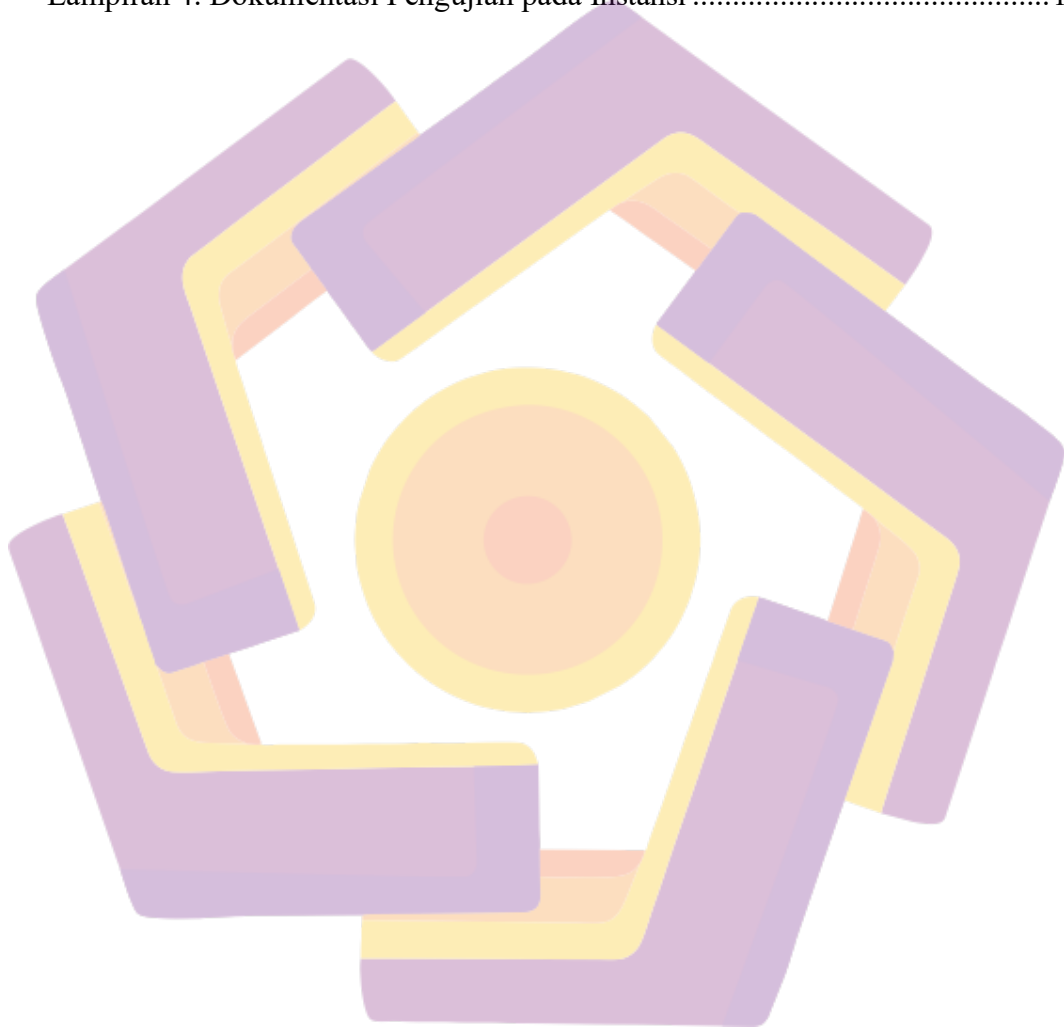


DAFTAR GAMBAR

Gambar 2. 1 Algoritma Random Forest	23
Gambar 3. 1 Alur Penelitian	34
Gambar 3. 2 Alur Deployment	55
Gambar 4. 1 Distribusi Variabel Numerik	62
Gambar 4. 2 Distribusi Jenis Kelamin	63
Gambar 4. 3 Distribusi Desa	64
Gambar 4. 4 Distribusi Dusun	65
Gambar 4. 5 Distribusi Status Gizi BB/U	66
Gambar 4. 6 Distribusi Status TB/U	67
Gambar 4. 7 Distribusi Status BB/TB	68
Gambar 4. 8 Korelasi Antar Fitur	70
Gambar 4. 9 Analisis Outlier	71
Gambar 4. 10 Sebelum Penanganan Outlier	73
Gambar 4. 11 Setelah Penanganan Outlier	74
Gambar 4. 12 Sebelum Normalisasi Data	77
