

**IMPLEMENTASI *MACHINE LEARNING* UNTUK PREDIKSI EMISI CO₂
DI KANADA**

TUGAS AKHIR



diajukan oleh:

Nama : Muhammad Daffa Miquilla

NIM : 22.01.4857

**PROGRAM DIPLOMA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**IMPLEMENTASI *MACHINE LEARNING* UNTUK PREDIKSI EMISI CO₂
DI KANADA**

TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat mencapai gelar Ahli Madya
Komputer Program Diploma – Program Studi Teknik Informatika



diajukan oleh

Nama : Muhammad Daffa Miquilla

NIM : 22.01.4857

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

TUGAS AKHIR

**IMPLEMENTASI *MACHINE LEARNING* UNTUK PREDIKSI EMISI CO₂ DI
KANADA**

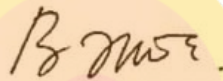
yang dipersiapkan dan disusun oleh

Muhammad Daffa Miquilla

22.01.4857

Telah disetujui oleh Dosen Pembimbing Tugas Akhir
pada tanggal 22 Mei 2025

Dosen Pembimbing,



Barka Satya, M.Kom
NIK. 190302126

HALAMAN PENGESAHAN
TUGAS AKHIR
IMPLEMENTASI *MACHINE LEARNING* UNTUK PREDIKSI EMISI CO₂ DI
KANADA

yang disusun dan diajukan oleh

Muhammad Daffa Miquilla

22.01.4857

Telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Mei 2025

Susunan Dewan Penguji

Nama Penguji

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393

Arvin C Frobenius, S.Kom., M.Kom.
NIK. 190302495

Tanda Tangan

Tugas Akhir ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Ahli Madya komputer
Tanggal 22 Mei 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Daffa Miquilla
NIM : 22.01.4857

Menyatakan bahwa Tugas Akhir dengan judul berikut:

IMPLEMENTASI *MACHINE LEARNING* UNTUK PREDIKSI EMISI CO₂ DI KANADA

Dosen Pembimbing : Barka Satya, 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, 22 Mei 2025

Yang Menyatakan,



Muhammad Daffa Miquilla

HALAMAN PERSEMBAHAN

Puji Syukur Kami panjatkan kepada Allah SWT, yang telah memberikan kesehatan, rahmat dan hidayah, sehingga penulis masih diberikan kesempatan untuk menyelesaikan tugas akhir ini, sebagai salah satu syarat untuk mendapatkan gelar Diploma III. Walaupun jauh dari kata sempurna, namun penulis bangga telah mencapai pada titik ini, yang akhirnya tugas akhir ini bisa selesai di waktu yang tepat, Seorang teman seangkatan pernah berkata, “Selesaikanlah apa yang sudah kau mulai”, sehingga hal inilah yang membuat penulis memacu dirinya untuk menyelesaikan semaksimal mungkin sehingga dapat menyelesaikan tugas akhir ini, diwaktu yang tepat. Tugas akhir ini saya persembahkan untuk :

1. Ayah dan Ibu, terimakasih atas doa, semangat, motivasi, pengorbanan, nasehat serta kasih sayang yang tidak pernah henti sampai saat ini.
2. Dosen Pembimbing kami Pak Barka Satya, M.Kom. yang sudah membimbing serta memberi masukan dan saran selama ini, sehingga saya dapat menyelesaikan tugas akhir ini.
3. Dosen D3 Teknik Informatika, yang telah memberikan ilmu dan motivasi dalam mengerjakan Tugas Akhir ini.
4. Teman - Teman Seangkatan D3 Teknik Informatika Angkatan 2022 teman-teman lainya yang sudah memberikan motivasi dalam mengerjakan Tugas Akhir ini.
5. Semua Komponen Amikom Yogyakarta, yang telah menerima kami dengan baik selama menempuh jenjang perkuliahan di Amikom Yogyakarta Hingga dapat menyelesaikan Tugas Akhir ini.

KATA PENGANTAR

Alhamdulillah, Puji dan Syukur kita panjatkan kepada Allah Subhanahu wa ta'ala. Dzat yang hanya kepada-Nya memohon pertolongan. Alhamdulillah atas segala pertolongan, rahmat, dan kasih sayang-Nya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “Implementasi *Machine Learning* untuk Prediksi Emisi CO₂ di Kanada”. Shalawat dan salam kepada Rasulullah Shallallāhu ‘Alaihi Wasallam yang senantiasa menjadi sumber inspirasi dan teladan terbaik untuk umat manusia.

