



**STUDI KASUS ANALISIS PREDIKTIF EMISI *CARBON* DI RWANDA
MENGUNAKAN *MACHINE LEARNING* MODEL *REGRESSION***

KERJA PRAKTIK



UNIVERSITAS
Dinamika

Oleh :

LUTHFI KRISNA BAYU

21410100005

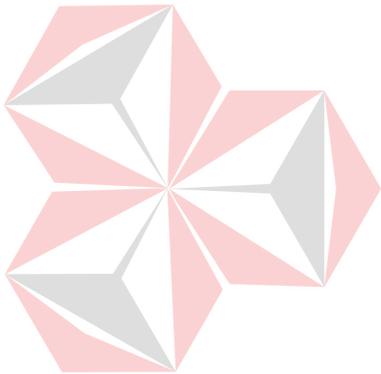
FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

2024

**STUDI KASUS ANALISIS PREDIKTIF EMISI *CARBON* DI RWANDA
MENGUNAKAN *MACHINE LEARNING* MODEL *REGRESSION***

Diajukan sebagai salah satu syarat untuk menyelesaikan
Program Sarjana



Disusun Oleh :

Nama : Luthfi Krisna Bayu

NIM : 21410100005

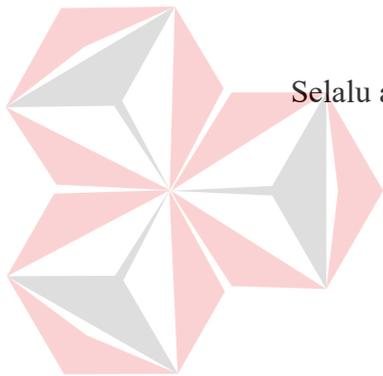
Program : S1 (Strata Satu)

Jurusan : Sistem Informasi

FAKULTAS TEKNOLOGI DAN INFORMATIKA

UNIVERSITAS DINAMIKA

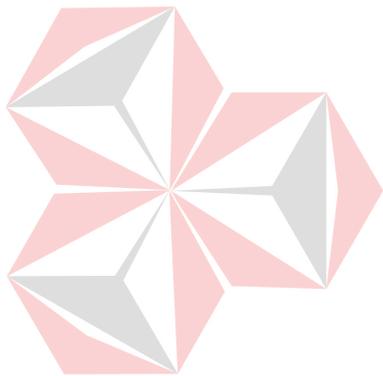
2024



Selalu andalkan Tuhan dalam setiap langkah hidup, maka Ia akan
memampukan

Luthfi Krisna Bayu

UNIVERSITAS
Dinamika



Laporan Kerja Praktik ini
Saya persembahkan kepada
Keluarga, Dosen Pembimbing, dan
Teman-teman yang saya kasihi

UNIVERSITAS
Dinamika

LEMBAR PENGESAHAN

**STUDI KASUS ANALISIS PREDIKTIF EMISI CARBON DI RWANDA
MENGUNAKAN MACHINE LEARNING MODEL REGRESSION**

Laporan Kerja Praktik oleh

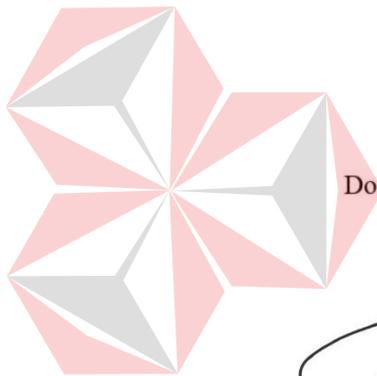
Luthfi Krisna Bayu

NIM: 21410100005

Telah diperiksa, diuji, dan disetujui

Surabaya, 01 Juli 2024

Disetujui



Dosen Pembimbing,

Tutut Wuriyanto, M.Kom.

NIDN. 0703056702

Penyelia,

M. Amrullah

NIP. 101805070009

Mengetahui,

Ketua Program Studi S1 Sistem Informasi

Digitally signed
by Julianto

Date: 2024.08.02

11:47:59 +07'00'

Julianto Lemantara, S.Kom., M.Eng.

NIDN. 0722108601

PERNYATAAN
PERSETUJUAN PUBLIKASI DAN KEASLIAN KARYA ILMIAH

Sebagai mahasiswa **Universitas Dinamika**, Saya :

Nama : **Luthfi Krisna Bayu**
NIM : **21410100005**
Program Studi : **S1 Sistem Informasi**
Fakultas : **Fakultas Teknologi dan Informatika**
Jenis Karya : **Laporan Kerja Praktik**
Judul Karya : **STUDI KASUS ANALISIS PREDIKTIF EMISI
CARBON DI RWANDA MENGGUNAKAN MACHINE
LEARNING MODEL REGRESSION**

Menyatakan dengan sesungguhnya bahwa :

1. Demi pengembangan Ilmu Pengetahuan, Teknologi dan Seni, Saya menyetujui memberikan kepada **Universitas Dinamika** Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas seluruh isi/sebagian karya ilmiah Saya tersebut diatas untuk disimpan, dialihmediakan, dan dikelola dalam bentuk pangkalan data (*database*) untuk selanjutnya didistribusikan atau dipublikasikan demi kepentingan akademis dengan tetap mencantumkan nama Saya sebagai penulis atau pencipta dan sebagai pemilik Hak Cipta.
2. Karya tersebut diatas adalah hasil karya asli Saya, bukan plagiat baik sebagian maupun keseluruhan. Kutipan, karya, atau pendapat orang lain yang ada dalam karya ilmiah ini semata-mata hanya sebagai rujukan yang dicantumkan dalam Daftar Pustaka Saya.
3. Apabila dikemudian hari ditemukan dan terbukti terdapat tindakan plagiasi pada karya ilmiah ini, maka Saya bersedia untuk menerima pencabutan terhadap gelar kesarjanaan yang telah diberikan kepada Saya.

Demikian surat pernyataan ini Saya buat dengan sebenar-benarnya.

Surabaya, 01 Juli 2024



Luthfi Krisna Bayu
NIM : 21410100005

ABSTRAK

Emisi karbon (CO₂) adalah pelepasan gas karbon dioksida (CO₂) ke atmosfer akibat pembakaran fosil dan berbagai macam proses industri. Kemampuan untuk memantau emisi karbon secara akurat merupakan langkah penting dalam upaya melawan perubahan iklim. Pembacaan karbon yang tepat memungkinkan peneliti dan pemerintah untuk memahami sumber dan pola keluaran massa karbon. Dewasa ini, Eropa dan Amerika Utara memiliki sistem yang luas untuk memantau emisi karbon di darat, namun sangat sedikit yang tersedia di Afrika termasuk Rwanda. Penelitian ini bertujuan untuk menerapkan model *machine learning Random Forest* menggunakan data emisi CO₂ yang didapat dari Kaggle berdasarkan pengamatan satelit Sentinel-5P untuk memprediksi emisi karbon di masa depan di Rwanda. Model regresi digunakan dalam analisis prediktif ini untuk memberikan estimasi emisi karbon yang dapat membantu pemerintah dan aktor lainnya dalam memperkirakan tingkat emisi karbon di Rwanda. Berdasarkan hasil penelitian, didapatkan hasil yang baik (nilai *error* yang kecil) dengan *evaluation metrics* RMSE (*Root Mean Squared Error*) sebesar 18,8 dan R^2 sebesar 0,98.

Kata Kunci: Analisis Prediktif, Emisi Karbon, *Machine Learning*, Model Regresi

KATA PENGANTAR

Puji syukur dengan kehadiran Tuhan YME yang telah memberikan berkat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan kerja praktik ini dengan judul “STUDI KASUS ANALISIS PREDIKTIF EMISI *CARBON* DI RWANDA MENGGUNAKAN *MACHINE LEARNING* MODEL *REGRESSION*” ini dengan baik dan lancar. Penyelesaian laporan Kerja Praktik ini sebagai syarat wajib untuk menyelesaikan program sarjana. Tidak terlepas dari bantuan dari pihak yang telah memberikan masukan, nasihat, saran, kritik kepada penulis. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan rasa terima kasih kepada:).

1. Ayah dan Ibu tercinta yang memberikan doa dan dukungan penuh kepada saya
2. Bapak Tutut Wuriyanto, M.Kom. selaku Dosen Pembimbing yang sudah memberikan bimbingan selama proses penyelesaian kerja praktik.
3. Kak Salsabila Nur Yasmin, selaku Mentor MSIB GreatEdu *Batch 6*
4. Untuk sahabat dan teman – teman perkuliahan di Universitas Dinamika Surabaya yang telah membantu dalam proses penyelesaian kerja praktik

Penulis menyadari bahwa laporan ini masih jauh dari kata sempurna. Dengan demikian penulis mengharapkan kritik dan saran yang membangun dari pembaca untuk penyempurnaan dalam menyelesaikan laporan. Semoga laporan Kerja Praktik ini dapat bermanfaat untuk penulis sendiri, dan para pembaca.

