

**ANALISIS KOMPARATIF AKURASI PREDIKSI STUNTING PADA
ANAK BERDASARKAN FAKTOR GIZI TERHADAP IBU HAMIL DAN
FAKTOR SOSIAL KELUARGA MENGGUNAKAN REGRESI RANDOM
FOREST DAN GRADIENT BOOSTING REGRESSION**

SKRIPSI

Diajukan untuk memenuhi salah satu persyaratan memperoleh gelar Sarjana pada
Program Studi Informatika



disusun oleh:

MUHAMMAD MAULANA ANTARIKSA

21.61.0212

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS KOMPARATIF AKURASI PREDIKSI STUNTING PADA
ANAK BERDASARKAN FAKTOR GIZI TERHADAP IBU HAMIL DAN
FAKTOR SOSIAL KELUARGA MENGGUNAKAN REGRESI RANDOM
FOREST DAN GRADIENT BOOSTING REGRESSION**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

MUHAMMAD MAULANA ANTARIKSA
21.61.0212

Kepada

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS KOMPARATIF AKURASI PREDIKSI STUNTING PADA ANAK BERDASARKAN FAKTOR GIZI TERHADAP IBU HAMIL DAN FAKTOR SOSIAL KELUARGA MENGGUNAKAN REGRESI RANDOM FOREST DAN GRADIENT BOOSTING REGRESSION

yang disusun dan diajukan oleh

Muhammad Maulana Antariksa

21.61.0212

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 28 mei 2025

Dosen Pembimbing,


Ferjan Fauzi Abdulloh, M.Kom

NIK. 190302276

HALAMAN PENGESAHAN**SKRIPSI****ANALISIS KOMPARATIF AKURASI PREDIKSI STUNTING PADA
ANAK BERDASARKAN FAKTOR GIZI TERHADAP IBU HAMIL DAN
FAKTOR SOSIAL KELUARGA MENGGUNAKAN REGRESI RANDOM
FOREST DAN GRADIENT BOOSTING REGRESSION**

yang disusun dan diajukan oleh

Muhammad Maulana Antariksa

21.61.0212

Telah dipertahankan di depan Dewan Penguji
pada tanggal 28 mei 2025

Susunan Dewan Penguji**Nama Penguji**

Ali Mustopa, S.Kom., M.Kom.
NIK. 190302192

Nafiatun Sholihah, S.Kom., M.Cs
NIK. 190302524

Ferian Fauzi Abdulloh, S.Kom., M.Kom.
NIK. 190302276

Tanda Tangan

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 28 mei 2025

DEKAN FAKULTAS ILMU KOMPUTER

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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Maulana Antariksa
NIM : 21.61.0212

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS KOMPARATIF AKURASI PREDIKSI STUNTING PADA ANAK BERDASARKAN FAKTOR GIZI TERHADAP IBU HAMIL DAN FAKTOR SOSIAL KELUARGA MENGGUNAKAN REGRESI RANDOM FOREST DAN GRADIENT BOOSTING REGRESSION

Dosen Pembimbing : Ferian Fauzi Abdulloh, 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, 28 Mei 2025

Yang Menyatakan,



Muhammad Maulana Antariksa

HALAMAN PERSEMBAHAN

Puji dan syukur kehadiran Allah SWT atas berkat dan rahmat -Nya sehingga saya bisa menyelesaikan Laporan ini dengan baik. Dengan segala kerendahan hati, saya persembahkan skripsi ini kepada:

1. Orang tua saya tercinta, yang selalu memberikan cinta, doa, dan dukungan tanpa henti.
2. Bapak Ferian Fauzi Abdulloh, M.kom yang telah membimbing saya dengan sabar dan penuh perhatian.
3. Teman – teman seperjuangan yang selalu memberikan semangat dan motivasi. Teman-teman kelas 21 BCI 01 yang selalu membantu dari semester awal sampai semester akhir ini
4. Seluruh pihak yang telah membantu, baik yang saya kenal maupun tidak, serta yang telah berkontribusi secara langsung maupun tidak dalam penelitian ini. Dukungan kalian sangat berarti bagi saya dalam menyelesaikan tugas akademik ini.
5. Kepada teman teman saya yang berada di luar negeri, mas ilham dan mas marsal dan yang lain. Terimakasih telah membantu dan memberikan motivasi saya untuk selalu semangat mengejar mimpi.