Penulis dengan tulus menyadari bahwa banyak pihak yang telah memberikan dukungan dan bantuan selama proses penyelesaian studi dan tugas akhir ini. Oleh karena itu, dengan rasa hormat dan terima kasih yang mendalam, penulis ingin menyampaikan penghargaan kepada

1. Kedua orang tua penulis, Ayahanda Muhammad Ismail Shaleh dan Ibu Sofia Roma Ulian yang selalu memberi motivasi, doa, dan bimbingannya kepada penulis dan menjadi sumber kekuatan terbesar dihidup penulis untuk terus berjuang dan melangkah sejauh ini serta penyemangat dalam menyelesaikan tugas akhir ini.
2. Bapak Prof. Dr. M.Suyanto, M.M. sebagai Rektor Universitas Amikom Yogyakarta yang telah menyediakan fasilitas dan memberikan kesempatan kepada penulis dalam menimba ilmu pada lembaga pendidikan yang Beliau pimpin.
3. Ibu Prof. Dr. Kursini, S.Kom., M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang telah menyediakan fasilitas dan memberikan kesempatan kepada penulis dalam menimba ilmu pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta
4. Bapak Barka Satya, M.Kom., selaku dosen pembimbing, atas bimbingan, masukan, dan saran yang sangat berarti dalam proses penyusunan tugas akhir ini.
5. Kepada kakak dan adik penulis yang selalu memberikan support dalam mengerjakan tugas akhir ini, Molida Citra dan Abdul Mughits yang saya sayangi.

6. Kepada teman-teman tongkrongan *Alex Sambo* yang saat ini juga sedang berjuang menyelesaikan Tugas Akhirnya. Terima kasih atas semangat, masukan, dan dukungan yang tak henti-hentinya kalian berikan kepada penulis.
7. Sahabat dan rekan-rekan seperjuangan di Universitas Amikom Yogyakarta angkatan 2022, atas semangat dan bantuan yang diberikan
8. Serta seluruh pihak yang tidak dapat penulis sebutkan satu-persatu yang telah membantu kelancaran dalam menyusun tugas akhir ini.
9. Penulis juga ingin menyampaikan terima kasih kepada karya-karya musik dari Daniel Baskara Putra Mahendra (*Hindia*), .Feast, dan Lomba Sihir yang telah menjadi teman setia selama proses penulisan ini. Melodi dan lirik dari lagu-lagu seperti *Everything U Are*, *Tarot*, *Nina*, *Evaluasi*, dan *Peradaban* bukan hanya mengisi keheningan, tetapi juga menjadi penyemangat di saat-saat sulit. Musik-musik tersebut membantu penulis untuk tetap tenang, merenung, dan perlahan bangkit mengingatkan bahwa setiap perjalanan, seberat apa pun, tetap layak dijalani dengan penuh keberanian.
10. Dan terakhir, untuk diri saya sendiri, Muhammad Daffa Miquilla terima kasih telah bertahan, berproses, dan tidak menyerah hingga Tugas Akhir ini terselesaikan. Semoga langkah awal ini menjadi pijakan yang baik menuju masa depan, dan Allah SWT senantiasa memudahkan setiap perjalanan ke depan. Aamiin ya Rabbal'alamin.

Penulis menyadari bahwa laporan ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan ke depannya. Semoga laporan tugas akhir ini dapat memberikan manfaat, baik bagi penulis sendiri maupun bagi pembaca.

