



**KLASIFIKASI KELULUSAN UJI EMISI KENDARAAN DENGAN
RANDOM FOREST, LOGISTIC REGRESSION, DAN SVM**

LAPORAN TUGAS AKHIR

SANDY MAHARDIKA

41521010037

**UNIVERSITAS
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**KLASIFIKASI KELULUSAN UJI EMISI KENDARAAN DENGAN
RANDOM FOREST, LOGISTIC REGRESSION, DAN SVM**

LAPORAN TUGAS AKHIR

SANDY MAHARDIKA

41521010037

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

**2025
MERCU BUANA**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : SANDY MAHARDIKA
NIM : 41521010037
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Klasifikasi Kelulusan Uji Emisi Kendaraan
Dengan Random Forest, Logistic Regression, Dan
SVM

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 24 Juli 2025



Sandy Mahardika.

UNIVERSITA
MERCU BUANA

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Sandy Mahardika
NIM : 41521010037
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Klasifikasi Kelulusan Uji Emisi Kendaraan
Dengan Random Forest, Logistic Regression,
Dan SVM

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Wawan Gunawan, S.Kom, MT.,
M.Kom
NIDN : 0424108104
Ketua Penguji : Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701
Penguji 1 : Ilham Nugraha, S.Kom, M.Sc
NIDN : 307098904
Penguji 2 : Siti Maesaroh, S.Kom., M.TI.
NIDN : 0413059003

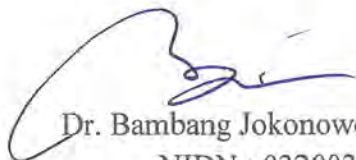
(GAS) ,
(Hadi Santoso) ,
(Ilham Nugraha) ,
(Siti Maesaroh) ,

UNIVERSITAS
MERCU BUANA

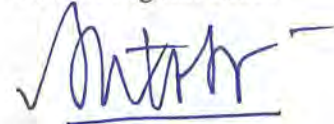
Jakarta, 25 Juli 2025

Mengetahui,

Dekan


Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002

Ketua Program Studi


Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Dosen Pembimbing, Wawan Gunawan, S.Kom, MT., M.Kom. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana..
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 24 Juli 2025



Sandy Mahardika

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Sandy Mahardika
NIM : 41521010037
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Klasifikasi Kelulusan Uji Emisi Kendaraan
Dengan Random Forest, Logistic Regression,
Dan SVM

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 24 Juli 2025



Sandy Mahardika.

ABSTRAK

Nama : Sandy Mahardika
NIM : 41521010037
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Klasifikasi Kelulusan Uji Emisi Kendaraan Dengan Random Forest, Logistic Regression, Dan SVM
Dosen Pembimbing : Wawan Gunawan, S.Kom, MT., M.Kom

Uji emisi kendaraan sangat penting untuk mitigasi polusi udara perkotaan. Memprediksi hasil uji dapat memberikan wawasan berharga bagi pemilik kendaraan untuk perawatan proaktif dan membantu regulator dalam menyusun kebijakan yang lebih terarah. Penelitian ini bertujuan untuk menyelidiki kemampuan prediktif model *machine learning* dalam mengklasifikasikan hasil uji emisi kendaraan. Saya menggunakan dataset historis yang mencakup fitur-fitur penting seperti merek kendaraan, jenis bahan bakar, tahun pembuatan, dan status kelulusan uji. Penelitian ini mengevaluasi tiga algoritma *machine learning* terkemuka: Random Forest, Logistic Regression, dan Support Vector Machine (SVM). Data melalui pra-pemrosesan menyeluruh dan kemudian dibagi menjadi set pelatihan dan pengujian untuk memastikan evaluasi model yang objektif. Kinerja model akan dinilai secara komprehensif menggunakan metrik klasifikasi standar seperti akurasi, presisi, *recall*, F1-score, dan Area di Bawah Kurva ROC (AUC). Studi ini berupaya mengidentifikasi algoritma yang memberikan kinerja prediktif terbaik untuk hasil uji emisi. Temuan diharapkan dapat berkontribusi pada literatur *machine learning* di bidang aplikasi lingkungan dan meletakkan dasar bagi pengembangan sistem cerdas. Sistem semacam itu dapat memungkinkan prediksi kepatuhan emisi kendaraan yang lebih cepat dan efisien, pada akhirnya mendukung upaya global menuju udara yang lebih bersih dan pengelolaan lingkungan yang berkelanjutan.