Semoga karya ini menjadi bukti rasa terima kasih dan penghargaan saya kepada kalian semua.

KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Tuhan Yang Maha Esa karena atas rahmat dan karunia-Nya, saya dapat menyelesaikan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Studi Informatika di Universitas Amikom Yogyakarta

Dalam penyusunan skripsi ini, saya menyadari bahwa banyak pihak yang telah memberikan dukungan, bimbingan, dan bantuan sehingga skripsi ini dapat terselesaikan. Oleh karena itu, saya ingin mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. M.Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta
2. Bapak Ferian Fauzi Abdulloh, M.kom, selaku dosen pembimbing yang telah memberikan arahan, bimbingan, dan motivasi selama proses penyusunan skripsi ini.
3. Keluarga tercinta yang selalu memberikan dukungan moral dan doa yang tiada henti.
4. Bapak Agung Riyadi dan Ibu Retno Dwi Lesteri, selaku orang tua penulis, telah memberikan dukungan, doa, kasih sayang, motivasi, serta memfasilitasi segala keperluan penulis.

Yogyakarta, 28 mei 2025

Muhammad Maulana Antariksa

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	6
BAB II	9
TINJAUAN PUSTAKA	9

2.1	Studi Literatur	9
2.2	Dasar Teori.....	14
BAB III		21
METODE PENELITIAN.....		21
3.1	Dataset	21
3.2	Alur Penelitian	25
3.6	Alat dan Bahan	33
3.6.1	Alat/instrumen.....	33
3.6.2	Bahan.....	33
BAB IV		35
HASIL DAN PEMBAHASAN		35
4.1	Pembersihan data.....	35
4.2	Exploratory Data Analysis (EDA)	47
4.3	Bivariate Analysis	51
4.4	Multivariate Analysis	55
4.5	Modeling.....	57
4.6	Evaluasi	60
4.7	Hasil Visualisasi Data	60
BAB V		63
PENUTUP.....		63
REFERENSI.....		65
LAMPIRAN		68