Yogyakarta, 22 Mei 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH.....	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan masalah.....	2
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	3
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 Literature Review.....	6
2.2 Landasan Teori.....	11
2.2.1 Emisi Karbon Dioksida (CO ₂)	11
2.2.2 <i>Cross-Industry Standard Process for Data Mining (CRISP-DM)</i>	11
2.2.3 <i>Machine Learning</i>	13
2.2.4 Algoritma Regresi dalam <i>Machine Learning</i>	14
2.2.5 Model Regresi	18
2.2.6 <i>Hyperparameter Tuning</i>	20
2.2.7 <i>Streamlit</i>	20
BAB III METODOLOGI PENELITIAN	22
3.1 Alat dan Bahan.....	22
3.1.1 Perangkat Keras (<i>Hardware</i>).....	22

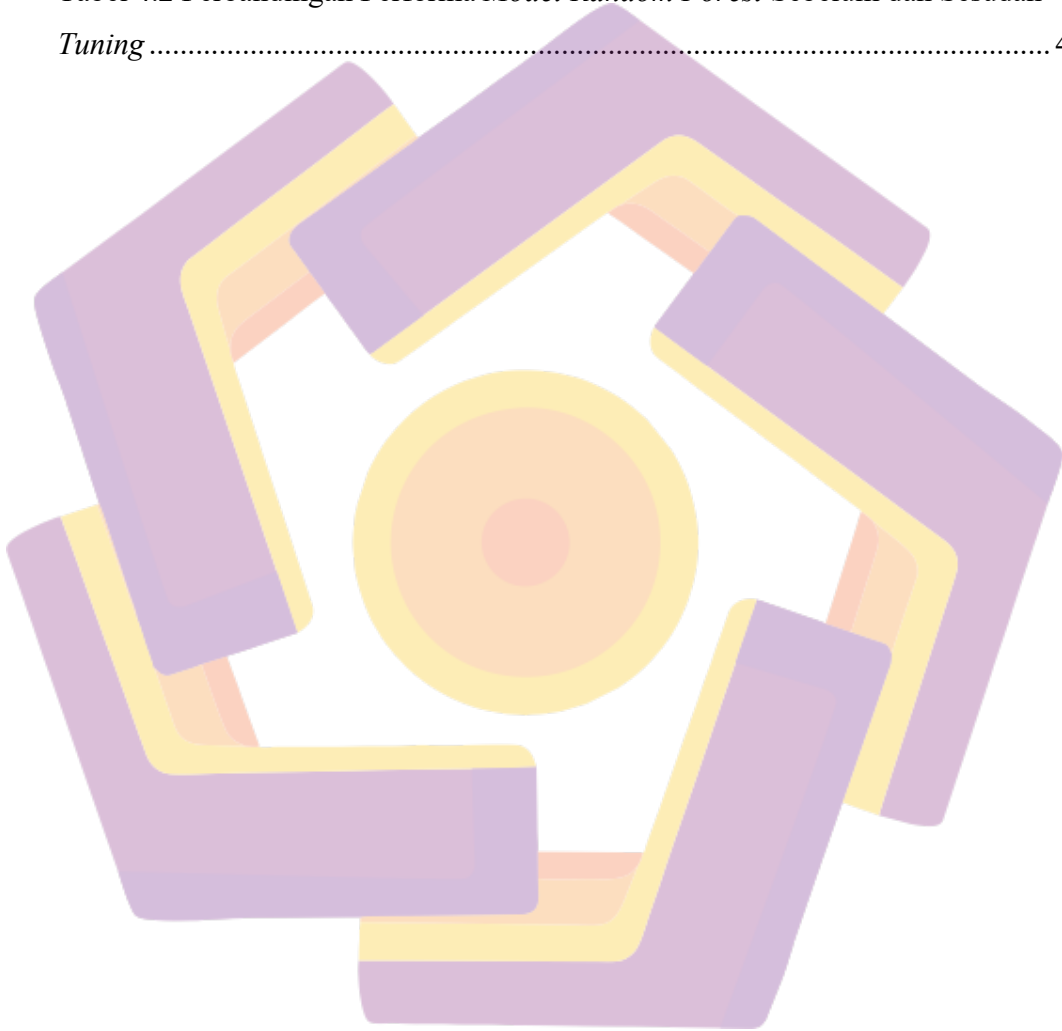
3.1.2 Perangkat Lunak (<i>Software</i>)	22
3.1.3 Spesifikasi Minimum.....	22
3.2 Alur Penelitian.....	23
3.2.1 <i>Business Understanding</i> (Pemahaman Bisnis)	24
3.2.2 <i>Data Understanding</i> (Pemahaman Data).....	25
3.2.3 <i>Data Preparation</i> (Persiapan Data).....	26
3.2.4 <i>Modeling</i> (Pemodelan).....	28
3.2.5 <i>Evaluation</i> (Evaluasi)	29
3.2.6 <i>Deployment</i> (Penyebaran).....	30
BAB IV HASIL DAN PEMBAHASAN	32
4.1 <i>Exploratory Data Analysis</i>	32
4.2 <i>Feature Engineering</i>	39
4.3 <i>Feature Selection</i>	39
4.4 <i>Data Preparation</i>	40
4.5 <i>Modeling</i>	41
4.6 <i>Evaluation</i>	44
4.7 <i>Deployment</i>	46
BAB V KESIMPULAN DAN SARAN	49
5.1 Kesimpulan.....	49
5.2 Saran.....	50
DAFTAR PUSTAKA	51