Gambar 4. 13 Setelah Normalisasi Data	79
Gambar 4. 14 Proporsi Data Latih dan Data Uji	84
Gambar 4. 15 Distribusi Kelas Data Latih Sebelum dan	85
Gambar 4. 16 Confusion Matrix Sebelum Tuning	93
Gambar 4. 17 Confusion Matrix Sesudah Tuning	94
Gambar 4. 18 Tampilan Dashboard	97
Gambar 4. 19 Tampilan Pendataan Balita	98
Gambar 4. 20 Tampilan Input Pengukuran	99
Gambar 4. 21 Tampilan Monitoring	99

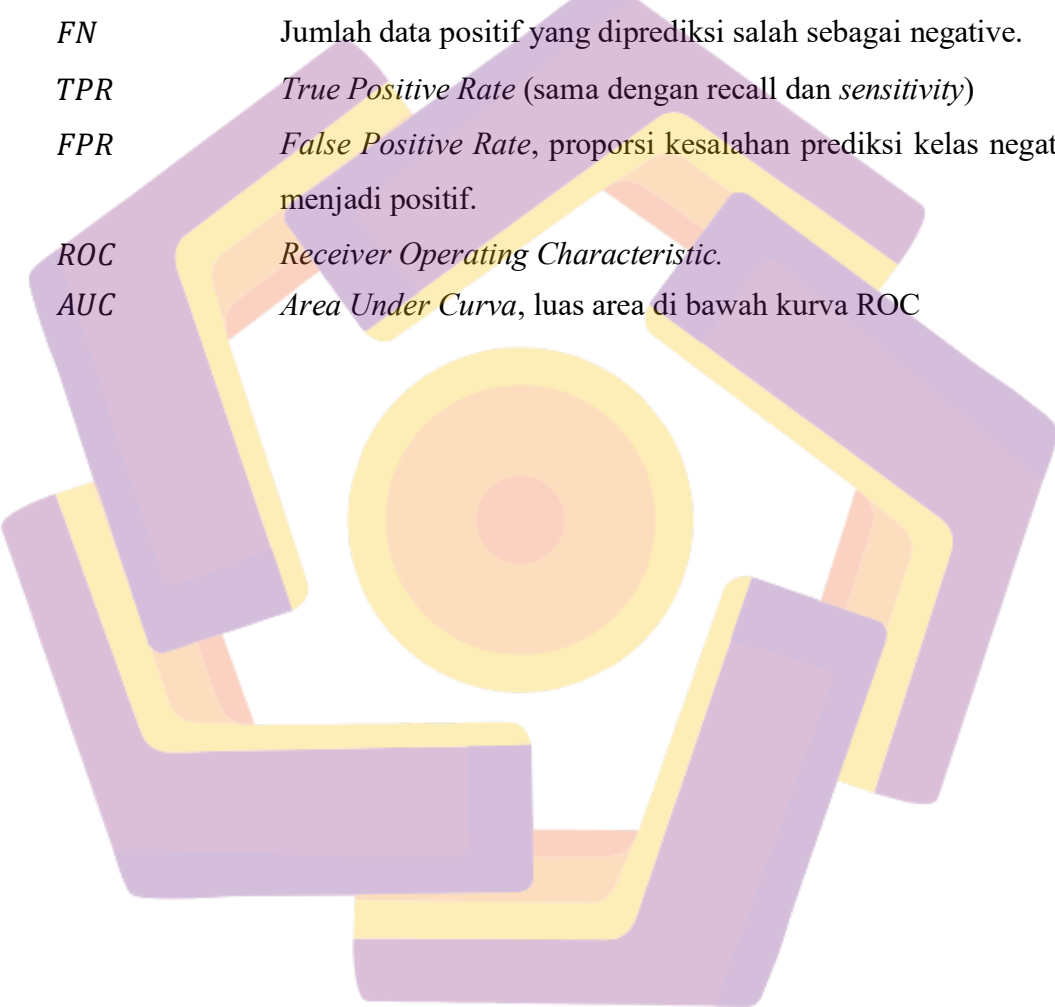
DAFTAR LAMPIRAN

Lampiran 1. Surat Ijin Penelitian.....	126
Lampiran 2. Surat Jawaban Persetujuan Penelitian.....	127
Lampiran 3. Dokumentasi Observasi Lapangan	128
Lampiran 4. Dokumentasi Pengujian pada Instansi	128