Surabaya, 10 Juli 2024

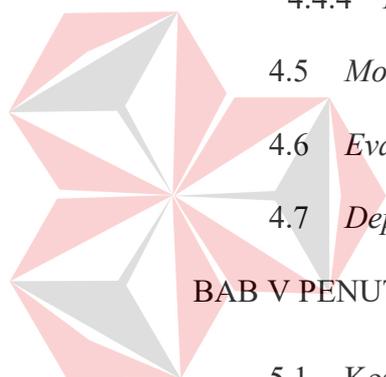


Penulis

DAFTAR ISI

	Halaman
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan.....	5
1.5 Manfaat.....	5
BAB II GAMBARAN UMUM.....	6
2.1 Latar Belakang Perusahaan	6
2.2 Identitas Perusahaan.....	6
2.3 Visi Perusahaan	7
2.4 Misi Perusahaan	7
2.5 Struktur Organisasi.....	7
BAB III LANDASAN TEORI.....	10
3.1 Emisi Karbon.....	10
3.2 <i>Machine Learning</i>	11
3.3 <i>Regression</i>	11
3.4 Python (<i>Data Science</i>).....	12

3.5	<i>Streamlit</i>	13
BAB IV DESKRIPSI PEKERJAAN		14
4.1	Metodologi Penelitian	14
4.2	<i>Business Understanding</i>	15
4.3	<i>Data Understanding</i>	16
4.4	<i>Data Preparation</i>	18
4.4.1	EDA	18
4.4.2	<i>Features Engineering</i>	24
4.4.3	<i>Select Features</i>	27
4.4.4	<i>Preparation</i>	27
4.5	<i>Modelling</i>	29
4.6	<i>Evaluation</i>	31
4.7	<i>Deployment</i>	32
BAB V PENUTUP		36
5.1	Kesimpulan	36
5.2	Saran	36
DAFTAR PUSTAKA		38

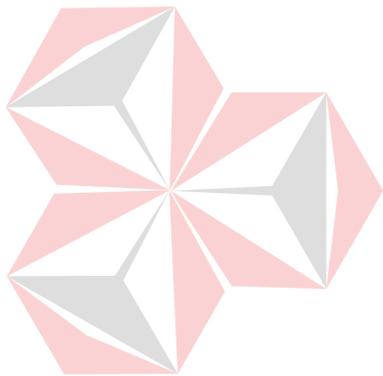


UNIVERSITAS
Dinamika

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Logo GreatEdu	6
Gambar 2.2 Struktur Organisasi GreatEdu	9
Gambar 4.1 CRISP-DM	14
Gambar 4.2 <i>Basic description train.head()</i>	18
Gambar 4.3 <i>Basic description train.info()</i>	19
Gambar 4.4 <i>Basic description train.info()</i>	19
Gambar 4.5 <i>Missing value</i>	20
Gambar 4.6 <i>Visualization of Emission by Date</i>	20
Gambar 4.7 Titik lokasi emisi.....	21
Gambar 4.8 Titik lokasi emisi 0.....	21
Gambar 4.9 <i>Map Plot</i>	22
Gambar 4.10 <i>Pearson's Correlation</i>	23
Gambar 4.11 <i>Date adjustment & drop columns</i>	24
Gambar 4.12 <i>Visualization emission by date after adjustment</i>	25
Gambar 4.13 <i>Season Feature</i>	25
Gambar 4.14 <i>Holiday Feature</i>	25
Gambar 4.15 <i>Cyclic Feature</i>	26
Gambar 4.16 <i>Rotate Location Feature</i>	26
Gambar 4.17 <i>Distance to max emission feature</i>	26
Gambar 4.18 <i>Select Features</i>	27
Gambar 4.19 <i>Train-Test Split</i>	28
Gambar 4.20 <i>Scaling</i>	28

Gambar 4.21 <i>Modelling</i>	29
Gambar 4.22 <i>Result of Modelling</i>	30
Gambar 4.23 <i>RandomForest with Hyperparameter Tuning</i>	30
Gambar 4.24 <i>Evaluation Model</i>	31
Gambar 4.25 <i>Feature Importance of RandomForest</i>	32
Gambar 4.26 Streamlit bagian Deskripsi <i>Dataset</i>	33
Gambar 4.27 Streamlit bagian EDA	34
Gambar 4.28 Streamlit bagian Prediksi Emisi CO2	35



UNIVERSITAS
Dinamika

BAB I

PENDAHULUAN

1.1 Latar Belakang

Project ini terlaksana saat mengikuti program Magang dan Studi Independen (MSIB) *Batch 6* pada PT Greatedu Global Mahardika atau yang lebih *familier* disebut GreatEdu yang merupakan mitra yang berlokasi di daerah Jakarta Selatan. Topik yang didapat terkait *project* ini merupakan prediksi emisi karbon di Rwanda dengan menggunakan model *machine learning* berjenis *regression*. Prediksi emisi karbon dengan *machine learning regression* adalah proses menggunakan algoritma pembelajaran mesin untuk memperkirakan jumlah emisi karbon dioksida (CO₂) yang dilepaskan ke atmosfer berdasarkan data historis dan *variabel-variabel* terkait, seperti konsumsi energi dan aktivitas industri (Larasati, 2023). Teknik ini melibatkan pengumpulan data dari data historis pada satelit Sentinel-5P, proses EDA (*Exploratory Data Analysis*), *pre-processing data*, pelatihan model *regression*, evaluasi kinerja model, serta penggunaan model untuk membuat prediksi yang akurat.

Perubahan iklim merupakan tantangan besar yang dihadapi dunia saat ini, dengan emisi karbon dioksida (CO₂) dari aktivitas manusia sebagai salah satu penyebab utamanya. Mengurangi emisi karbon menjadi prioritas global untuk mengatasi dampak perubahan iklim (Mardani, 2023). Oleh karena itu, kemampuan untuk memantau emisi karbon secara akurat sangat penting. Pengukuran karbon yang tepat memungkinkan peneliti dan pemerintah memahami sumber serta pola keluaran massa karbon, sehingga dapat merumuskan kebijakan yang efektif. Di

Eropa dan Amerika Utara telah mengembangkan dan menerapkan sistem pemantauan emisi karbon yang canggih (Van Der Woude dkk., 2023). Namun, banyak negara di Afrika, termasuk Rwanda, masih kekurangan infrastruktur pemantauan yang memadai. Hal ini menyebabkan kesulitan dalam memperoleh data emisi karbon yang akurat dan berkelanjutan. Rwanda, sebagai negara yang sedang berkembang di Afrika Timur, menghadapi tantangan ini secara langsung. Oleh karena itu, penelitian ini memfokuskan pada Rwanda untuk mengembangkan metode pemantauan emisi karbon yang efektif.

Salah satu pendekatan modern yang digunakan untuk memantau dan menganalisis emisi karbon adalah melalui model *machine learning* untuk menganalisis jumlah emisi karbon dengan menggunakan *variabel-variabel* faktor pendukung emisi. *Machine learning*, terutama model regresi, memiliki potensi besar dalam analisis prediktif emisi karbon (Geevaretnam dkk., 2022). Model regresi menggunakan algoritma untuk menganalisis dan memprediksi *data* emisi karbon berdasarkan pola yang terdeteksi dari *data training*. Metode ini dapat memproses *data* dalam jumlah besar dengan cepat dan mendeteksi pola yang kompleks, menjadikannya alat yang kuat dalam memahami dan memprediksi emisi karbon.

Model regresi memiliki beberapa kelebihan dan kelemahan. Keunggulan utama dari model regresi adalah kesederhanaan dan interpretabilitasnya (Yang dkk., 2023), yang memungkinkan peneliti dan pembuat kebijakan untuk memahami hubungan antara *variabel input* dan *output* dengan jelas. Selain itu, model regresi dapat diterapkan pada berbagai jenis data dan relatif mudah diimplementasikan (Obagbuwa dkk., 2023). Namun, model ini juga memiliki kelemahan, seperti risiko

overfitting, keterbatasan data yang dapat mengurangi akurasi prediksi (Kurniadi & Larasati, 2022), dan pengaruh *variabel* lingkungan yang mungkin tidak tercakup dalam model.

Dalam konteks penelitian ini, model regresi diterapkan untuk memprediksi emisi karbon di Rwanda menggunakan data dari pengamatan satelit Sentinel-5P. Meskipun model regresi memiliki potensi besar, terdapat beberapa masalah yang perlu diatasi. *Overfitting* merupakan masalah umum dalam model *machine learning*, di mana model terlalu menyesuaikan diri dengan data *training* sehingga performanya menurun pada data baru (Bengnga & Ishak, 2022). Keterbatasan data juga menjadi tantangan, karena data yang kurang memadai dapat mengurangi akurasi prediksi. Selain itu, *variabel* lingkungan seperti kondisi cuaca, lingkungan dan aktivitas manusia yang tidak diperhitungkan dalam model dapat mempengaruhi hasil prediksi.

Untuk mengatasi masalah tersebut, beberapa solusi perbaikan diusulkan. Pertama, teknik Transformasi, penyesuaian *outliers*, dan *scaling* agar dapat digunakan untuk mengurangi *overfitting*. Kedua, melakukan *hyperparameter tuning*, agar model yang dihasilkan dapat lebih optimal dan *robust*. Ketiga, menggunakan metrik evaluasi seperti RMSE (*Root Mean Squared Error*) dan R^2 untuk mengetahui seberapa *robust* model kita dalam melakukan prediksi. Keempat, mengintegrasikan *variabel* eksternal seperti kondisi cuaca, lingkungan dan aktivitas manusia agar dapat meningkatkan akurasi prediksi model terhadap emisi karbon. Hal ini karena faktor-faktor tersebut dapat mempengaruhi tingkat emisi karbon, sehingga dengan memperhitungkannya, model dapat memberikan prediksi yang lebih akurat dan representatif.

Penelitian ini bertujuan untuk menerapkan model *machine learning* menggunakan data emisi CO₂ dari pengamatan satelit Sentinel-5P untuk memprediksi emisi karbon di masa depan di Rwanda. Model regresi yang digunakan dalam analisis prediktif ini diharapkan dapat memberikan estimasi emisi karbon yang akurat, membantu pemerintah dan aktor lainnya dalam memperkirakan tingkat emisi karbon di seluruh Afrika, bahkan di tempat-tempat yang sulit dipantau secara langsung.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, berikut rumusan masalah yang dihasilkan, yaitu bagaimana menghasilkan emisi karbon di Rwanda dengan mengukur keakuratan dan prediksi dengan menggunakan model-model *machine learning regression*

1.3 Batasan Masalah

Berdasarkan uraian di atas, maka dalam pelaksanaan Kerja Praktik terdapat beberapa batasan masalah, antara lain :

1. Penelitian ini akan menggunakan data emisi CO₂ dari pengamatan satelit Sentinel-5P yang didapat dari *Kaggle*.
2. Penelitian ini fokus pada penggunaan model regresi untuk analisis prediktif.
3. Penelitian ini memiliki batasan waktu yang spesifik yaitu 2019 – 2021 untuk pengumpulan data, pengembangan model, dan analisis hasil.
4. Model yang digunakan adalah *Random Forest Regressor*
5. *Deployment* yang digunakan adalah *Streamlit*

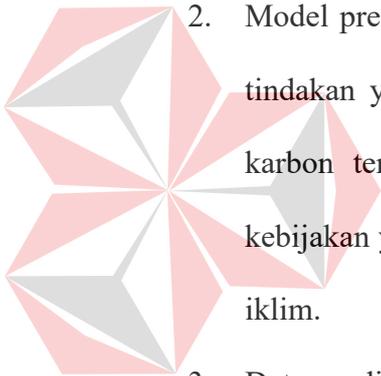
1.4 Tujuan

Berdasarkan uraian dari latar belakang dan rumusan masalah, maka dapat disesuaikan bahwa, tujuan dari kerja praktik ini yaitu untuk mengembangkan model *Machine Learning* menggunakan data emisi karbon dari observasi satelit Sentinel-5P untuk memprediksi emisi karbon daerah Rwanda di masa yang akan datang.

