

**Pemodelan Prediksi *Stunting* di Jakarta: Studi Komparatif Naive
Bayes, Decision Tree, Random Forest, SVM, dan *Ensemble*
Methode**

SKRIPSI SARJANA TEKNOLOGI INFORMATIKA

Oleh

Ferdinand Marudut Tua Pane

217064516079



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL**

2024

**Pemodelan Prediksi *Stunting* di Jakarta: Studi Komparatif Naive
Bayes, Decision Tree, Random Forest, SVM, dan *Ensemble*
Methode**

SKRIPSI SARJANA

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar
Sarjana Teknik Teknologi Informatika dari Fakultas Teknologi Komunikasi dan Informatika

Oleh

Ferdinand Marudut Tua Pane

217064516079



PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI KOMUNIKASI DAN
INFORMATIKA
UNIVERSITAS NASIONAL

2024

KATA PENGANTAR

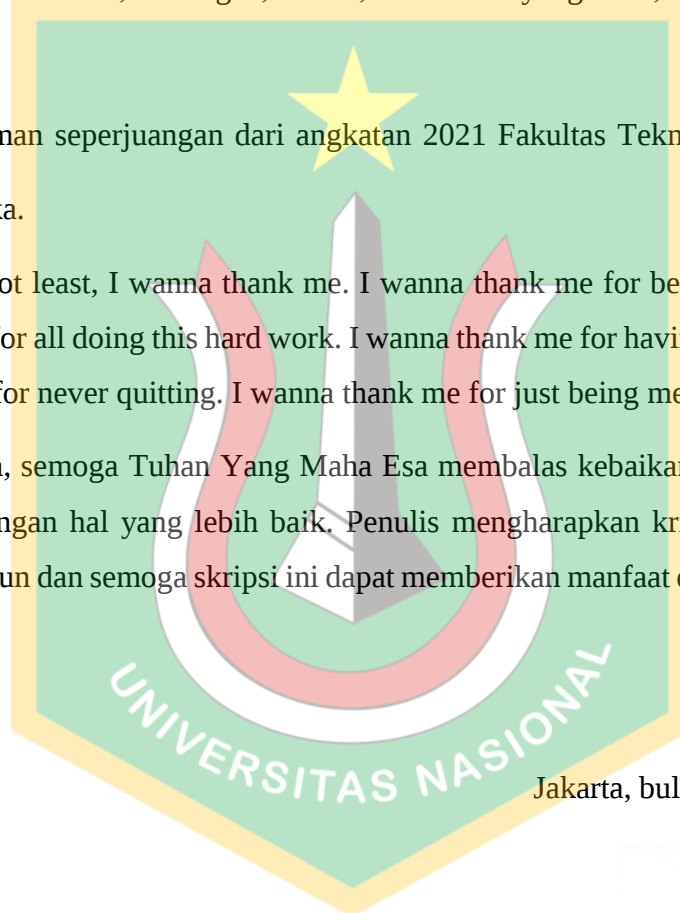
Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat dan karunia sehingga penulis dapat menyelesaikan skripsi dengan judul **“Pemodelan Prediksi *Stunting* di Jakarta: Studi Komparatif Naive Bayes, Decision Tree, Random Forest, SVM, dan *Ensemble Methods*”** sebagai salah satu syarat kelulusan Program Studi Sarjana Fakultas Teknologi Komunikasi dan Informatika.

Penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih terutama kepada dosen pembimbing Tugas Akhir, Djarot Hindarto, S.Kom., M.Kom yang telah meluangkan banyak waktu, tenaga, pikiran, bimbingan, arahan, motivasi serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan skripsi. Penulis juga mengucapkan banyak terima kasih kepada :