DAFTAR GAMBAR

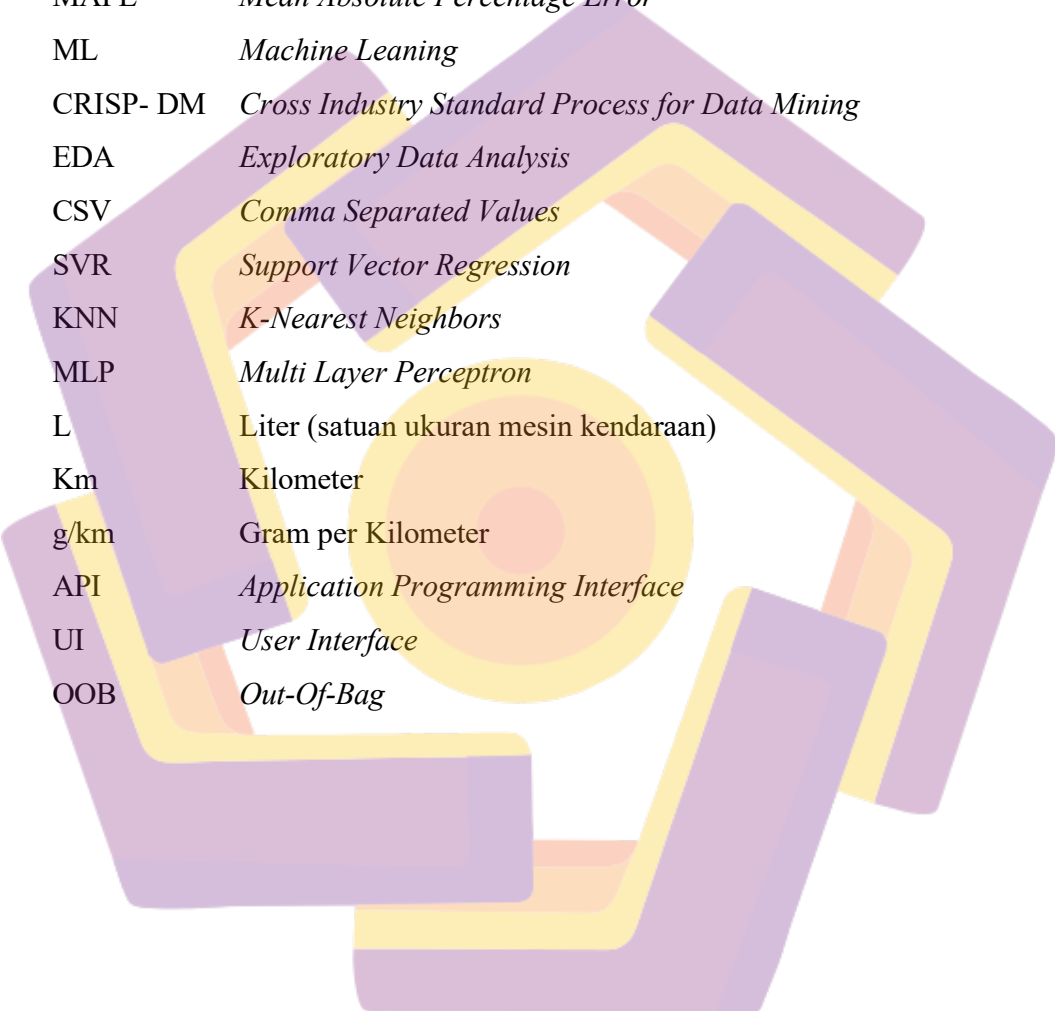
Gambar 2.1 Tahapan dalam <i>CRISP-DM</i>	12
Gambar 2.2 Jenis <i>Machine Learning</i>	14
Gambar 2.3 Mekanisme Prediksi <i>Random Forest Regressor</i>	16
Gambar 3.1 Tahapan Metode <i>CRISP-DM</i>	23
Gambar 4.1 Pratinjau Lima Baris Data Teratas dari <i>Dataset</i>	32
Gambar 4.2 Informasi <i>Dataset</i>	32
Gambar 4.3 Jumlah <i>Missing Values</i>	34
Gambar 4. 4 Distribusi Emisi CO ₂	35
Gambar 4.5 Matriks Korelasi Fitur Numerik.....	36
Gambar 4.6 <i>Scatter Plot</i> antara <i>Engine Size</i> dan <i>CO₂ Emissions</i>	37
Gambar 4.7 <i>Scatter Plot</i> antara <i>Fuel Consumption Comb</i> dan <i>CO₂ Emissions</i>	37
Gambar 4.8 Rata-rata Emisi CO ₂ per Kelas Kendaraan	38
Gambar 4.9 <i>Scatter Plot Engine Size per Cylinder vs CO₂ Emissions</i>	39
Gambar 4.10 Korelasi Fitur dan Identifikasi Multikolinearitas	40
Gambar 4.11 Pemisahan Fitur dan Target serta <i>Train-Test Split</i>	40
Gambar 4.12 Penerapan <i>Preprocessing Pipeline</i> dengan <i>StandardScaler</i> dan <i>OneHotEncoder</i>	41
Gambar 4.13 <i>Modeling</i>	42
Gambar 4.14 <i>Result of Modeling</i>	42
Gambar 4.15 <i>Hyperparameter Tuning</i>	43
Gambar 4.16 Evaluasi Model	44
Gambar 4.17 <i>Feature Importance of Random Forest</i>	46
Gambar 4.18 <i>Streamlit</i> bagian <i>Insight Analysis</i>	47
Gambar 4.19 <i>Streamlit</i> bagian <i>Predictions Emissions CO₂</i>	47
Gambar 4.20 <i>Streamlit</i> bagian <i>Chatbot</i>	48