1.5 Manfaat

Adapun manfaat dari pelaksanaan Kerja Praktik ini antarlain sebagai berikut:

1. Pemantauan Emisi yang Lebih Efektif
2. Model prediksi ini membantu pemerintah dan masyarakat dalam mengambil tindakan yang lebih tepat untuk mengelola dan mengurangi dampak emisi karbon terhadap perubahan iklim. Didukung dengan data yang akurat, kebijakan yang lebih tepat sasaran dapat diterapkan untuk mitigasi perubahan iklim.
3. Data prediksi emisi karbon dapat digunakan untuk mendukung perencanaan kebijakan lingkungan yang lebih baik. Dengan memahami pola emisi, pemerintah dapat merancang dan mengimplementasikan strategi yang lebih efektif untuk menurunkan emisi karbon di masa depan.
4. Proyek ini menunjukkan bagaimana teknologi, seperti machine learning dan platform visualisasi data, dapat digunakan untuk tujuan lingkungan. Hal ini dapat mendorong lebih banyak inovasi dan penggunaan teknologi dalam upaya konservasi lingkungan.



UNIVERSITAS
Dinamika

BAB II

GAMBARAN UMUM

2.1 Latar Belakang Perusahaan

Perusahaan yang memberikan tugas untuk mengerjakan *Project* ini selama program MSIB *Batch* 6 adalah GreatEdu, yang merupakan sebuah *online educational marketplace & learning management system* (LMS) platform untuk berbagi keahlian dan pengetahuan yang terstruktur dan dapat dimanfaatkan secara luas baik untuk individu maupun komunitas dengan harga yang terjangkau. GreatEdu memiliki kantor yang berlokasi di Jl. Duren Tiga Raya No.09, RT.2/RW.1, Duren Tiga, Kec. Pancoran, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12760. Logo dari GreatEdu sendiri dapat dilihat pada Gambar 2.1.



Gambar 2.1 Logo GreatEdu

2.2 Identitas Perusahaan

Nama Instansi : PT Greatedu Global Mahardika (GreatEdu)
Alamat : Jl. Duren Tiga Raya No.09, RT.2/RW.1, Duren Tiga, Kec.
Pancoran, Kota Jakarta Selatan, Daerah Khusus Ibukota
Jakarta 12760
No. Telepon : 0816-338-441
Website : <https://greatedu.co.id/>

Email : class@greatedu.co.id

2.3 Visi Perusahaan

GreatEdu memiliki visi untuk hadir sebagai platform dan ekosistem bagi para pelajar, pengajar, instruktur dan praktisi profesional yang membutuhkan kelas dan penyaluran potensi dalam berbagai ilmu menghadapi dunia kerja di era 4.0.

2.4 Misi Perusahaan

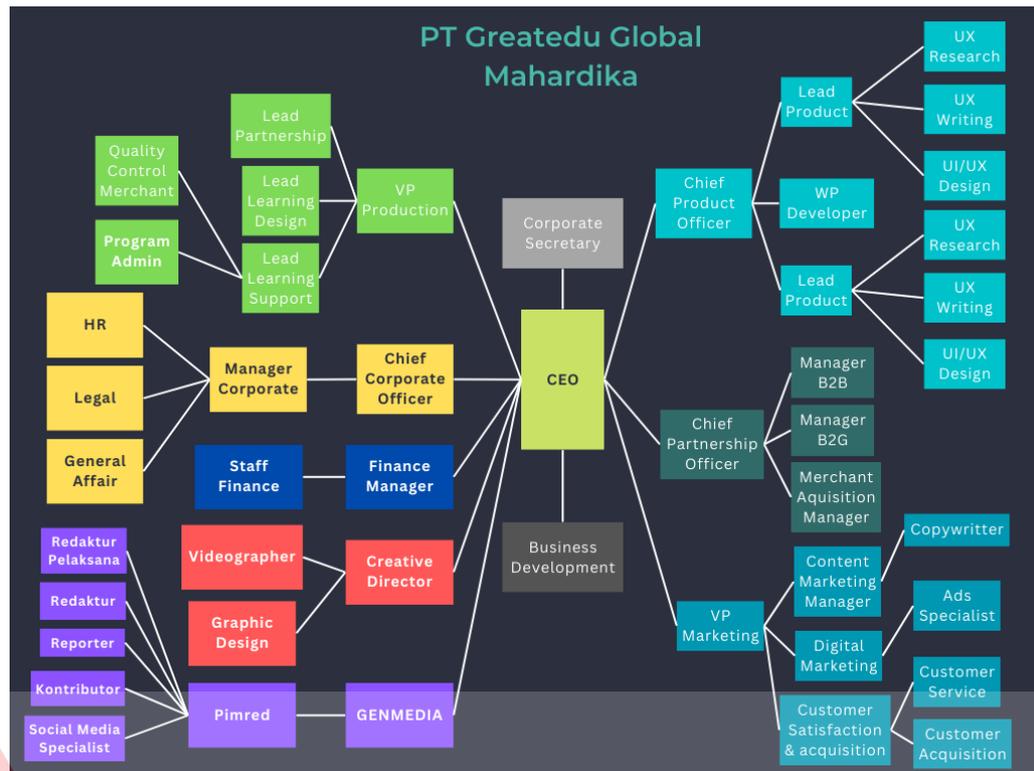
Misi dari GreatEdu adalah sebagai berikut:

1. Berkolaborasi dengan semua pihak yang peduli dan berkecimpung di dunia pendidikan untuk membangun ekosistem pendidikan yang saling melengkapi.
2. Meningkatkan kualitas pendidikan di Indonesia melalui kemitraan strategis dan sinergi antar lembaga pendidikan.
3. Mengakui bahwa tanggung jawab peningkatan kualitas pendidikan tidak hanya berada pada pemerintah, tetapi juga memerlukan kontribusi dari berbagai pihak.
4. Memposisikan diri sebagai kontributor aktif dalam upaya peningkatan kualitas pendidikan di Indonesia.
5. Mendukung paradigma baru dalam dunia pendidikan yang menekankan kebebasan dan otonomi institusi pendidikan, sejalan dengan visi Menteri Pendidikan Nadiem Makarim.

2.5 Struktur Organisasi

Struktur organisasi dari GreatEdu diperlihatkan pada Gambar 2.2. Pada posisi tertinggi/*central* terdapat *Chief Executive Officer* (CEO). Di bawah CEO, terdapat 10 divisi lainnya yaitu *VP Production*, *Chief Corporate Officer*, *Finance*

Manager, Creative Director, Genmedia, Chief Production Officer, Chief Partnership Officer, VP Marketing, Business Development, dan Corporate Secretary. Divisi VP Production dibagi menjadi 3, yaitu Lead Partnership, Lead Learning Design, dan Lead Learning Support (Quality Control Merchant dan Program Admin). Divisi Chief Corporate Officer membawahi 1 divisi, yaitu Manager Corporate (HR, Legal, General Affair). Finance Manager membawahi 1 divisi, yaitu Staff Finance. Creative Director membawahi 2 divisi, yaitu Videographer dan Graphic Design. Genmedia membawahi 1 divisi Pimred yang terbagi menjadi Redaktur Pelaksana, Redaktur, Reporter, Kontributor, dan Social Media Specialist. Chief Production Officer dibagi menjadi 3, yaitu Lead Product (UX Research, UX Writing, dan UI/UX Design), Lead Engineer (Flutter Dev, Backend Dev, dan Frontend Dev), dan WP Developer. Chief Partnership Officer membawahi 3 divisi, yaitu Manager B2B, Manager B2G, Merchant Acquisition Manager. VP Marketing dibagi menjadi 4 bagian, yaitu Content Marketing Manager (Copywriter), Social Media Specialist, Digital Marketing (Ads Specialist), Customer Satisfaction & Acquisition (Customer Service dan Customer Acquisition). Divisi-divisi tersebut bekerja secara sinergis untuk mencapai visi dan misi dari GreatEdu, dengan setiap divisi memainkan peran penting dalam keberhasilan keseluruhan organisasi. Struktur yang terorganisir dengan baik tersebut memungkinkan GreatEdu untuk beroperasi secara efisien dan responsif terhadap kebutuhan pasar serta perkembangan industri pendidikan. Lebih jelasnya dapat dilihat pada gambar 2.2.



Gambar 2.2 Struktur Organisasi GreatEdu

Berdasarkan gambar 2.2, bagian/divisi yang menaungi kegiatan Studi Independen ini adalah divisi *Production*. Divisi *Production* secara umum merancang dan menghasilkan program-program yang sesuai dengan hasil riset. Dalam divisi ini, terdapat tiga sub-divisi yang bertanggung jawab atas berbagai aspek dari produksi program, yaitu *Lead Partnership*, *Lead Learning Design*, dan *Lead Learning Support*. *Lead Partnership* bertanggung jawab menjalin dan mengelola kemitraan strategis untuk mendukung pelaksanaan program. *Lead Learning Design* merancang materi dan kurikulum untuk program Studi Independen, memastikan kualitas konten edukatif. *Lead Learning Support*, terdiri dari *Quality Control Merchant* dan *Program Admin*, memastikan kelancaran program, melakukan pengawasan kualitas, serta menyediakan dukungan administratif.