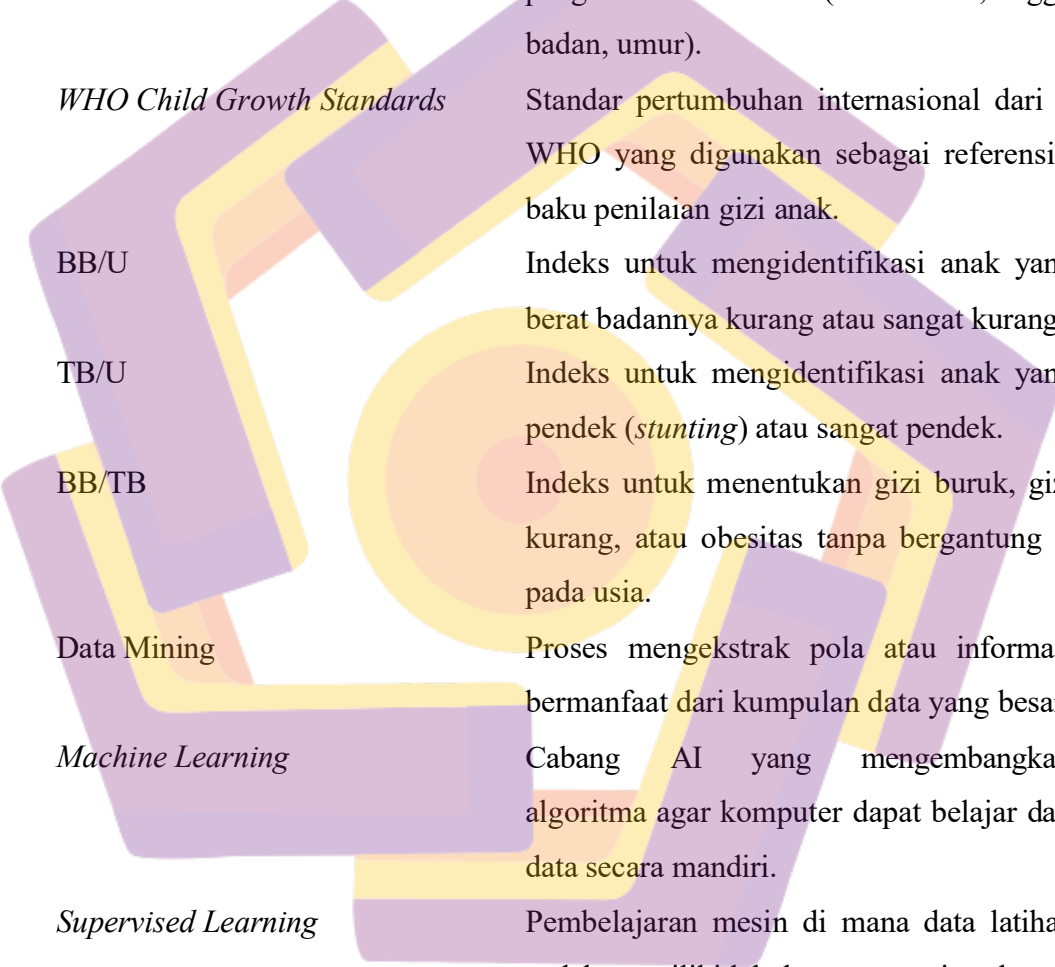
DAFTAR LAMBANG DAN SINGKATAN

Z	Nilai <i>Z-Score</i> .
x	Nilai individu (hasil pengukuran berat atau tinggi badan anak).
σ	Nilai simpangan baku (<i>standard deviation</i>) dari populasi rujukan.
IQR	<i>Interquartile Range</i> (selisih antara kuartil atas dan bawah).
Q_1	Kuartil pertama.
Q_2	Kuartil ketiga.
$Lower_Bound$	Batas bawah untuk menentukan data <i>outlier</i> .
$Upper_Bound$	Batas atas untuk menentukan data <i>outlier</i> .
x_{new}	Data sintetis baru yang dihasilkan algoritma.
x_i	Salah satu data dari kelas minoritas yang dipilih acak.
x_{zi}	Salah satu <i>k-nearest neighbors</i> dari x_i dalam kelas minoritas.
δ	Bilangan acak yang bernilai 0 dan 1.
r_i	Tingkat kesulitan klasifikasi lokal sampel minoritas ke- i .
Δ_i	Jumlah tetangga terdekat dari kelas minoritas.
k	Jumlah <i>k-nearest neighbors</i> .
g_i	Bobot proporsi sampel sintetis untuk sampel minoritas ke- i .
n	Jumlah total sampel kelas minoritas.
G_i	Jumlah sampel sintetis untuk sampel minoritas ke- i .
G	Total jumlah sampel sintetis yang dibangkitkan.
λ	Bilangan acak dalam rentang $[0,1]$.
$Entropy(S)$	Ukuran ketidakmurnian dalam himpunan data S
p_i	Proporsi kelas ke- i pada satu <i>node</i> .
S	Himpunan data pada satu <i>node</i> .
$Gain(S, A)$	Ukuran efektivitas atribut A dalam mengklasifikasi data.
S_v	Subset data dengan nilai atribut v .
$ S_v / S $	Proporsi jumlah data pada subset v terhadap total data.
$Gini(S)$	Ukuran keragaman atau tingkat kemurnian <i>node</i> .



$f_{RF}(X)$	Hasil prediksi akhir <i>Random Forest</i> .
$h_i(X)$	Hasil prediksi dari pohon ke-i.
<i>mode</i>	Nilai yang paling sering muncul (voting mayoritas).
<i>TP</i>	Jumlah data positif yang diprediksi benar.
<i>TN</i>	Jumlah data negative yang diprediksi benar.
<i>FP</i>	Jumlah data negative yang diprediksi salah sebagai positif.