DAFTAR TABEL

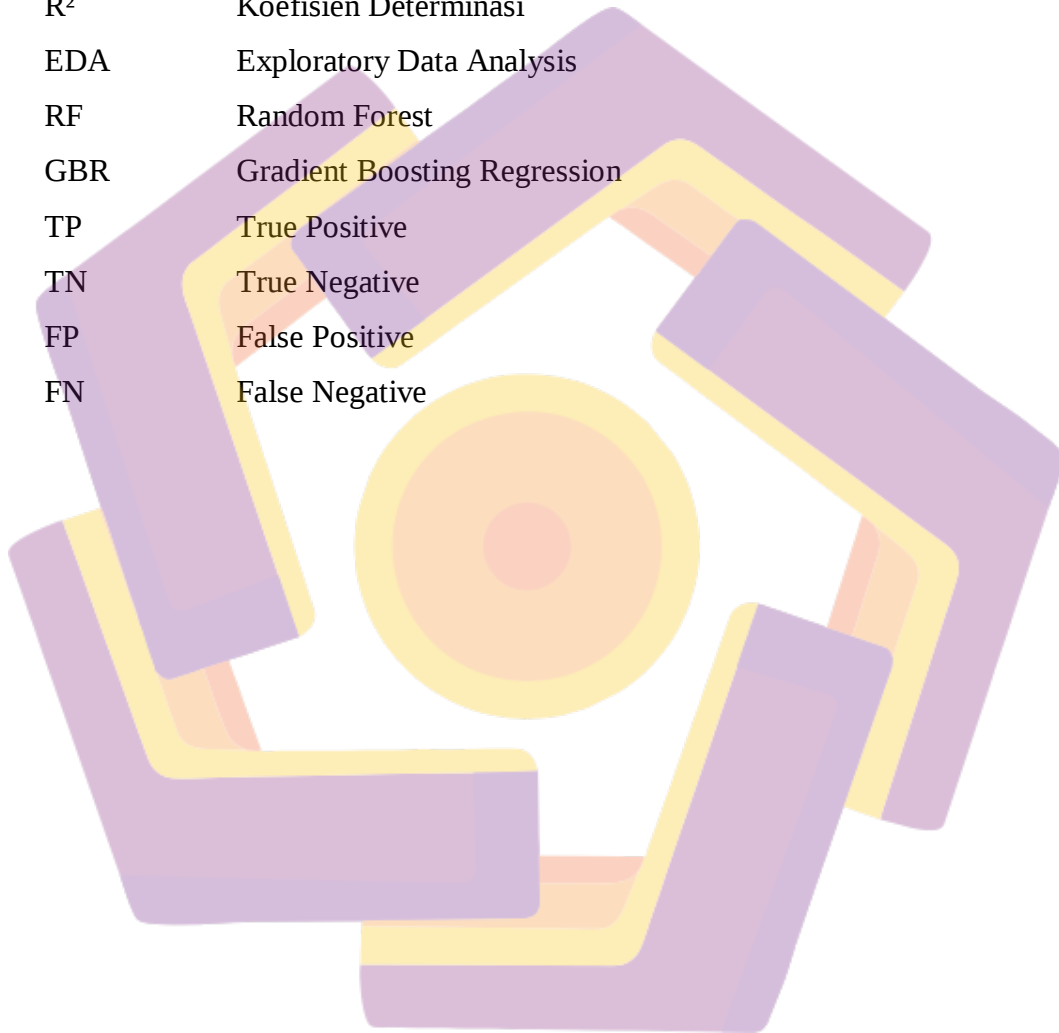
<i>Tabel 2. 1 Keaslian Penelitian</i>	<i>12</i>
<i>Tabel 3. 1 Tabel Featur Selection</i>	<i>21</i>
<i>Tabel 3. 2 Tabel Data Faktor Stunting</i>	<i>33</i>
<i>Tabel 4. 1 Tabel Informasi Data</i>	<i>36</i>
<i>Tabel 4. 2 Tabel Feature Selection.....</i>	<i>39</i>
<i>Tabel 4. 3 Tabel Informasi Data</i>	<i>44</i>
<i>Tabel 4. 4 Tabel Hasil Metrik Evaluasi</i>	<i>60</i>

DAFTAR GAMBAR

<i>Gambar 3. 1 Alur Penelitian.....</i>	25
<i>Gambar 3. 2 Heatmap Distribusi</i>	27
<i>Gambar 3. 3 Grafik Persebaran Data Residence</i>	28
<i>Gambar 3. 4 Grafik Persebaran Data Poverty Rating</i>	29
<i>Gambar 3. 5 Grafik Persebaran Data Maternal Education</i>	30
<i>Gambar 4. 1 Kode String Variabel Feature.....</i>	43
<i>Gambar 4. 2 Kode DataFrame Dataset.....</i>	43
<i>Gambar 4. 3 Kode Hasil Seleksi Fitur Indicator</i>	46
<i>Gambar 4. 4 Heatmap Distribusi Data</i>	48
<i>Gambar 4. 5 Grafik Persebaran Data Residence</i>	49
<i>Gambar 4. 6 Grafik Persebaran Poverty Rating</i>	50
<i>Gambar 4. 7 Grafik Persebaran Maternal Education.....</i>	51
<i>Gambar 4. 8 Kode Hubungan Indicator–Maternal Education.....</i>	52
<i>Gambar 4. 9 Heatmap Hubungan Indicator–Maternal Education.....</i>	52
<i>Gambar 4. 10 Heatmap Hubungan Indicator–Poverty Rating.....</i>	54
<i>Gambar 4. 11 Trend Stunting Indonesia.....</i>	55
<i>Gambar 4. 12 Trend Wasting Indonesia.....</i>	56
<i>Gambar 4. 13 Trend Overweight Indonesia</i>	56
<i>Gambar 4. 14 Trend Underweight Indonesia</i>	57
<i>Gambar 4. 15 Kode Data Setelah LabelEncoder.....</i>	58
<i>Gambar 4. 16 Kode Menentukan X dan Y</i>	58
<i>Gambar 4. 17 Seleksi Fitur</i>	58
<i>Gambar 4. 18 Kode Data Splitting.....</i>	59
<i>Gambar 4. 19 Hasil Standard Scaller</i>	59
<i>Gambar 4. 20 Visualisasi Predict value</i>	61
<i>Gambar 4. 21 Visualisasi Predict Value.....</i>	62