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	7
Tabel 3.1 Perangkat Keras (<i>Hardware</i>)	22
Tabel 3.2 Perangkat lunak (<i>Software</i>).....	22
Tabel 3.3 Spesifikasi Minimum.....	23
Tabel 4.1 Statistik Deskriptif Fitur Numerik pada Dataset Emisi CO ₂	33
Tabel 4.2 Perbandingan Performa <i>Model Random Forest</i> Sebelum dan Sesudah <i>Tuning</i>	45

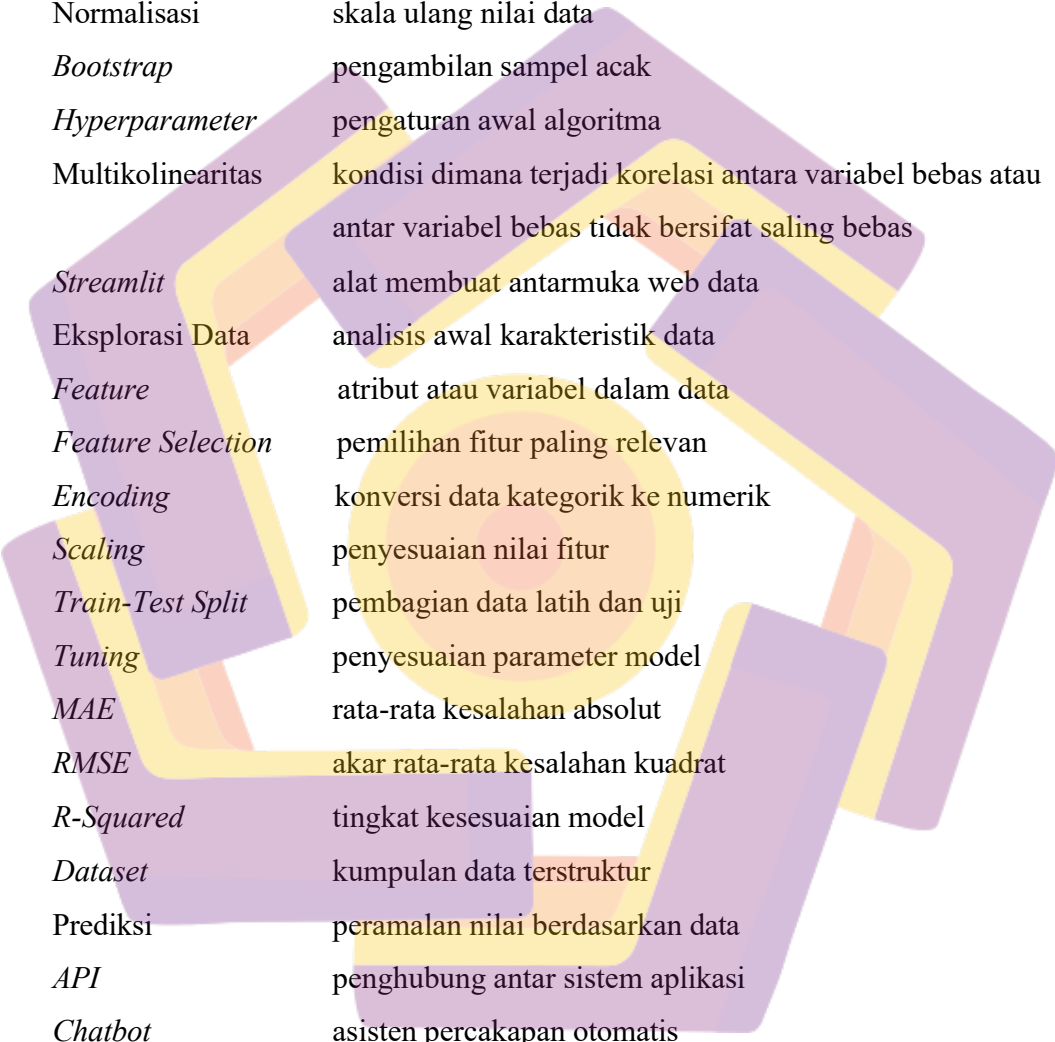


DAFTAR LAMBANG DAN SINGKATAN



CO ₂	Karbon Dioksida
SVM	<i>Support Vector Machines</i>
RMSE	<i>Root Mean Square Error</i>
MAE	<i>Mean Absolute Error</i>
R ²	Koefisien Determinasi (<i>R-Squared</i>)
MAPE	<i>Mean Absolute Percentage Error</i>
ML	<i>Machine Learning</i>
CRISP- DM	<i>Cross Industry Standard Process for Data Mining</i>
EDA	<i>Exploratory Data Analysis</i>
CSV	<i>Comma Separated Values</i>
SVR	<i>Support Vector Regression</i>
KNN	<i>K-Nearest Neighbors</i>
MLP	<i>Multi Layer Perceptron</i>
L	Liter (satuan ukuran mesin kendaraan)
Km	Kilometer
g/km	Gram per Kilometer
API	<i>Application Programming Interface</i>
UI	<i>User Interface</i>
OOB	<i>Out-Of-Bag</i>

DAFTAR ISTILAH



Implementasi	penerapan metode ke sistem