<i>FN</i>	Jumlah data positif yang diprediksi salah sebagai negative.
<i>TPR</i>	<i>True Positive Rate</i> (sama dengan recall dan <i>sensitivity</i>)
<i>FPR</i>	<i>False Positive Rate</i> , proporsi kesalahan prediksi kelas negatif menjadi positif.
<i>ROC</i>	<i>Receiver Operating Characteristic</i> .
<i>AUC</i>	<i>Area Under Curva</i> , luas area di bawah kurva ROC

DAFTAR ISTILAH



Status Gizi	Keadaan tubuh yang mencerminkan tingkat konsumsi zat gizi dan penggunaan zat gizi tersebut dalam tubuh.
Antropometri	Metode penilaian status gizi berdasarkan pengukuran fisik tubuh (berat badan, tinggi badan, umur).
<i>WHO Child Growth Standards</i>	Standar pertumbuhan internasional dari WHO yang digunakan sebagai referensi baku penilaian gizi anak.
BB/U	Indeks untuk mengidentifikasi anak yang berat badannya kurang atau sangat kurang.
TB/U	Indeks untuk mengidentifikasi anak yang pendek (<i>stunting</i>) atau sangat pendek.
BB/TB	Indeks untuk menentukan gizi buruk, gizi kurang, atau obesitas tanpa bergantung pada usia.
Data Mining	Proses mengekstrak pola atau informasi bermanfaat dari kumpulan data yang besar.
<i>Machine Learning</i>	Cabang AI yang mengembangkan algoritma agar komputer dapat belajar dari data secara mandiri.
<i>Supervised Learning</i>	Pembelajaran mesin di mana data latihan sudah memiliki label atau target jawaban.
<i>Unsupervised Learning</i>	Pembelajaran mesin untuk menemukan pola pada data yang tidak memiliki label atau target.
Klasifikasi	Tugas machine learning untuk memprediksi kategori atau label diskret dari suatu data.

Random Forest

Algoritma yang terdiri dari kumpulan banyak pohon keputusan (*Decision Trees*) untuk meningkatkan akurasi.

SMOTE (Synthetic Minority Over-sampling Technique)

Teknik menciptakan data buatan (sintetis) untuk memperbanyak jumlah kelas minoritas.