1. Bapak Dr. El Amry Bermawi Putera, M.A., selaku Rektor Universitas Nasional.
2. Bapak Dr. Agung Triayudi, S.Kom., M.Kom., selaku Dekan Fakultas Teknologi Komunikasi dan Informatika sekaligus Dosen Pembimbing Skripsi, yang telah memberikan waktu, tenaga, bimbingan, nasihat, dan pengarahan selama penyusunan skripsi ini.
3. Ibu Ir. Endah Tri Esti Handayani, MMSI, selaku Wakil Dekan Fakultas Teknologi Komunikasi dan Informatika.
4. Ibu Ratih Titi Komala Sari, S.T., M.M., MMSI, selaku Ketua Program Studi Informatika.
5. Bapak Djarot Hindarto, S.Kom., M.Kom, selaku Dosen Pembimbing Skripsi, yang telah memberikan waktu, tenaga, bimbingan, nasihat, dan pengarahan selama penyusunan skripsi ini
6. Bapak Yunan Fauzi Wijaya, S.Kom., MMSI., selaku Dosen Pembimbing Akademik Program Studi Informatika, yang telah memberikan saran, bimbingan, dan pendapat yang sangat berharga selama perkuliahan.

7. Seluruh pengajar Fakultas Teknologi Komunikasi dan Informatika yang tidak dapat penulis sebutkan satu per satu, karena ilmu dan pengalaman yang telah diberikan di bidang Ilmu Informatika selama masa perkuliahan.
8. Seluruh staf Sekretariat Fakultas Teknologi Komunikasi dan Informatika, yang telah mendukung proses perkuliahan dan administrasi selama penulis menempuh pendidikan.
9. Kedua orang tua yang sangat berharga, Bapak dan Mama, yang selalu membimbing, memberikan arahan, dukungan, materi, cinta kasih yang tulus, sejak masa kecil hingga saat ini.
10. Teman-teman seperjuangan dari angkatan 2021 Fakultas Teknologi Komunikasi dan Informatika.
11. Last but not least, I wanna thank me. I wanna thank me for believing in me. I wanna thank me for all doing this hard work. I wanna thank me for having no days off. I wanna thank me for never quitting. I wanna thank me for just being me at all times.

Akhir kata, semoga Tuhan Yang Maha Esa membalas kebaikan dan bantuan yang telah diberikan dengan hal yang lebih baik. Penulis mengharapkan kritik dan saran yang bersifat membangun dan semoga skripsi ini dapat memberikan manfaat di bidang Teknologi Informatika.



Jakarta, bulan - tanggal skripsi

A handwritten signature in black ink, appearing to read 'Ferdinand Marudut Tua Pane'.

Ferdinand Marudut Tua Pane

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Program Studi Teknik Informatika, Fakultas Teknologi Komunikasi dan Informatika, saya yang bertanda tangan di bawah ini:

Nama : Ferdinand Marudut Tua Pane

NIM : 217064516079

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Fakultas Teknologi Komunikasi dan Informatika, Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalti Free Right*) atas karya ilmiah saya yang berjudul:

PEMODELAN PREDIKSI *STUNTING* DI JAKARTA: STUDI KOMPARATIF NAIVE BAYES, DECISION TREE, RANDOM FOREST, SVM, DAN *ENSEMBLE* METHODE

Beserta perangkat yang ada (jika diperlukan). Dengan Hak ini Fakultas Teknologi Komunikasi dan Informatika berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkal data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 28 Februari 2025

Yang menyatakan



(Ferdinand Marudut Tua Pane)

HALAMAN PENGESAHAN
TUGAS AKHIR

**Pemodelan Prediksi Stunting di Jakarta: Studi Komparatif Naive Bayes,
Decision Tree, Random Forest, SVM, dan Ensemble Methods**



Ferdinand Marudut Tua Pane
217064516079

Dosen Pembimbing 1

A handwritten signature in black ink, appearing to read 'Djarot', is written over the name and title of the supervisor.

Djarot Hindarto, S.Kom., M.Kom.

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

Pemodelan Prediksi Stunting di Jakarta: Studi Komparatif Naive Bayes, Decision Tree, Random Forest, SVM, dan Ensemble Methods

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 28 Febuari 2025


METERAI TEMPEL
12AMX144184715
Ferdinand Marudut Tua Pane

217064516079

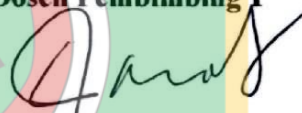
LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

Pemodelan Prediksi Stunting di Jakarta: Studi Komparatif Naive Bayes, Decision Tree, Random Forest, SVM, dan Ensemble Methods

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Informatika Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2024-2025 pada tanggal 25 Februari Tahun 2025

Dosen Pembimbing 1



Djarot Hindarto, S.Kom., M.Kom.