<i>Overfitting</i>	model terlalu cocok pada data latih
Regresi	prediksi nilai kontinu
<i>Modeling</i>	pembuatan model dari data
<i>Deployment</i>	penerapan model ke aplikasi
Normalisasi	skala ulang nilai data
<i>Bootstrap</i>	pengambilan sampel acak
<i>Hyperparameter</i>	pengaturan awal algoritma
Multikolinearitas	kondisi dimana terjadi korelasi antara variabel bebas atau antar variabel bebas tidak bersifat saling bebas
<i>Streamlit</i>	alat membuat antarmuka web data
Eksplorasi Data	analisis awal karakteristik data
<i>Feature</i>	atribut atau variabel dalam data
<i>Feature Selection</i>	pemilihan fitur paling relevan
<i>Encoding</i>	konversi data kategorik ke numerik
<i>Scaling</i>	penyesuaian nilai fitur
<i>Train-Test Split</i>	pembagian data latih dan uji
<i>Tuning</i>	penyesuaian parameter model
<i>MAE</i>	rata-rata kesalahan absolut
<i>RMSE</i>	akar rata-rata kesalahan kuadrat
<i>R-Squared</i>	tingkat kesesuaian model
<i>Dataset</i>	kumpulan data terstruktur
Prediksi	peramalan nilai berdasarkan data
<i>API</i>	penghubung antar sistem aplikasi
<i>Chatbot</i>	asisten percakapan otomatis
<i>Web App</i>	aplikasi berbasis web