Tomek Links

Metode untuk membersihkan data dengan menghapus pasangan data yang tumpang tindih antar kelas agar batas keputusan lebih jelas.

ADASYN (Adaptive Synthetic Sampling)
Decision Tree

Metode menciptakan data baru di area yang dianggap sulit dipelajari oleh model.

Model prediksi berbentuk struktur pohon yang mewakili serangkaian aturan keputusan.

Preprocessing Data

Tahap pengolahan data mentah menjadi format yang bersih dan siap digunakan untuk pemodelan.

Exploratory Data Analysis (EDA)

Analisis awal untuk memahami karakteristik, pola, dan anomali dalam dataset.

Outlier

Data ekstrem yang nilainya sangat jauh berbeda dari sebagian besar data lainnya.

Feature Engineering

Proses memanipulasi atau menciptakan variabel baru agar model dapat bekerja lebih baik.

Feature Selection

Proses memilih variabel yang paling relevan untuk mengurangi kompleksitas model.

Hyperparameter Tuning

Proses mencari parameter terbaik untuk mengoptimalkan performa algoritma.

RandomizedSearchCV

Metode pencarian kombinasi parameter terbaik secara acak untuk efisiensi waktu.

Confusion Matrix

Tabel evaluasi yang menunjukkan perbandingan antara nilai aktual dan hasil prediksi model.

User Acceptance Test (UAT)

Pengujian untuk memastikan sistem sudah sesuai dengan kebutuhan dan harapan pengguna.

Dataset

Kumpulan data yang digunakan sebagai objek dalam penelitian.

Deployment

Tahap peluncuran atau penerapan model/sistem ke lingkungan yang dapat diakses pengguna.

e-PPBGM

Aplikasi elektronik Pencatatan dan Pelaporan Gizi Berbasis Masyarakat yang digunakan oleh instansi kesehatan.

Overfitting

Kondisi di mana model terlalu cerdas menghafal data latihan tapi gagal memprediksi data baru.

Overlapping

Kondisi di mana data dari kelas yang berbeda berada di posisi yang sangat berdekatan atau bertumpuk.

Node

Titik dalam pohon keputusan yang mewakili pengujian pada atribut atau hasil keputusan.

Redundan

Data atau variabel yang berulang dan tidak memberikan informasi baru yang unik.

Generalisasi (Robustness)

Kemampuan model untuk tetap akurat saat diberikan data baru yang belum pernah dilihat.

Imputasi

Teknik mengisi nilai yang hilang (*missing values*) dengan nilai pengganti (seperti median atau mean).

Encoding

Proses mengubah data kategori (teks) menjadi angka agar dapat diproses algoritma.

Data Transformation

Mengubah format atau distribusi data (seperti normalisasi) agar sesuai untuk analisis.

Class Imbalance

Kondisi di mana jumlah sampel antar kelas target tidak seimbang secara drastis.

Multiclass

Klasifikasi yang memiliki lebih dari dua kategori target (seperti 6 status gizi dalam penelitian Anda).

Data Leakage

Kesalahan di mana informasi dari luar dataset latihan digunakan untuk membuat model (menyebabkan hasil tes tidak valid).

Bagging & Bootstrapping

Teknik pengambilan sampel secara acak dengan pengembalian untuk membangun model *ensemble*.

Entropy & Gini Index

Rumus statistik untuk mengukur tingkat ketidakmurnian atau keragaman dalam suatu dataset.

API (Application Programming Interface)

Antarmuka yang memungkinkan dua aplikasi/sistem saling berkomunikasi secara otomatis.

Skewness

Ukuran statistik yang menunjukkan tingkat ketidaksamaan atau kemiringan distribusi data.

Linear positif

Hubungan di mana kenaikan satu variabel diikuti oleh kenaikan variabel lainnya secara searah.

Cross Validation

Metode pengujian performa model dengan membagi data menjadi beberapa bagian yang berbeda secara berulang.

Subset

Sekumpulan data yang diambil dari kumpulan data yang lebih besar (dataset utama).

Support Vector Machine (SVM)

Algoritma *machine learning* yang mencari bidang pemisah (*hyperplane*) terbaik untuk memisahkan dua atau lebih kelas data.