NIDN. 0317016905

Ketua Program Studi



Ratih Titik Komala Sari, S.T., M.M., MMSI.

NIDN. 0301038302

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Ferdinand Marudut Tua Pane
NPM : 217064516079
Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
Program Studi : Informatika
Tanggal Sidang : 25 Febuari 2025

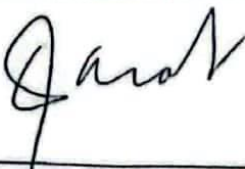
JUDUL DALAM BAHASA INDONESIA :

PEMODELAN PREDIKSI STUNTING DI JAKARTA: STUDI KOMPARATIF
NAIVE BAYES, DECISION TREE, RANDOM FOREST, SVM, DAN ENSEMBLE
METHODS

JUDUL DALAM BAHASA INGGRIS :

STUNTING PREDICTION MODELING IN JAKARTA: A COMPARATIVE STUDY
OF NAIVE BAYES, DECISION TREE, RANDOM FOREST, SVM, AND
ENSEMBLE METHODS

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 04/03/2025	TGL : 28/2/2025	TGL : 04/03/2025
	 RATIH TITI K.S.	 Ferdinand Marudut T.P.

Abstrak

Stunting masih menjadi permasalahan kesehatan masyarakat yang serius di Indonesia, dengan 6,3 juta anak terdampak, termasuk di wilayah Jakarta yang memiliki prevalensi tinggi. Stunting disebabkan oleh malnutrisi kronis, sanitasi yang buruk, dan keterbatasan akses layanan kesehatan, yang dapat mengakibatkan gangguan perkembangan kognitif dan fisik dalam jangka panjang. Meskipun berbagai program intervensi telah dilakukan, tantangan utama dalam penanggulangan stunting adalah kesulitan dalam mendeteksi anak yang berisiko sejak dini akibat kompleksitas faktor sosial ekonomi dan lingkungan. Oleh karena itu, diperlukan model prediksi yang akurat dan andal untuk meningkatkan efektivitas intervensi dini.

Penelitian ini membandingkan beberapa algoritma machine learning, yaitu Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), dan Ensemble Method dalam mengklasifikasikan risiko stunting pada anak. Dataset yang digunakan mencakup fitur penting seperti usia, tinggi badan, jenis kelamin, riwayat kesehatan keluarga, pendapatan rumah tangga, akses layanan kesehatan, dan kondisi sanitasi. Berbagai teknik pra-pemrosesan data, seleksi fitur, serta penyeimbangan kelas diterapkan untuk meningkatkan performa model.

Hasil evaluasi menunjukkan bahwa Random Forest memiliki akurasi tertinggi sebesar 98%, diikuti oleh Decision Tree dengan 97%, serta Ensemble Method (Naïve Bayes + Decision Tree) yang juga mencapai 97%. Namun, Naïve Bayes hanya mencapai akurasi 38%, dan SVM memiliki akurasi terendah sebesar 37%, menunjukkan bahwa metode ini kurang efektif dalam menangani dataset yang digunakan. Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan model prediksi stunting berbasis machine learning, yang dapat digunakan sebagai referensi bagi tenaga kesehatan dan pembuat kebijakan dalam upaya pencegahan dan intervensi dini.

Kata Kunci: *Stunting, Machine Learning, Naïve Bayes, Decision Tree, Random Forest, Ensemble Method, Prediksi*

Abstract

Stunting remains a significant public health issue in Indonesia, affecting 6.3 million children, with Jakarta experiencing a high prevalence of cases. Stunting is caused by chronic malnutrition, poor sanitation, and limited access to healthcare services, which can lead to long-term cognitive and physical development impairments. Despite various intervention programs, early detection of children at risk remains a challenge due to the complex interplay of socioeconomic and environmental factors. Therefore, an accurate and reliable predictive model is essential to enhance the effectiveness of early intervention strategies.

