

**TUGAS AKHIR
SKEMA SKRIPSI**

**PREDIKSI TINGKAT KEPARAHAN CEDERA AKIBAT
KECELAKAAN LALU LINTAS DENGAN RANDOM FOREST**



**FEBRIAN PRAMUDYA NUGRAHA
NIM : 215410012**

**PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA**

2025

**TUGAS AKHIR
SKEMA SKRIPSI**

**PREDIKSI TINGKAT KEPARAHAN CEDERA AKIBAT
KECELAKAAN LALU LINTAS DENGAN RANDOM FOREST**

Diajukan sebagai salah satu syarat untuk menyelesaikan studi pada

Program Sarjana

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Disusun Oleh

FEBRIAN PRAMUDYA NUGRAHA

NIM : 215410012

PROGRAM STUDI INFORMATIKA

PROGRAM SARJANA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2025

HALAMAN PERSETUJUAN UJIAN TUGAS AKHIR

Judul : Prediksi Tingkat Keparahan Cedera Akibat Kecelakaan
Lalu Lintas Dengan Random Forest

Nama : Febrian Pramudya Nugraha

NIM : 215410012

Program Studi : Informatika

Program : Sarjana

Semester : 7

Tahun Akademik : 2025

Telah diperiksa dan disetujui untuk diujikan
di hadapan Dewan Penguji Tugas Akhir

Yogyakarta, 16 Februari 2025

Dosen Pembimbing



Febri Nova Lenti, S.Si., M.T.

NIDN : 0505027101

HALAMAN PENGESAHAN

PREDIKSI TINGKAT KEPARAHAN CEDERA AKIBAT KECELAKAAN LALU LINTAS DENGAN RANDOM FOREST

Telah dipertahankan di depan Dewan Penguji dan dinyatakan diterima untuk memenuhi sebagian persyaratan guna memperoleh

Gelar S.Kom

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta, 18 Februari 2025

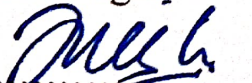
Dewan Penguji

NIDN

Tandatangan


1. Ariesta Damayanti, S.Kom., M.Cs.

0020047801

.....


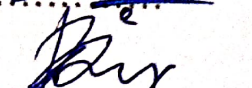
2. Febri Nova Lenti, S.Si., M.T.

0505027101

.....


3. Erna Hudianti Pujiarini, S.Si., M.Si.

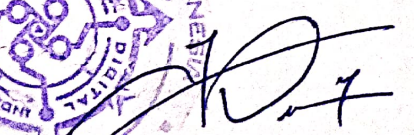
0528097101

.....


Mengetahui

Ketua Program Studi Informatika




Dini Fakta Sari, S.T., M.T.

NIDN : 0507108401

PERNYATAAN KEASLIAN TUGAS AKHIR

Dengan ini saya menyatakan bahwa naskah Tugas Akhir ini belum pernah diajukan untuk memperoleh gelar S.Kom disuatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 16 Februari 2025



Febrian Pramudya Nugraha

NIM : 215410012

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT yang telah memberikan kekuatan dan kesehatan serta kesempatan untuk menyelesaikan karya ini, akan dipersembahkan skripsi ini kepada :

1. **Kedua orang tua tercinta**, ayah dan ibu yang dengan kasih sayang, doa, dukungan dan pengorbanannya tidak pernah berhenti memberikan inspirasi kepada saya untuk terus berjuang dan mencapai mimpi-mimpi saya.
2. **Sahabat-sahabat terbaik** yang telah memberikan semangat, canda tawa dan dorongan moral selama proses pengerjaan ini. Dukungan kalian sangat berarti bagi saya.
3. **Dosen pembimbing** yang telah memberikan ilmu, arahan dan bimbingan dengan penuh kesabaran. Tanpa panduan dan masukan berharga dari kalian, karya ini tidak akan mungkin terselesaikan.

Karya ini adalah wujud dari perjalanan panjang yang penuh tantangan dan pelajaran berharga. Semoga skripsi ini dapat menjadi langkah awal untuk memberikan kontribusi yang berarti, baik bagi diri saya sendiri maupun orang lain di masa yang akan datang.

PRAKATA

Puji syukur saya panjatkan kehadirat Allah SWT atas limpahan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan skripsi ini dengan judul PREDIKSI TINGKAT KEPARAHAN CEDERA AKIBAT KECELAKAAN LALU LINTAS DENGAN RANDOM FOREST. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Sarjana di Universitas Teknologi Digital Indonesia. Penyusunan skripsi ini tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu saya ingin mengucapkan terimakasih yang sebesar-besarnya kepada :

1. **Ibu Febri Nova Lenti S.Si. , M.T.** selaku dosen pembimbing yang telah memberikan arahan, bimbingan dan motivasi selama proses penyusunan skripsi ini.
2. **Kedua orang tua** saya yang selalu mendoakan mendukung dan memberikan semangat tiada hentinya selama saya menempuh pendidikan ini.
3. **Sahabat dan teman** seperjuangan yang telah memberikan bantuan, dukungan moral dan semangat selama proses penelitian dan penulisan skripsi.
4. **Seluruh dosen di Universitas Teknologi Digital Indonesia** yang telah memberikan ilmu dan pengalaman berharga selama masa studi.

Saya menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu saya sangat terbuka terhadap saran dan kritik yang membangun. Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi pengembangan ilmu pengetahuan dan penelitian lebih lanjut.