BAB III

LANDASAN TEORI

3.1 Emisi Karbon

Emisi karbon adalah pelepasan gas karbon dioksida (CO₂) ke atmosfer akibat aktivitas manusia seperti pembakaran bahan bakar fosil, deforestasi, dan berbagai proses industri (PGN LNG Indonesia, 2023). Emisi karbon dianggap sebagai salah satu kontributor utama perubahan iklim karena CO₂ adalah gas rumah kaca yang memerangkap panas di atmosfer, menyebabkan peningkatan suhu global (Siregar & Hasbi, 2023). Oleh karena itu, peningkatan komponen CO₂ di atmosfer memiliki dampak yang signifikan terhadap perubahan pola cuaca, peningkatan frekuensi dan intensitas bencana alam, serta kerusakan ekosistem dan keanekaragaman hayati secara global.

Standard pengukuran emisi karbon yang sering digunakan dalam global yaitu ISO 14064. ISO 14064 menyediakan spesifikasi dan panduan untuk organisasi dalam mengukur dan melaporkan emisi gas rumah kaca yang terdiri dari tiga bagian: spesifikasi dan panduan untuk pengukuran dan pelaporan emisi gas rumah kaca, spesifikasi untuk validasi dan verifikasi data emisi, dan panduan untuk pengurangan emisi (Cano dkk., 2023). Lalu untuk satuan emisi karbon yaitu tCO₂ yang merupakan satuan paling umum dalam mengukur jumlah karbon dioksida yang dilepaskan ke atmosfer.

3.2 *Machine Learning*

Machine Learning adalah cabang dari kecerdasan buatan (*Artificial Intelligence*) yang berfokus pada pengembangan algoritma dan teknik yang memungkinkan komputer untuk belajar dari dan membuat prediksi atau keputusan berdasarkan data (Huarng dkk., 2023). Proses ini melibatkan pelatihan model pada *dataset* tertentu untuk mengenali pola atau hubungan dalam data tersebut. *Machine Learning* dibagi menjadi beberapa jenis, termasuk *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. Algoritma *machine learning* umum meliputi regresi linier, regresi logistik, pohon keputusan, jaringan saraf tiruan, dan lain-lain.

Implementasi dari masing-masing jenis *machine learning* yaitu, *Supervised learning*: terdapat klasifikasi untuk deteksi *spam email*, diagnosis penyakit, dan juga terdapat regresi untuk prediksi harga rumah, dan prediksi emisi karbon. Lalu *Unsupervised learning*: seperti segmentasi pelanggan, pengelompokan siswa, dll. Terakhir terdapat *Reinforcement learning*: untuk implementasi seperti *game playing*, robotika, dan sistem kendali.

3.3 *Regression*

Prediksi emisi karbon dapat dilakukan dengan menggunakan data historis yang didapat dari *Kaggle* dengan menggunakan satelit Sentinel-5P. Lalu jenis *machine learning* yang digunakan yaitu *Regression*, karena data historis tersebut beratribut *numeric* tanpa ada *categorical* pada *features* maupun *target*. Pada *project* kali ini, model yang digunakan adalah Random Forest Regressor. *Regression* adalah teknik statistik yang digunakan dalam *machine learning* untuk memodelkan dan menganalisis hubungan antara *variabel* dependen dan satu atau lebih *variabel*

independen (Halim & Bunyamin, 2023). Tujuan dari regresi adalah untuk memprediksi atau memperkirakan nilai *variabel* dependen berdasarkan nilai *variabel* independen. Jenis-jenis regresi yang umum digunakan meliputi *Linier Regression*, *Ridge Regression*, *Lasso Regression*, *XGBoost Regressor* dan *Random Forest Regressor*. *Linier Regression*, misalnya, mencari garis terbaik yang meminimalkan kesalahan kuadrat antara nilai prediksi dan nilai sebenarnya.

3.4 Python (*Data Science*)

Python adalah bahasa pemrograman yang banyak digunakan dalam bidang *data science* karena sintaksnya yang sederhana dan dukungan yang luas terhadap berbagai pustaka dan *framework* (Maki dkk., 2022). Pustaka Python seperti *NumPy*, *Pandas*, *Matplotlib*, *Seaborn*, dan *Scikit-Learn* menyediakan alat yang kuat untuk analisis data, visualisasi, dan penerapan algoritma *machine learning*. Python memudahkan para ilmuwan data / *data scientist* untuk melakukan pembersihan data, eksplorasi data, pembangunan model prediktif, dan evaluasi model. Popularitas Python dalam *data science* juga didukung oleh komunitas yang aktif dan sumber daya pembelajaran yang melimpah.

Selain itu, Python memiliki ekosistem yang kaya akan alat dan pustaka yang memfasilitasi pengolahan data dalam skala besar. Pustaka seperti *Dask* dan *PySpark* memungkinkan pengolahan data secara paralel dan distribusi, sehingga mampu menangani kumpulan data yang sangat besar dengan efisien. Dalam hal pemrosesan data teks atau tekstual, pustaka seperti *NLTK* dan *spaCy* menyediakan beragam fungsi untuk analisis teks dan pemrosesan bahasa alami (NLP). Di bidang *deep learning*, pustaka seperti *TensorFlow* dan *PyTorch* menyediakan kerangka kerja yang kuat untuk membangun dan melatih *model neural network* yang

kompleks. Kombinasi dari alat-alat ini membuat Python menjadi pilihan utama bagi para profesional *data science* yang ingin menangani berbagai jenis analisis data dan penerapan algoritma canggih.

3.5 *Streamlit*

Streamlit adalah *framework open-source* yang digunakan untuk membangun aplikasi web interaktif secara cepat dan mudah, khususnya dalam konteks *data science* dan *machine learning* (Putranto dkk., 2023). Dengan *Streamlit*, pengguna dapat membuat antarmuka *web* yang intuitif untuk memvisualisasikan data, memantau model *machine learning*, dan berbagi hasil analisis dengan mudah. *Streamlit* memungkinkan pengguna untuk menulis aplikasi *web* dengan kode Python sederhana, tanpa perlu pengetahuan mendalam tentang pengembangan *web*. Fitur-fitur seperti integrasi langsung dengan pustaka visualisasi Python, kemampuan untuk menangani *input* pengguna, dan pembaruan otomatis membuat *Streamlit* menjadi alat yang populer di kalangan *data scientist*.

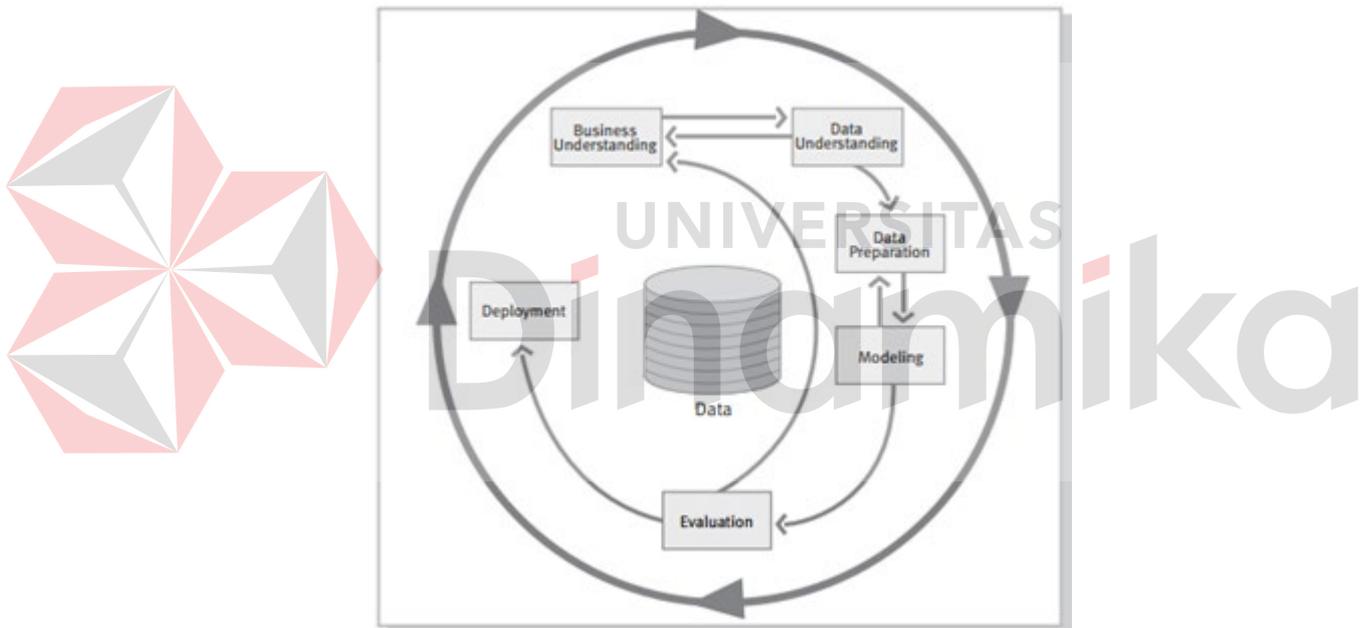
Selain itu, *Streamlit* juga mendukung berbagai *format input* dan *output*, seperti *file CSV*, gambar, dan video, sehingga memudahkan pengguna dalam berinteraksi dengan data mereka. *Framework* ini memungkinkan integrasi yang mulus dengan berbagai pustaka Python lainnya, seperti *NumPy*, *Pandas*, dan *Matplotlib*, sehingga analisis data dapat dilakukan dengan lebih efisien. *Streamlit* juga menawarkan dukungan untuk *deployment* aplikasi secara cepat ke berbagai *platform hosting*, sehingga hasil analisis dapat diakses oleh berbagai pemangku kepentingan dengan mudah.

BAB IV

DESKRIPSI PEKERJAAN

4.1 Metodologi Penelitian

Dalam proses pembuatan *data mining* atau *machine learning*, dapat menggunakan metodologi CRISP-DM seperti pada gambar 4.1, karena metodologi tersebut digunakan oleh mayoritas *data scientist* untuk membuat laporan *data science*.



Gambar 4.1 CRISP-DM

Ada beberapa tahap pada metodologi tersebut, yaitu: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*.