This study compares several machine learning algorithms, including Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), and Ensemble Methods, to classify stunting risk in children. The dataset consists of key features such as age, height, gender, family health history, household income, healthcare access, and sanitation conditions. Various data preprocessing techniques, feature selection, and class balancing methods were applied to improve model performance.

Evaluation results show that Random Forest achieved the highest accuracy of 98%, followed by Decision Tree with 97%, and Ensemble Method (Naïve Bayes + Decision Tree) also achieving 97%. However, Naïve Bayes only reached an accuracy of 38%, while SVM recorded the lowest accuracy at 37%, indicating that these methods were less effective for the given dataset. This study contributes to the development of machine learning-based stunting prediction models, which can serve as a reference for healthcare professionals and policymakers in designing more effective early prevention and intervention strategies.

Keywords: Stunting, Machine Learning, Naïve Bayes, Decision Tree, Random Forest, Ensemble Method, Prediction

Daftar Isi

KATA PENGANTAR	i
Daftar Isi	vi
Daftar Gambar	viii
Daftar Tabel	v
BAB I	1
Pendahuluan	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	2
1.3 Tujuan Penelitian	2
1.4 Batasan Masalah	3
1.5 Kontribusi	3
BAB II	4
Studi Literatur	4
2.1 Landasan Teori	4
2.1.1 <i>Stunting</i>	4
2.1.2 <i>Data Mining</i>	4
2.1.3 Klasifikasi	5
2.1.4 <i>Naïve Bayes Gaussian</i>	5
2.1.5 <i>Ensemble Methode</i>	5
2.1.6 <i>Decision Tree</i>	5
2.1.7 <i>Support Vector Machine (SVM)</i>	6
2.1.8 Dataset	6
2.2 Kajian Penelitian Terdahulu	7
BAB III	12
Metodologi Penelitian	12
3.1 Lokasi Penelitian	12
3.2 Waktu Penelitian	12
3.3 Alur Penelitian	13
3.4 Pengumpulan Data	14
3.4.1 Sumber Data	14
3.4.2 Fitur dalam <i>Dataset</i>	14
3.5 Pra-Pemrosesan Data	16
3.5.1 Pembersihan Data	16
3.5.2 Transformasi Data	16
3.5.3 Normalisasi Data	17

3.5.4	Pembagian Data	17
3.5.5	Penanganan <i>Outlier</i>	17
3.5.6	Pemilihan Fitur.....	17
BAB IV.....		18
Hasil Dan Diskusi		18
4.1	Deskripsi Dataset	19
4.1.2	Distribusi Kelas Probabilitas	20
4.1.3	Statistik Deskriptif Fitur Utama	21
4.2	Hasil Pra-Pemrosesan Data	23
4.2.1	Memuat Dataset	24
4.2.2	Melihat Data Awal	24
4.2.3	Mengubah Variabel Kategorikal Menjadi Numerik	24
4.2.4	Memisahkan Fitur dan Target	24
4.2.5	Normalisasi Data.....	24
4.2	Hasil Pengembangan Model.....	25
BAB V.....		29
Kesimpulan dan Saran		29
5.1	Kesimpulan	29
5.2	Saran	29
DAFTAR PUSTAKA.....		30
LAMPIRAN		31
KODINGAN.....		31



Daftar Gambar

Gambar 4. 1 Distribusi Probabilitas	21
Gambar 4. 2 Scatter plot untuk hubungan antara umur dan tinggi badan	22
Gambar 4. 3 Distribusi Berdasarkan Jenis Kelamin	23
Gambar 4. 4 Confussion Matrix Ensemble Methode	28



Daftar Tabel

Tabel 2. 1 Penelitian Terdahulu.....	7
Tabel 3. 1 Desain Sistem Model	13
Tabel 3. 2 Dataset.....	16
Tabel 4. 1 Model Naive Bayes	25
Tabel 4. 2 Model Random Forest.....	25
Tabel 4. 3 Model SVM.....	26
Tabel 4. 4 Model Decision Tree.....	26
Tabel 4. 5 Ensemble Methode	28