Yogyakarta, 16 Februari 2025

Penulis

DAFTAR ISI

	Hal
PREDIKSI TINGKAT KEPARAHAN CEDERA AKIBAT KECELAKAAN LALU LINTAS DENGAN RANDOM FOREST	i
HALAMAN PERSETUJUAN UJIAN TUGAS AKHIR	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN TUGAS AKHIR	iv
HALAMAN PERSEMBAHAN	v
PRAKATA.....	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
INTISARI	xi
ABSTRACT	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Ruang Lingkup.....	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	5
2.1 Tinjauan Pustaka.....	5
2.2 Dasar Teori.....	7
2.2.1 Kecelakaan Lalu Lintas.....	7
2.2.2 Kaggle	7
2.2.3 Google Colab.....	7
2.2.4 Python.....	7
2.2.5 Data Mining	8
2.2.6 Analisis Korelasi.....	8
2.2.7 Machine Learning.....	9

2.2.8	Random Forest.....	10
2.2.9	Teknik SMOTE	11
2.2.10	Evaluasi.....	11
BAB III	METODE PENELITIAN	13
3.1	Data.....	13
3.2	Peralatan	13
3.3	Kebutuhan Perangkat Keras.....	13
3.4	Kebutuhan Perangkat Lunak.....	14
3.5	Pengumpulan Data.....	14
3.6	Analisis dan Rancangan Sistem	14
3.7	Diagram Proses	14
BAB IV	IMPLEMENTASI DAN PEMBAHASAN	19
4.1	Implementasi dan Pengujian Model	19
4.1.1	Pengumpulan Data	19
4.1.2	Preprocessing	20
4.1.3	Data Splinting	26
4.1.4	Model Training	29
4.1.5	Evaluasi.....	30
4.1.6	Pengujian Model.....	32
4.2	Pembahasan.....	34
BAB V	PENUTUP	36
5.1	SIMPULAN	36
5.2	SARAN	37
	DAFTAR PUSTAKA	38
	CARA MENJALANKAN PROGRAM	42
	LAMPIRAN	44

DAFTAR GAMBAR

	Hal
Gambar 3.1 Diagram Proses	14
Gambar 4. 1 Penghapusan Variabel Yang Tidak Relevan	20
Gambar 4. 2 Penanganan Data Tidak Sesuai	20
Gambar 4. 3 Penanganan Data Ambigu	21
Gambar 4.4 Mengatasi Missing Value Dengan Metode Imputasi.....	21
Gambar 4.5 Konversi Format Waktu	22
Gambar 4. 6 Konversi Data Kategorikal.....	23
Gambar 4.7 Pemisahan Fitur dan Target.....	27
Gambar 4. 8 Penerapan SMOTE	27
Gambar 4. 9 Data Sebelum SMOTE	27
Gambar 4. 10 Data Setelah SMOTE	28
Gambar 4. 11 Split Data	28
Gambar 4. 12 Model Training	29
Gambar 4. 13 Confusion Matrix.....	30
Gambar 4. 14 Hasil Akurasi Model.....	31
Gambar 4. 15 Hasil Klasifikasi Model.....	31
Gambar 4. 16 Feature Importance.....	32
Gambar 4. 17 Pengujian Model.....	33
Gambar 4. 18 Hasil Pengujian	33

DAFTAR TABEL

	Hal
Tabel 4. 1 Raw Data	19
Tabel 4. 2 Hasil Konversi Kategorikal.....	22
Tabel 4. 3 Analisis Korelasi (Spearman) Fitur Dengan Target.....	24
Tabel 4. 4 Hasil Percobaan n_estimators	35

INTISARI

Kecelakaan lalu lintas merupakan masalah serius yang mengancam keselamatan pengguna jalan dan menyebabkan korban jiwa setiap tahunnya. Berbagai faktor seperti kondisi cuaca, jenis kendaraan, kecepatan, serta kondisi pengemudi, memiliki peran penting dalam menentukan tingkat keparahan cedera yang dialami korban. Dengan memanfaatkan algoritma machine learning seperti Random Forest prediksi tingkat keparahan cedera dapat dilakukan secara lebih akurat, sehingga membantu pemangku kebijakan dalam merancang strategi pencegahan yang lebih efektif.

Penelitian ini menggunakan Algoritma Random Forest Classifier yang memiliki keunggulan dalam menangani data dengan banyak variabel serta secara alami mampu mengelola ketidakseimbangan data. Dataset yang digunakan diperoleh dari Kaggle berisi data kecelakaan lalu lintas dari Amerika Serikat dalam rentang waktu 2016 hingga 2023 dengan total 172.105 sampel dan 43 fitur yang mencakup berbagai aspek kecelakaan. Hasil penelitian ini menunjukkan bahwa model Random Forest berhasil mencapai akurasi sebesar 91%.

Kata Kunci : ketidakseimbangan data, prediksi kecelakaan, Random Forest, SMOTE, tingkat keparahan kecelakaan

ABSTRACT

Traffic accidents are a serious problem that threatens the safety of road users and causes fatalities every year. Various factors such as weather conditions, vehicle type, speed, and driver conditions play an important role in determining the severity of injuries experienced by victims. By utilizing machine learning algorithms such as Random Forest, the severity of injuries can be done more accurately, thereby helping policymakers in designing more effective prevention strategies.

This study uses the Random Forest Classifier Algorithm which has the advantage of handling data with many variables and is naturally able to manage data imbalance. The dataset used was obtained from Kaggle containing traffic accident data in the United States in the period 2016 to 2023 with a total of 172.105 samples and 43 features covering various aspects of accidents. The results of this study show that the Random Forest model managed to achieve an accuracy of 91%.

Keywords: accident severity, machine learning, predictive modeling, SMOTE, traffic safety