4.2 *Business Understanding*

Pada tahap ini, merupakan tahapan awal dari metodologi CRISP-DM untuk memahami tujuan bisnis dan persyaratan dari *Project* kali ini, yaitu untuk memprediksi emisi karbon. Tujuan dari tahap ini sebagai berikut:

- 1) Menentukan Tujuan Bisnis
 - a) Memahami dan memprediksi emisi karbon (CO₂) di Rwanda
 - b) Mengidentifikasi faktor-faktor yang mempengaruhi emisi karbon di Rwanda
- 2) Penilaian Situasi
 - a) Data emisi karbon di Rwanda didapat dari Kaggle berdasarkan hasil observasi data satelit Sentinel-5P
 - b) Tujuan utama adalah mengembangkan model regresi *machine learning* untuk memprediksi emisi karbon di Rwanda dengan akurasi yang baik.
 - c) Proyek ini dapat membantu pemerintah Rwanda dan pemangku kepentingan lainnya dalam perencanaan dan pengambilan keputusan terkait isu perubahan iklim.
- 3) Pengembangan Model
 - a) Mengembangkan model regresi *machine learning* yang dapat memprediksi emisi karbon di Rwanda dengan akurasi yang dapat diandalkan.
- 4) Rencana *Project*
 - a) Mengumpulkan dan mempersiapkan data yang didapat dari *dataset* Kaggle tentang prediksi emisi karbon di Rwanda.

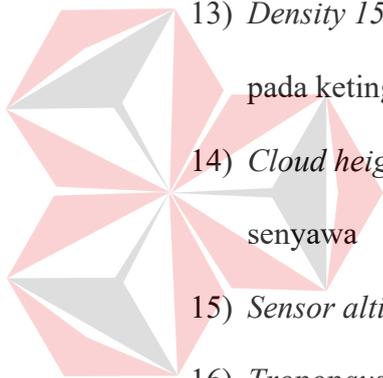
- b) Melakukan eksplorasi dan pemahaman data untuk mengidentifikasi pola, tren, dan hubungan antara *variabel*.
- c) Memilih dan menerapkan 1 *model* regresi *machine learning* dari 8 *model*, yaitu: *Linear*, *Lasso*, *Ridge*, *ElasticNet*, *RandomForest*, *SVR*, *XGBoost*, dan *LightGBM Regressor*. Berdasarkan hasil analisis, didapatkan *model* terbaik yaitu *RandomForest* yang akan digunakan sebagai *model* untuk memprediksi emisi karbon di Rwanda
- d) Mengevaluasi kinerja *model* untuk mencapai *error* yang lebih kecil
- e) Melakukan *deployment model* untuk dapat digunakan memprediksi emisi karbon di Rwanda.

4.3 Data Understanding

Pada tahap ini, melakukan identifikasi terhadap data yang digunakan. *Dataset* yang digunakan dalam *Project* ini diambil dari Kaggle menggunakan satelit observasi yaitu Sentinel-5P. Pada Rwanda, Afrika, dipilih sebanyak 497 lokasi beragam seperti daerah peternakan, kota, dan pembangkit listrik. Menggunakan satelit Sentinel-5P, diambil data emisi setiap lokasinya dari Januari 2019 hingga November 2021. Data awal terdiri dari 75 kolom dan 79.023 baris. Berikut beberapa kolom pada *dataset*:

- 1) *Latitude* : Koordinat Lintang dari lokasi yang diamati
- 2) *Longitude* : Koordinat bujur dari lokasi yang diamati
- 3) *Year* : Tahun pengukuran dilakukannya pengamatan
- 4) *Week No* : Nomor minggu dalam tahun saat pengukuran dilakukan
- 5) *Density* : Kepadatan atau massa jenis kolom jumlah senyawa di atmosfer

- 6) *Density Amf* : Kepadatan atau massa jenis dari sebuah senyawa yang telah dikoreksi menggunakan *Air Mass Factor* (AMF).
- 7) *Cloud Fraction* : Fraksi awan yang mempengaruhi pengukuran senyawa.
- 8) *Azimuth* : Sudut *azimuth* sensor saat mengukur senyawa
- 9) *Zenith* : Sudut *zenith* sensor saat mengukur senyawa
- 10) *Solar Azimuth* : Sudut *azimuth* matahari saat pengukuran senyawa
- 11) *Solar Zenith* : Sudut *zenith* matahari saat pengukuran senyawa
- 12) *Slant Column Number Density* : Kepadatan atau massa jenis sebuah senyawa dalam kolom atmosfer sepanjang lintasan miring.
- 13) *Density 15 Km* : Kepadatan atau massa jenis kolom jumlah senyawa pada ketinggian 15 Km
- 14) *Cloud height* : Tinggi awan yang mempengaruhi pengukuran senyawa
- 15) *Sensor altitude* : Ketinggian sensor saat mengukur senyawa
- 16) *Tropopause pressure* : Tekanan di *tropopause* saat pengukuran senyawa.
- 17) *Tropospheric* : Lapisan atmosfer bumi yang paling dekat dengan permukaan bumi.
- 18) *Cloud Top Pressure* : Tekanan di puncak awan.
- 19) *Cloud Top Height* : Ketinggian puncak awan.
- 20) *Cloud Base Pressure* : Tekanan di dasar awan.
- 21) *Cloud Base Height* : Ketinggian dasar awan.
- 22) *Cloud Optical Depth* : Kedalaman optik awan.
- 23) *Cloud Surface Albedo* : *Albedo* permukaan yang mempengaruhi pengukuran awan.



- 24) *Emission* : Emisi CO2 yang diukur pada lokasi dan waktu tertentu.
- 25) *Absorbing aerosol index* : Indeks *aerosol* yang menyerap dan mempengaruhi pengukuran senyawa.

4.4 Data Preparation

Pada tahap ini, melakukan persiapan terhadap data agar dapat digunakan pada tahap *modelling*. Pada tahap ini terdapat beberapa hal yang dilakukan untuk mempersiapkan data, yaitu: EDA (*Exploratory Data Analysis*), *Features Engineering*, *Select Features*, dan *Preparation*.

4.4.1 EDA

Tahap ini adalah proses awal dalam analisis data untuk memahami struktur, karakteristik, dan pola dalam *dataset*.

- Langkah pertama yaitu: *Statistical Description* yang isinya terdapat *basic description* dan *missing value*. Gambar 4.2 tersebut untuk menampilkan lima data teratas dari *dataframe train*.

```
[ ] 1 train.head()
latitude longitude year week_no SulphurDioxide_S02_column_number_density SulphurDioxide_S02_column_number_density
ID_LAT_LON_YEAR_WEEK
ID_0.510_29.290_2019_00 -0.51 29.29 2019 0 -0.000108 0.60
ID_0.510_29.290_2019_01 -0.51 29.29 2019 1 0.000021 0.72
ID_0.510_29.290_2019_02 -0.51 29.29 2019 2 0.000514 0.74
ID_0.510_29.290_2019_03 -0.51 29.29 2019 3 NaN
ID_0.510_29.290_2019_04 -0.51 29.29 2019 4 -0.000079 0.67
5 rows x 75 columns
```

Gambar 4.2 *Basic description train.head()*

- 2) Lalu untuk mengetahui atribut pada *dataframe*, dapat menggunakan *method* seperti pada gambar 4.3.

```
[14] 1 train.info()

<class 'pandas.core.frame.DataFrame'>
Index: 79023 entries, ID_-0.510_29.290_2019_00 to ID_-3.299_30.301_2021_52
Data columns (total 75 columns):
#   Column                                                                 Non-Null Count  Dtype
---  -
0   latitude                                                                79023 non-null  float64
1   longitude                                                                79023 non-null  float64
2   year                                                                    79023 non-null  int64
3   week_no                                                                79023 non-null  int64
4   SulphurDioxide_S02_column_number_density                            64414 non-null  float64
5   SulphurDioxide_S02_column_number_density_amf                      64414 non-null  float64
6   SulphurDioxide_S02_slant_column_number_density                    64414 non-null  float64
7   SulphurDioxide_cloud_fraction                                        64414 non-null  float64
8   SulphurDioxide_sensor_azimuth_angle                               64414 non-null  float64
9   SulphurDioxide_sensor_zenith_angle                                64414 non-null  float64
10  SulphurDioxide_solar_azimuth_angle                               64414 non-null  float64
11  SulphurDioxide_solar_zenith_angle                                 64414 non-null  float64
12  SulphurDioxide_S02_column_number_density_15km                    64414 non-null  float64
```

Gambar 4.3 Basic description *train.info()*

- 3) Lalu untuk mengetahui statistik sederhana (*mean*, standar deviasi, *min*, dan *max*) dari data, dapat menggunakan *method describe*. Pada gambar 4.4 tersebut, dapat dilihat bahwa terdapat beberapa *statistics method*, seperti: *Count*, *Mean*, *Standard Deviation*, *Min*, dan *Max* dari masing-masing kolom *dataset* tersebut. Hasilnya cenderung normal karena tidak ada anomali yang signifikan dari *method describe* tersebut.

```
[17] 1 train.describe()

   latitude  longitude  year  week_no  SulphurDioxide_S02_column_number_density  SulphurDioxide_S02_column_number_densit
count  79023.000000  79023.000000  79023.000000  79023.000000  64414.000000  64414.0
mean    -1.891072    29.880155    2020.000000    26.000000    0.000048    0.8
std      0.694522     0.810375     0.816502     15.297155    0.000272    0.1
min     -3.299000    28.228000    2019.000000    0.000000    -0.000996    0.2
25%     -2.451000    29.262000    2019.000000    13.000000    -0.000096    0.7
50%     -1.882000    29.883000    2020.000000    26.000000    0.000024    0.8
75%     -1.303000    30.471000    2021.000000    39.000000    0.000153    0.9
max     -0.510000    31.532000    2021.000000    52.000000    0.004191    1.8

8 rows x 75 columns
```