K-Nearest Neighbors (KNN)

Algoritma yang mengklasifikasikan data berdasarkan kemiripan pola dengan data terdekat di sekitarnya.

Noise

Data yang tidak relevan, salah, atau mengganggu yang dapat mengurangi akurasi hasil analisis.

Ensemble Learning

Teknik menggabungkan beberapa model pembelajaran untuk mendapatkan hasil prediksi yang lebih akurat

G-Mean

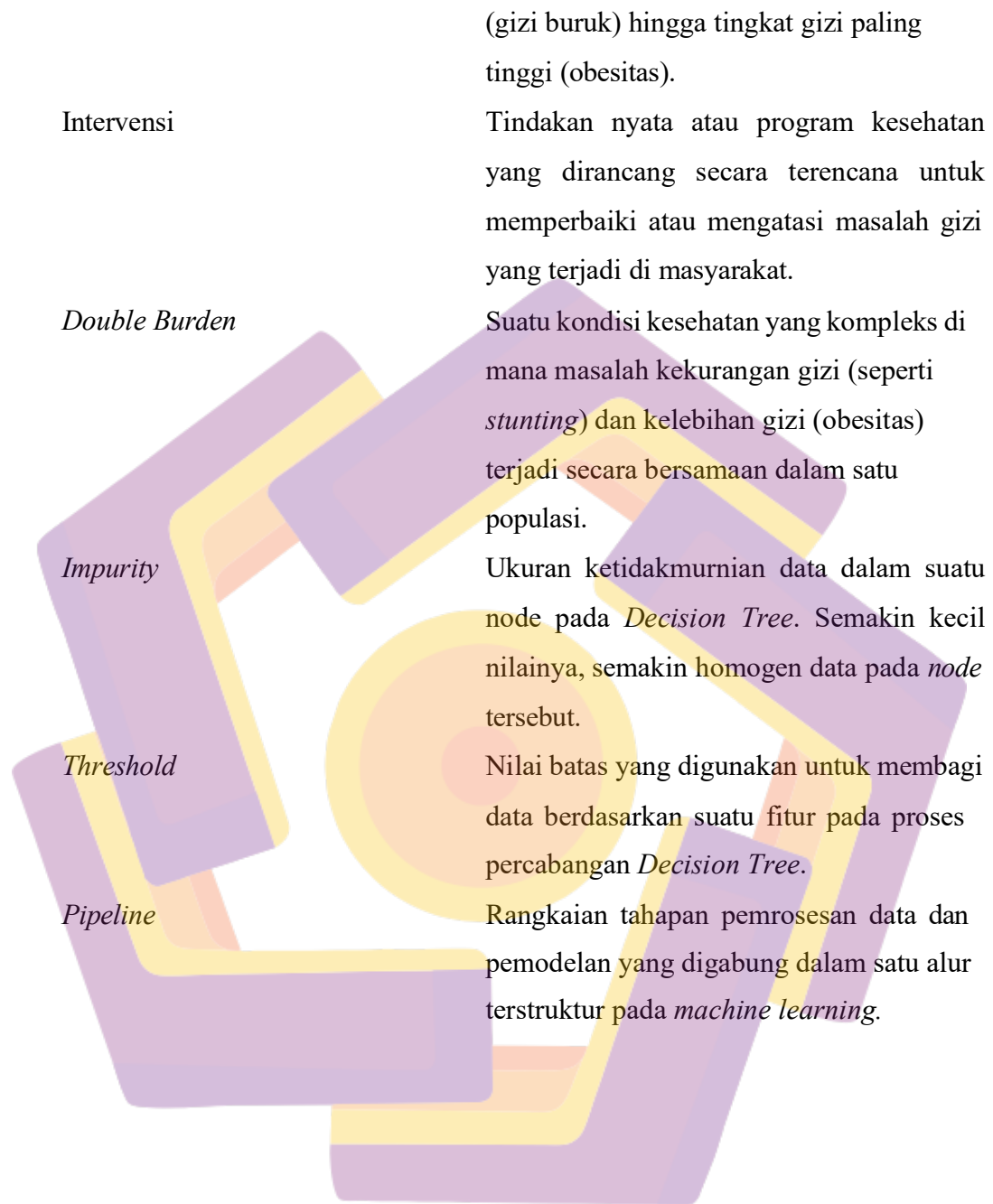
Metrik yang menghitung data tidak seimbang karena mengukur seberapa adil model memperlakukan kelas mayoritas dan minoritas.