INTISARI

Perubahan iklim menjadi masalah lingkungan yang serius akibat peningkatan gas rumah kaca, terutama karbon dioksida (CO₂). Pada tahun 2023, sektor transportasi di Kanada menyumbang sekitar 23% dari total emisi gas rumah kaca nasional, yaitu sekitar 157 megaton CO₂ ekuivalen. Emisi ini banyak berasal dari kendaraan berbahan bakar fosil dan berdampak langsung pada pemanasan global serta kerusakan lingkungan. Karena itu, diperlukan sistem yang dapat membantu memprediksi emisi kendaraan agar dapat digunakan untuk mendukung pengambilan keputusan dan kebijakan pengurangan emisi karbon. Penelitian ini menggunakan pendekatan *CRISP-DM* untuk membangun model prediksi emisi CO₂ kendaraan berbasis *machine learning*. Tujuh algoritma regresi yang digunakan adalah *Linear Regression*, *Ridge Regression*, *Random Forest*, *Gradient Boosting*, *Support Vector Regression (SVR)*, *K-Nearest Neighbor (KNN)*, dan *Neural Network*. Dataset diperoleh dari *Open Government Canada* melalui *Kaggle*, terdiri dari 7.385 data kendaraan. Proses penelitian meliputi *preprocessing data*, eksplorasi fitur, pelatihan model, evaluasi menggunakan metrik R², MAE, RMSE, dan MAPE, serta *tuning hyperparameter* untuk meningkatkan performa. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* menghasilkan performa terbaik dengan nilai R² sebesar 0,9982 dan RMSE sebesar 2,49 g/km. Fitur paling berpengaruh adalah konsumsi bahan bakar kombinasi (L/100 km). Model ini diimplementasikan dalam aplikasi web berbasis *Streamlit* yang dapat dimanfaatkan oleh konsumen, produsen kendaraan, dan pembuat kebijakan. Ke depan, sistem ini bisa dikembangkan dengan data kendaraan listrik, hybrid, dan hidrogen, serta mempertimbangkan pola berkendara dan kondisi cuaca.

Kata kunci: Emisi Karbon, *Machine Learning*, *CRISP-DM*, Transportasi, *Streamlit*

ABSTRACT

Climate change is becoming a serious environmental issue due to the increase in greenhouse gases, particularly carbon dioxide (CO₂). By 2023, the transportation sector in Canada accounts for about 23% of total national greenhouse gas emissions, which is about 157 megatonnes of CO₂ equivalent. Many of these emissions come from fossil fuel vehicles and have a direct impact on global warming and environmental damage. Therefore, a system that can help predict vehicle emissions is needed so that it can be used to support decision-making and policies to reduce carbon emissions. This research uses the CRISP-DM approach to build a machine learning-based vehicle CO₂ emission prediction model. The seven regression algorithms used are Linear Regression, Ridge Regression, Random Forest, Gradient Boosting, Support Vector Regression (SVR), K-Nearest Neighbor (KNN), and Neural Network. The dataset was obtained from Open Government Canada via Kaggle, consisting of 7.385 vehicle data. The research process includes data preprocessing, feature exploration, model training, evaluation using R², MAE, RMSE, and MAPE metrics, and hyperparameter tuning to improve performance. The results show that the Random Forest algorithm produces the best performance with an R² value of 0.9982 and an RMSE of 2.49 g/km. The most influential feature is the combined fuel consumption (L/100 km). The model is implemented in a Streamlit-based web application that can be utilized by consumers, vehicle manufacturers, and policy makers. In the future, the system can be extended with electric, hybrid, and hydrogen vehicle data, as well as considering driving patterns and weather conditions.

Keywords: Carbon Emissions, Machine Learning, CRISP-DM, Transportation, Streamlit