Gambar 4.4 Basic description *train.info()*

- 4) Pada gambar 4.5, terdapat beberapa kolom *missing value* dalam bentuk persentase. Kolom tersebut nantinya akan dilakukan *adjustment*.

```

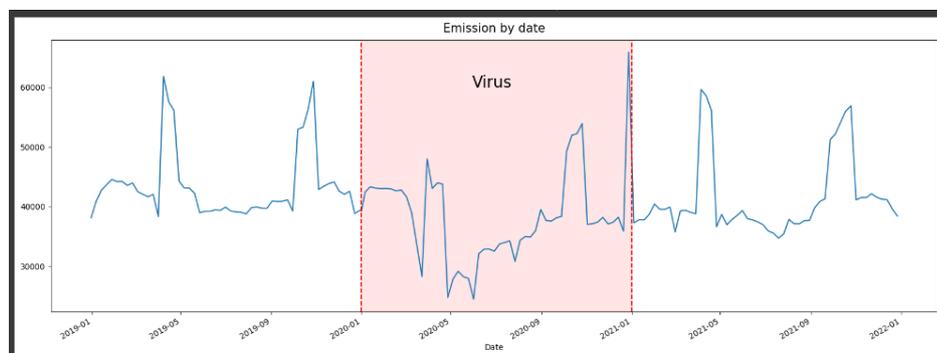
1 # Mengecek missing value train dalam bentuk persentase
2 missValTrain = train.isnull().sum()
3 missValTrain = missValTrain[missValTrain > 0].sort_values(ascending=False).head(20)
4 missValTrain = missValTrain / len(train)
5 missValTrain

```

UvAerosolLayerHeight_aerosol_pressure	0.994445
UvAerosolLayerHeight_solar_zenith_angle	0.994445
UvAerosolLayerHeight_aerosol_height	0.994445
UvAerosolLayerHeight_aerosol_optical_depth	0.994445
UvAerosolLayerHeight_sensor_zenith_angle	0.994445
UvAerosolLayerHeight_sensor_azimuth_angle	0.994445
UvAerosolLayerHeight_solar_azimuth_angle	0.994445
NitrogenDioxide_NO2_slant_column_number_density	0.231831
NitrogenDioxide_tropopause_pressure	0.231831
NitrogenDioxide_tropospheric_NO2_column_number_density	0.231831
NitrogenDioxide_absorbing_aerosol_index	0.231831
NitrogenDioxide_cloud_fraction	0.231831
NitrogenDioxide_sensor_altitude	0.231831
NitrogenDioxide_sensor_azimuth_angle	0.231831
NitrogenDioxide_sensor_zenith_angle	0.231831
NitrogenDioxide_solar_azimuth_angle	0.231831
NitrogenDioxide_stratospheric_NO2_column_number_density	0.231831
NitrogenDioxide_solar_zenith_angle	0.231831
NitrogenDioxide_NO2_column_number_density	0.231831
SulphurDioxide_SO2_column_number_density_amf	0.184870
dtype: float64	

Gambar 4.5 Missing value

- 5) Pada gambar 4.6, efek Covid-19 memberikan pengaruh yang sangat kuat di Q2 tahun 2020. Karena tahun ini memiliki tren yang unik, maka akan menyebabkan *overfitting* jika menggunakan data ini. Maka dari itu, akan dilakukan *adjustment*.



Gambar 4.6 Visualization of Emission by Date

- 6) Pada gambar 4.7, titik geografis pada *dataset*, didapatkan 497 titik.

```
2.3.1. Geographic Points of Emission  
  
[ ] 1 train.drop_duplicates(subset=['latitude', 'longitude']).shape[0]  
↔ 497
```

Gambar 4.7 Titik lokasi emisi

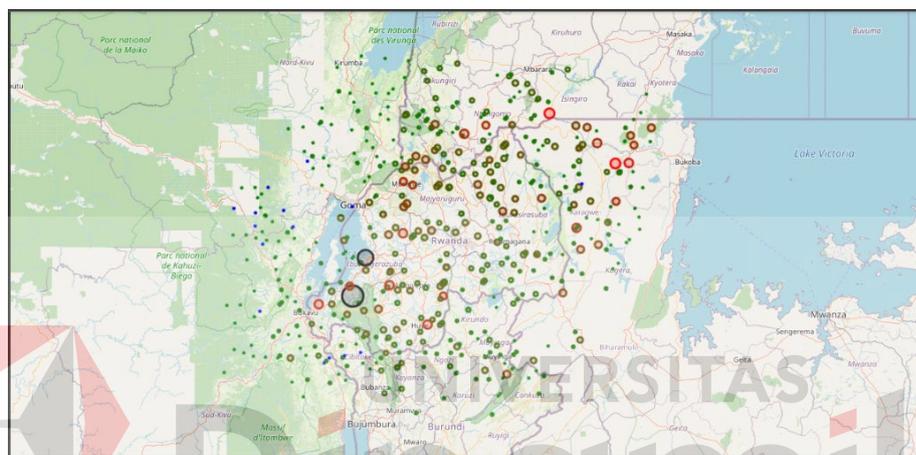
- 7) Pada gambar 4.8, didapati bahwa dari 497 titik geografis tersebut, terdapat 15 titik yang mempunyai emisi 0.

```
2.3.2. Location of Zero Emission  
  
[ ] 1 zero_emissions = train.groupby(['latitude', 'longitude'])['emission'].mean().to_frame()  
2 zero_emissions = zero_emissions[zero_emissions['emission'] == 0]  
3 zero_emissions
```

latitude	longitude	emission
-2.898	29.002	0.0
-2.859	29.041	0.0
-2.841	29.159	0.0
-2.817	29.283	0.0
-2.496	28.504	0.0
-1.974	28.526	0.0
-1.839	28.761	0.0
-1.833	28.467	0.0
-1.730	28.470	0.0
-1.712	28.688	0.0
-1.696	28.304	0.0
-1.682	29.218	0.0
-1.532	28.568	0.0
-1.505	30.995	0.0
-1.328	28.872	0.0

Gambar 4.8 Titik lokasi emisi 0

- 8) Pada gambar 4.9, merupakan titik-titik dari Lokasi emisi di Rwanda. Berdasarkan hal tersebut, terdapat 2 lokasi yang berwarna hitam, dapat diasumsikan bahwa ini memiliki emisi yang lebih tinggi dari yang lain (> 3000 tCO₂). Selain warna hitam, terdapat warna hijau dan merah. Warna hijau melambangkan jumlah emisi 0 – 1.000 tCO₂, sedangkan warna merah melambangkan jumlah emisi 1.001 – 3.000 tCO₂.



Gambar 4.9 Map Plot

- 9) Berdasarkan korelasi yang terdapat pada gambar 4.10, dapat diambil kesimpulan bahwa mayoritas sangat kecil hubungan keterkaitan antara *variabel features* dengan *target*, yakni kolom *emission*, akan tetapi ada beberapa *variabel* yang masih bisa diikuti seperti, *variabel Longitude* dan *Latitude* yang memiliki nilai tidak terlalu kecil, sehingga masih dapat berguna pada saat pemodelan. Selain itu, meskipun korelasi langsung dengan target tidak terlalu tinggi, *variabel* tersebut mungkin memiliki hubungan non-linear dengan *variabel* lain yang dapat dieksplorasi lebih lanjut. Oleh karena itu, *variabel Longitude* dan *Latitude* tetap penting untuk dipertimbangkan dalam proses pemodelan untuk memastikan prediksi yang lebih akurat dan informatif.

```

2.2.4. Correlation

[ ] 1 # Top 20 correlated features to the target
     2 # Pilih hanya kolom-kolom numerik
     3 numeric_train = train.select_dtypes(include=['float64', 'int64'])
     4
     5 # Hitung korelasi
     6 top20_corr = abs(numeric_train.corr()['emission']).sort_values(ascending=False).head(20)
     7 top20_corr

emission      1.000000
longitude      0.102746
UvAerosolLayerHeight_aerosol_height  0.069008
UvAerosolLayerHeight_aerosol_pressure  0.068138
Cloud_surface_albedo  0.046587
CarbonMonoxide_H2O_column_number_density  0.043217
CarbonMonoxide_CO_column_number_density  0.041328
Formaldehyde_tropospheric_HCHO_column_number_density_amf  0.040263
UvAerosolLayerHeight_aerosol_optical_depth  0.040156
UvAerosolLayerHeight_sensor_azimuth_angle  0.035142
NitrogenDioxide_solar_azimuth_angle  0.033417
Formaldehyde_tropospheric_HCHO_column_number_density  0.033333
SulphurDioxide_solar_azimuth_angle  0.032338
Formaldehyde_solar_azimuth_angle  0.030815
NitrogenDioxide_sensor_altitude  0.027540
UvAerosolLayerHeight_solar_azimuth_angle  0.027214
NitrogenDioxide_sensor_azimuth_angle  0.027104
CarbonMonoxide_solar_azimuth_angle  0.026282
SulphurDioxide_sensor_azimuth_angle  0.025079
Ozone_solar_azimuth_angle  0.024847
Name: emission, dtype: float64

```

Gambar 4.10 Pearson's Correlation

10) Berdasarkan hal-hal yang dilakukan dalam proses EDA, maka berikut adalah rangkumannya:

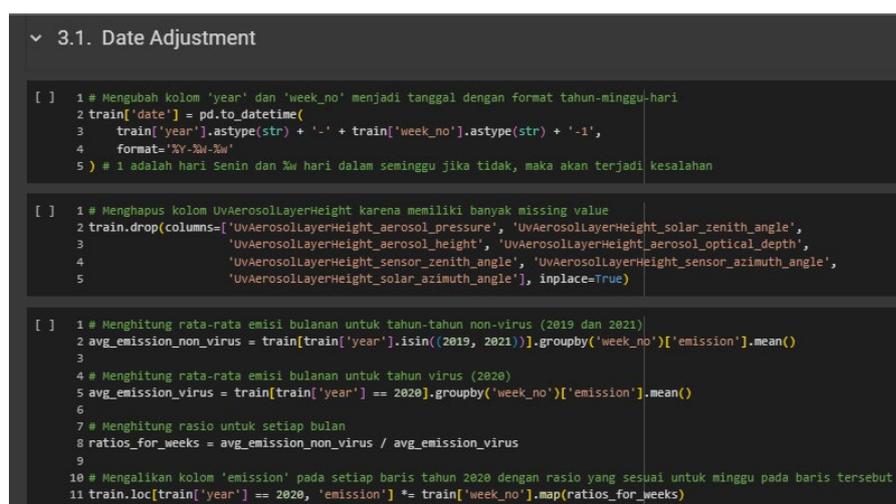
- Hanya ada fitur numerik
- Contoh pelatihan yang diberikan berubah-ubah dari tahun 2019 hingga 2021
- Pembacaan emisi CO₂ tunggal dapat berkisar dari 0,3 hingga hampir 3200
- Kelompok fitur *Uv Aerosol* memiliki tingkat nilai yang hilang yang sangat tinggi dan penyimpangan data yang signifikan
- Nilai emisi diambil dari 497 lokasi yang berbeda
- Emisi sangat bergantung pada tren musiman/liburan
- Terdapat 2 lokasi dengan emisi yang sangat tinggi yang tercatat

- h) Korelasi antara *variabel features* dengan *target* sangat kecil, sehingga perlu dilakukan *feature engineering* untuk mendapatkan kolom yang lebih berguna bagi model nantinya.