Kata Kunci: *Random Forest, Logistic Regression, SVM.*

ABSTRACT

Nama : Sandy Mahardika
NIM : 41521010037
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Classification of Vehicle Emission Tes Compliance
Using Random Forest, Logistic Regression, and SVM
Dosen Pembimbing : Wawan Gunawan, S.Kom, MT., M.Kom

Vehicle emission testing is crucial for mitigating urban air pollution. Predicting test outcomes offers valuable insights for proactive vehicle maintenance and assists regulators in crafting more targeted environmental policies. This study investigates the predictive capabilities of machine learning models for classifying vehicle emission test results. Utilize a historical dataset comprising key features such as vehicle make, fuel type, manufacturing year, and the definitive pass/fail status of the tests. Our research evaluates three prominent machine learning algorithms: Random Forest, Logistic Regression, and Support Vector Machine (SVM). The data undergoes rigorous preprocessing and is then strategically split into training and testing sets to ensure objective model evaluation. Model performance will be comprehensively assessed using standard classification metrics, including accuracy, precision, recall, F1-score, and the Area Under the ROC Curve (AUC). This study aims to identify the algorithm that delivers the best predictive performance for emission test outcomes. The findings are expected to contribute to the machine learning literature in environmental applications and lay foundational groundwork for developing intelligent systems. Such systems could enable faster, more efficient prediction of vehicle emission compliance, ultimately supporting global efforts towards cleaner air and sustainable environmental management.

Keywords: *Random Forest, Logistic Regression, SVM.*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA	4
2.1 Penelitian Terdahulu	4
2.2 Teori Pendukung.....	11
2.2.1 Random Forest	11
2.2.2 Logistic Regression.....	11
2.2.3 SVM.....	13
2.2.3 Uji Emisi	14
2.2.4 Python	14
BAB III METODE PENELITIAN	16
3.1 Jenis Penelitian.....	16
3.2 Tahapan Penelitian.....	16
a) Pengumpulan Data	17

b)	Eksplorasi Data.....	17
c)	Preprocessing Data	17
d)	Split Data	17
e)	Pelatihan Model.....	18
f)	Evaluasi Model.....	18
g)	Analisis Hasil	18
h)	Kesimpulan.....	19
BAB IV PEMBAHASAN		20
4.1	Dataset.....	20
4.2	Eksplorasi Data	21
4.3	Preprocessing Data.....	23
4.4	Split Data	26
4.5	Pelatihan Model	29
4.6	Evaluation Model.....	33
BAB V KESIMPULAN DAN SARAN		40
5.1	Kesimpulan	40
5.2	Saran	40
DAFTAR PUSTAKA.....		41
LAMPIRAN.....		44



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	4
Tabel 4. 1 Perbandingan 20% Dan 30%	36
Tabel 4. 2 Nilai Evaluasi (20%).....	37
Tabel 4. 3 Nilai Evaluasi (30%).....	37



DAFTAR GAMBAR

Gambar 2. 1 Random Forest Architechture	11
Gambar 3 1 Tahapan Penelitian	16
Gambar 4. 1 Dataframe	21
Gambar 4. 2 Informasi Atribut.....	22
Gambar 4. 3 Informasi Nilai Statistik Dataframe	22
Gambar 4. 4 Salah Satu Dari Value.Count	23
Gambar 4. 5 Missing Value	24
Gambar 4. 6 Handling Missing Value	25
Gambar 4. 7 Distribusi Uji Emisi	25
Gambar 4. 8 Distribusi Fitur Numerik	26
Gambar 4. 9 Distribusi Uji Emisi Berdasarkan Fitur Numerik	26
Gambar 4. 10 Perbandingan Data Setelah Di Smote	27
Gambar 4. 11 Output Mean(20% Data Test) F1 Scores	30
Gambar 4. 12 Output Mean(30% Data Test) F1 Scores	30
Gambar 4. 13 Komparasi Model Setelah Dilakukan Tuning Data Test(20%)	32
Gambar 4. 14 Komparasi Model Setelah Dilakukan Tuning Data Test(30%)	32
Gambar 4. 15 Confusion Matrix Model (20%).....	34
Gambar 4. 16 Confusion Matrix Model (30%).....	35
Gambar 4. 17 Features Importance Random Forest	36
Gambar 4. 18 Perbandingan Kinerja Model	38
Gambar 4. 19 Perbandingan 20% Dan 30%	39

DAFTAR LAMPIRAN

lampiran 1	44
Lampiran 2	45
Lampiran 3	47
Lampiran 4	48
Lampiran 5	49
Lampiran 6	51
Lampiran 7	52
Lampiran 8	53