Default

Nilai standar atau pengaturan awal yang sudah disediakan oleh sistem/pustaka pemrograman sebelum dilakukan perubahan atau optimasi.

Spektrum

Rentang variasi kondisi nutrisi seseorang yang mencakup tingkat gizi paling rendah



INTISARI

Masalah gizi pada balita di Indonesia, termasuk di Desa Mlese, merupakan tantangan signifikan yang berdampak pada gangguan tumbuh kembang dan penurunan kualitas sumber daya manusia di masa depan. Evaluasi status gizi konvensional di Desa Mlese saat ini masih memiliki keterbatasan administratif dan belum optimal dalam memprediksi pola data antropometri yang kompleks, sehingga diperlukan metode yang lebih responsif dan akurat. Penelitian ini menerapkan pendekatan *machine learning* dengan algoritma *Random Forest* untuk mengklasifikasikan status gizi balita melalui tahapan pengumpulan data sekunder periode 2020–2024, pra-pemrosesan data, serta seleksi fitur menggunakan metode *Random Forest Importance*, *Chi-Square*, dan *Recursive Feature Elimination (RFE)*. Untuk mengatasi ketidakseimbangan data, penelitian ini membandingkan teknik *SMOTE*, *SMOTE-TomekLinks*, dan *ADASYN*, yang dioptimasi menggunakan *hyperparameter tuning* melalui *Randomized Search Cross Validation*. Hasil penelitian menunjukkan bahwa integrasi *ADASYN* dengan *Random Forest* memberikan performa terbaik dengan akurasi 0,8757 dan F1-Score 0,6898. Sistem yang dikembangkan tidak hanya berfungsi sebagai media pencatatan data, tetapi juga menonjolkan analisis frekuensi kunjungan balita untuk membantu pemantauan kepatuhan kunjungan posyandu, menyediakan visualisasi data yang lebih analitik untuk melihat tren pertumbuhan dan distribusi status gizi secara individu maupun populasi, serta mendukung deteksi dini risiko masalah gizi berbasis *machine learning*. Berdasarkan pengujian *User Acceptance Test (UAT)* dengan skor 88,17%, sistem ini dinyatakan sangat layak untuk membantu kader posyandu dan pihak puskesmas dalam melakukan pemantauan berkelanjutan serta perencanaan intervensi gizi yang lebih tepat sasaran.

Kata kunci: Status Gizi Balita, Random Forest, ADASYN, Machine Learning, Klasifikasi.

ABSTRACT

Nutritional problems among toddlers in Indonesia, including in Mlese Village, represent a significant challenge that impacts growth disturbances and the decline of human resource quality in the future. Conventional nutritional status evaluation in Mlese Village currently faces administrative limitations and is not yet optimal in predicting complex anthropometric data patterns, necessitating a more responsive and accurate method. This study applies a machine learning approach using the Random Forest algorithm to classify the nutritional status of toddlers through stages of secondary data collection for the 2020–2024 period, data preprocessing, and feature selection using Random Forest Importance, Chi-Square, and Recursive Feature Elimination (RFE) methods. To address data imbalance, this research compares SMOTE, SMOTE-TomekLinks, and ADASYN techniques, optimized through hyperparameter tuning via Randomized Search Cross-Validation. The results show that the integration of ADASYN with Random Forest provides the best performance with an accuracy of 0.8757 and an F1-Score of 0.6898. The developed system functions not only as a data recording medium but also emphasizes toddler visit frequency analysis to assist in monitoring Posyandu attendance compliance, provides analytical data visualization to observe growth trends and nutritional status distribution individually and across the population, and supports early detection of nutritional risk based on machine learning. Based on the User Acceptance Test (UAT) with a score of 88.17%, this system is declared highly feasible for assisting Posyandu cadres and Puskesmas authorities in conducting continuous monitoring and planning more targeted nutritional interventions.

Keywords: *Toddler Nutritional Status, Random Forest, ADASYN, Machine Learning, Classification.*