DAFTAR LAMBANG DAN SINGKATAN

MAE	Mean Absolute Error
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
R^2	Koefisien Determinasi
EDA	Exploratory Data Analysis
RF	Random Forest
GBR	Gradient Boosting Regression
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative



DAFTAR ISTILAH

<i>Stunting</i>	Kondisi gagal tumbuh pada anak akibat kekurangan gizi
<i>Nutrisi</i>	proses pengambilan dan penggunaan makanan oleh tubuh
<i>Preprocessing</i>	Pembersihan dan pemrosesan teks sebelum analisis
<i>Labeling</i>	Proses memberikan label atau kategori pada data
<i>Predictive Model</i>	sistem yang digunakan untuk memprediksi hasil berdasarkan data
<i>Validasi Model</i>	proses evaluasi kinerja model prediksi
<i>Regresi</i>	metode statistik untuk memprediksi nilai variabel dependen
<i>Overfitting</i>	kondisi model terlalu kompleks dan menangkap noise dalam data
<i>Fitur</i>	Karakteristik yang digunakan dalam analisis data
<i>Mean Absolute Error</i>	rata-rata selisih absolut antara nilai yang diprediksi dan nilai aktual
<i>Mean Squared Error</i>	rata-rata kuadrat selisih antara nilai yang diprediksi dan nilai aktual
<i>Root Mean Squared</i>	akar kuadrat dari MSE yang mengukur kesalahan prediksi dalam satuan yang sama dengan data asli
<i>Mikronutrien</i>	Nutrisi yang dibutuhkan oleh tubuh dalam jumlah kecil
<i>Ensemble learning</i>	Teknik pembelajaran mesin yang menggabungkan beberapa model untuk meningkatkan akurasi dan keandalan prediksi
<i>Standard Deviation</i>	Ukuran statistik yang digunakan untuk menentukan seberapa jauh tinggi badan anak dari rata-rata populasi sehat

INTISARI

Penelitian ini bertujuan untuk menganalisis dan memprediksi status stunting pada anak menggunakan teknik data mining. Dataset diproses terlebih dahulu dengan menangani nilai yang hilang, melakukan normalisasi fitur, dan mengubah variabel kategori menjadi format numerik. Beberapa algoritma seperti Random Forest dan Gradient Boosting diterapkan untuk menentukan model yang paling efektif. Evaluasi dilakukan menggunakan metrik Mean Absolute Error (MAE), Mean Squared Error (MSE), dan R-squared. Algoritma Random Forest dan Gradient Boosting menunjukkan hasil akurasi yang baik, yang menandakan kecocokannya dalam memprediksi risiko stunting. Faktor-faktor utama yang mempengaruhi stunting meliputi asupan gizi harian, status ekonomi keluarga, akses ke fasilitas kesehatan, dan pola pemberian makanan. Temuan ini memberikan wawasan penting untuk intervensi yang lebih tepat sasaran dalam mengurangi prevalensi stunting. Keterbatasan penelitian dan saran untuk penelitian selanjutnya juga dibahas.

Kata kunci: Stunting, Gizi anak, Data mining, Algoritma, Preprocessing data, Random Forest, Gradient Boosting.

ABSTRACT

This research aims to analyze and predict stunting status in children using data mining techniques. The dataset is preprocessed by treating missing values, performing feature normalization, and converting categorical variables into numerical format. Several algorithms such as Random Forest and Gradient Boosting are applied to determine the most effective model. Evaluation is done using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared metrics. Random Forest and Gradient Boosting algorithms showed good accuracy results, indicating their suitability in predicting stunting risk. The main factors influencing stunting include daily nutrient intake, family economic status, access to health facilities, and feeding patterns. These findings provide important insights for more targeted interventions to reduce the prevalence of stunting. Limitations of the study and suggestions for future research are also discussed.

Keyword: *Stunting, Child Nutrition, Data Mining, Classification Algorithms, Data Preprocessing, Random Forest, Gradient Boosting.*