4.4.2 Features Engineering

Tahap ini bertujuan untuk merekayasa fitur agar lebih dapat menjadi lebih baik, sehingga fitur-fitur yang akan digunakan nantinya pada saat *modelling* dapat baik hasilnya.

- 1) Pada gambar 4.11, dilakukan proses *drop* kolom yang tadinya memiliki *missing value*, lalu juga melakukan *adjustment* pada tanggal karena tadi distribusi tren datanya tidak baik. Maksud dari *code python adjustment date* bertujuan untuk menyesuaikan emisi tahun 2020 dengan memperhitungkan rasio perbandingan dengan tahun-tahun non-virus (2019 dan 2021), sehingga menghasilkan data emisi yang mungkin lebih representatif atau sesuai dengan tren normal.



```

3.1. Date Adjustment

[ ] 1 # Mengubah kolom 'year' dan 'week_no' menjadi tanggal dengan format tahun-minggu-hari
2 train['date'] = pd.to_datetime(
3     train['year'].astype(str) + '-' + train['week_no'].astype(str) + '-1',
4     format='%Y-%w-%w'
5 ) # 1 adalah hari Senin dan %w hari dalam seminggu jika tidak, maka akan terjadi kesalahan

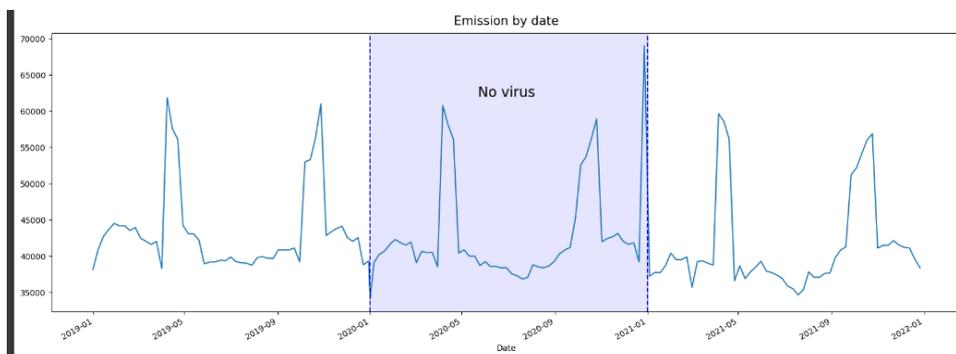
[ ] 1 # Menghapus kolom UvAerosolLayerHeight karena memiliki banyak missing value
2 train.drop(columns=['UvAerosolLayerHeight_aerosol_pressure', 'UvAerosolLayerHeight_solar_zenith_angle',
3     'UvAerosolLayerHeight_aerosol_height', 'UvAerosolLayerHeight_aerosol_optical_depth',
4     'UvAerosolLayerHeight_sensor_zenith_angle', 'UvAerosolLayerHeight_sensor_azimuth_angle',
5     'UvAerosolLayerHeight_solar_azimuth_angle'], inplace=True)

[ ] 1 # Menghitung rata-rata emisi bulanan untuk tahun-tahun non-virus (2019 dan 2021)
2 avg_emission_non_virus = train[train['year'].isin((2019, 2021))].groupby('week_no')['emission'].mean()
3
4 # Menghitung rata-rata emisi bulanan untuk tahun virus (2020)
5 avg_emission_virus = train[train['year'] == 2020].groupby('week_no')['emission'].mean()
6
7 # Menghitung rasio untuk setiap bulan
8 ratios_for_weeks = avg_emission_non_virus / avg_emission_virus
9
10 # Mengalikan kolom 'emission' pada setiap baris tahun 2020 dengan rasio yang sesuai untuk minggu pada baris tersebut
11 train.loc[train['year'] == 2020, 'emission'] *= train['week_no'].map(ratios_for_weeks)

```

Gambar 4.11 *Date adjustment & drop columns*

- 2) Hasilnya akan terlihat seperti gambar 4.12, bahwa distribusi tren datanya sudah lebih baik.



Gambar 4.12 *Visualization emission by date after adjustment*

- 3) Gambar 4.13 merupakan *feature engineering* untuk membuat kolom *season/musim*. Musim : Secara keseluruhan, Rwanda memiliki 4 musim : Musim hujan panjang Maret - Mei, Musim hujan pendek September - November, Musim kemarau panjang Juni - Agustus, dan Musim kemarau pendek Desember - Februari. Kita akan menambahkan fitur musim ini ke dalam data dari 1 hingga 4

```

3.2. Season Feature
[ ] 1 train['season'] = train['date'].dt.month.apply(lambda x: 1 if 3 <= x <= 5 else 2 if 6 <= x <= 8 else 3 if 9 <= x <= 11 else 4)

```

Gambar 4.13 *Season Feature*

- 4) Lalu menambahkan *feature* hari libur yang dapat dilihat pada gambar 4.14.

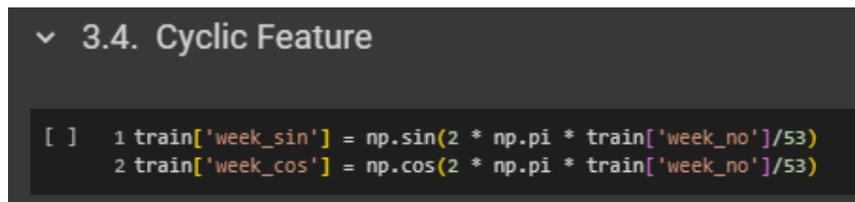
```

3.3. Holiday Feature
[ ] 1 train['holidays'] = (train['week_no'].isin([0, 51, 12, 30]))

```

Gambar 4.14 *Holiday Feature*

- 5) Pada gambar 4.15, merupakan langkah untuk menambahkan fitur siklik pada data. Sifat waktu yang siklik: Pada semua kasus, kecuali Tahun, fitur harus dibagi menjadi dua bagian: *sinus* dan *cosinus*, untuk mencerminkan siklus, misalnya tanggal 1 Januari mendekati tanggal 31 Desember.



```

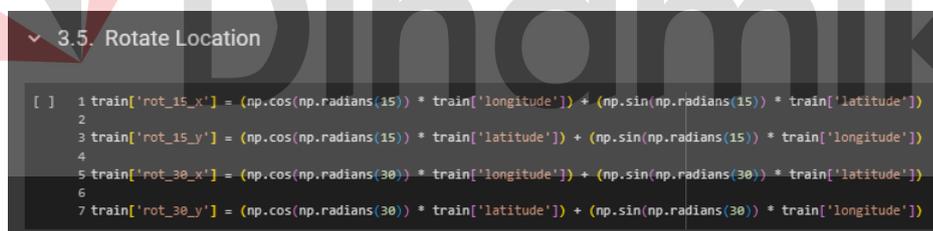
3.4. Cyclic Feature

[ ] 1 train['week_sin'] = np.sin(2 * np.pi * train['week_no']/53)
     2 train['week_cos'] = np.cos(2 * np.pi * train['week_no']/53)

```

Gambar 4.15 *Cyclic Feature*

- 6) Pada gambar 4.16, merupakan langkah untuk menambahkan fitur *rotate location* ke dalam data. Rotasi lokasi: Mengubah perspektif data dan menemukan pola atau hubungan baru yang tidak terlihat dalam data asli.



```

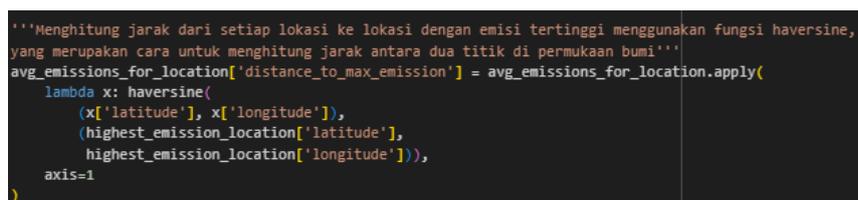
3.5. Rotate Location

[ ] 1 train['rot_15_x'] = (np.cos(np.radians(15)) * train['longitude']) + (np.sin(np.radians(15)) * train['latitude'])
     2
     3 train['rot_15_y'] = (np.cos(np.radians(15)) * train['latitude']) + (np.sin(np.radians(15)) * train['longitude'])
     4
     5 train['rot_30_x'] = (np.cos(np.radians(30)) * train['longitude']) + (np.sin(np.radians(30)) * train['latitude'])
     6
     7 train['rot_30_y'] = (np.cos(np.radians(30)) * train['latitude']) + (np.sin(np.radians(30)) * train['longitude'])

```

Gambar 4.16 *Rotate Location Feature*

- 7) Pada gambar 4.17, merupakan langkah untuk menambahkan fitur jarak Lokasi dengan emisi tertinggi.



```

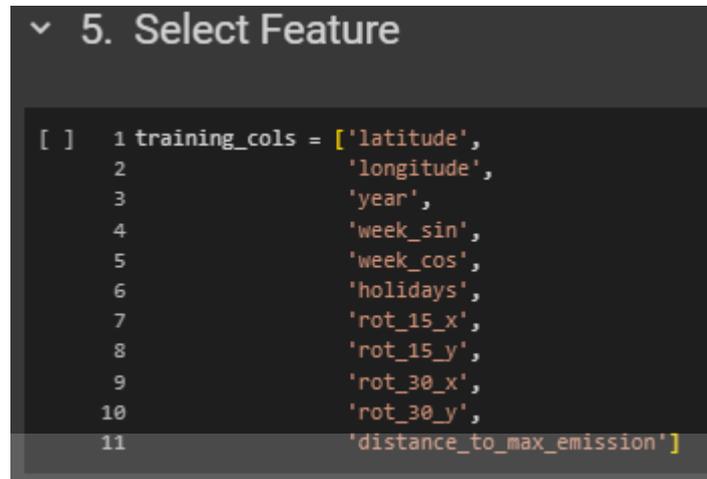
'''Menghitung jarak dari setiap lokasi ke lokasi dengan emisi tertinggi menggunakan fungsi haversine,
yang merupakan cara untuk menghitung jarak antara dua titik di permukaan bumi'''
avg_emissions_for_location['distance_to_max_emission'] = avg_emissions_for_location.apply(
    lambda x: haversine(
        (x['latitude'], x['longitude']),
        (highest_emission_location['latitude'],
         highest_emission_location['longitude'])),
        axis=1
)

```

Gambar 4.17 *Distance to max emission feature*

4.4.3 *Select Features*

Pada tahap ini, merupakan penyeleksian fitur yang akan digunakan ke dalam *modelling*.



```

5. Select Feature

[ ] 1 training_cols = ['latitude',
2      'longitude',
3      'year',
4      'week_sin',
5      'week_cos',
6      'holidays',
7      'rot_15_x',
8      'rot_15_y',
9      'rot_30_x',
10     'rot_30_y',
11     'distance_to_max_emission']

```

Gambar 4.18 *Select Features*

1) Fitur-fitur yang akan digunakan ke dalam *modelling* yaitu terdapat 11 kolom seperti pada gambar 4.18. Alasan menggunakan hanya 11 *features* tersebut, karena berdasarkan analisis pada bagian EDA, didapati bahwa *feature-feature* awal memiliki nilai *correlation* yang sangat kecil, sehingga dipilihlah *feature-feature* yang sudah dibuat pada bagian *feature engineering* dan juga beberapa *feature* awal yang nilai *correlation*-nya tidak terlalu rendah.

4.4.4 *Preparation*

Pada tahap ini, merupakan langkah terakhir dalam *Data Preparation*, yaitu persiapan data sebelum dilakukan *modelling*. Proses ini mencakup beberapa langkah penting untuk memastikan bahwa data siap digunakan dalam *model machine learning*.

- 1) Pada gambar 4.19, dilakukan proses *train-test split data* dengan rincian pembagian 70% *train* dan 30% untuk *test*. Lalu data di *transform* ke dalam bentuk diagram X dan y, yang di dalam sumbu X, kolom “*emission*” di drop atau dihapus, karena sumbu X hanya berisi fitur dan di sumbu y baru berisi target yaitu kolom “*emission*” saja.

```

6.1. Train-Test Split

[ ] 1 x = train.drop(columns='emission')
     2 y = train['emission']

[ ] 1 x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=rs)

```

Gambar 4.19 *Train-Test Split*

- 2) Setelah itu, dilakukan proses *scaling* data seperti pada gambar 4.20. Pertama, membuat *pipeline* dengan beberapa *parameter preprocessing*, yaitu *SimpleImputer* untuk mengisi nilai yang *missing* dengan *mean*, *QuantileTransformer* untuk merubah distribusi data menjadi distribusi normal, dan *StandardScaler* untuk menskalakan fitur agar rentang nilai dapat sama. Setelah itu, melakukan *fit pipeline* pada *X_train* dan *X_test*

```

6.2. Scaling

[ ] 1 # Buat pipeline dengan preprocessing
     2 pipeline = Pipeline([
     3     ('imputer', SimpleImputer(strategy='mean')), # Mengisi nilai yang hilang dengan mean
     4     ('quantile_transform', QuantileTransformer(output_distribution='normal')),
     5     ('scaler', StandardScaler()) # Scaling fitur
     6 ])

[ ] 1 # Preprocessing data pelatihan dan pengujian
     2 x_train = pipeline.fit_transform(x_train[training_cols])
     3 x_test = pipeline.transform(x_test[training_cols])

```

Gambar 4.20 *Scaling*

4.5 Modelling

Setelah melakukan tahap *Preparation*, maka dilanjutkan dengan tahap *Modelling*. Untuk model yang akan diujikan terdapat 8 *model*, yaitu *Linear*, *Lasso*, *Ridge*, *ElasticNet*, *RandomForest*, *SVR*, *XGBoost*, dan *LightGBM*.

- 1) Pada gambar 4.21, tahap *modelling* dilakukan dengan menggunakan *model* tanpa *hyperparameter tuning* untuk mengetahui model mana yang hasilnya lebih baik.



```
[ ] 1 # Inisialisasi model
2 models = {
3     "Linear": LinearRegression(),
4     "Lasso": Lasso(alpha=0.01),
5     "Ridge": Ridge(alpha=1.0),
6     "ElasticNet": ElasticNet(alpha=0.1, l1_ratio=0.5),
7     "RandomForest": RandomForestRegressor(n_estimators=100, random_state=42),
8     "SVR": SVR(kernel='linear'),
9     "XGBoost": XGBRegressor(),
10    "LightGBM": LGBMRegressor()
11 }
12
13 # Fit model ke data pelatihan dan evaluasi
14 results = []
15 for name, model in models.items():
16     model.fit(X_train, y_train)
17     y_pred = model.predict(X_test)
18     mae = mean_absolute_error(y_test, y_pred)
19     mse = mean_squared_error(y_test, y_pred)
20     rmse = np.sqrt(mse)
21     r2 = r2_score(y_test, y_pred)
22     results.append([name, mae, mse, rmse, r2])
23
24 # Membuat DataFrame untuk hasil evaluasi
25 results_df = pd.DataFrame(results, columns=['Model', 'MAE', 'MSE', 'RMSE', 'R2'])
26
27 # Cetak hasil evaluasi dalam bentuk tabel
28 results_df
```

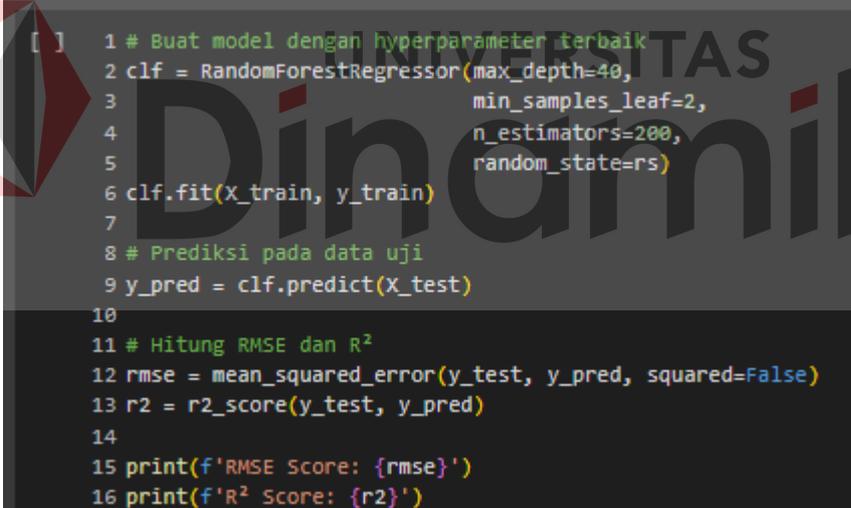
Gambar 4.21 *Modelling*

- 2) Berdasarkan gambar 4.22, dapat diambil Kesimpulan bahwa *model* yang baik adalah *RandomForest*, karena berdasarkan *evaluation metrics* MAE, MSE, dan RMSE, didapat bahwa hasilnya lebih baik karena mendekati 0. Lalu R2 *score* juga bagus karena mendekati 1.

	Model	MAE	MSE	RMSE	R2
0	Linear	69.422468	16021.019911	126.574168	0.187052
1	Lasso	69.442105	16020.545954	126.572295	0.187076
2	Ridge	69.421676	16020.893043	126.573666	0.187058
3	ElasticNet	67.899310	16033.901671	126.625044	0.186398
4	RandomForest	4.152948	455.361980	21.339212	0.976894
5	SVR	60.873443	17482.739510	132.222311	0.112880
6	XGBoost	8.318377	473.505151	21.760173	0.975973
7	LightGBM	15.539003	735.292181	27.116272	0.962689

Gambar 4.22 *Result of Modelling*

- 3) Pada gambar 4.23, *model* terbaik yaitu *RandomForest* dilakukan *hyperparameter tuning* untuk membuat *model* lebih baik lagi.



```
[ ] 1 # Buat model dengan hyperparameter terbaik
2 clf = RandomForestRegressor(max_depth=40,
3                             min_samples_leaf=2,
4                             n_estimators=200,
5                             random_state=rs)
6 clf.fit(X_train, y_train)
7
8 # Prediksi pada data uji
9 y_pred = clf.predict(X_test)
10
11 # Hitung RMSE dan R2
12 rmse = mean_squared_error(y_test, y_pred, squared=False)
13 r2 = r2_score(y_test, y_pred)
14
15 print(f'RMSE Score: {rmse}')
16 print(f'R2 Score: {r2}')
```

Gambar 4.23 *RandomForest with Hyperparameter Tuning*

Pada Gambar 4.23 tersebut, terdapat *hyperparameter* yang digunakan seperti *max_depth* sebesar 40, *min_samples_leaf* sebesar 2, *n_estimator* sebesar 200, dan *random_state* menggunakan *variabel rs*. Selanjutnya dilakukan *fit* ke *train* agar dapat diolah dan diuji dengan *test*.

4.6 Evaluation

Tahap ini merupakan bagian evaluasi dari model yang sudah dibuat. Evaluasi dapat menggunakan *evaluation metrics*, karena *Project* ini menggunakan *Regression*, maka *evaluation metrics* yang digunakan yaitu RMSE dan R^2 .

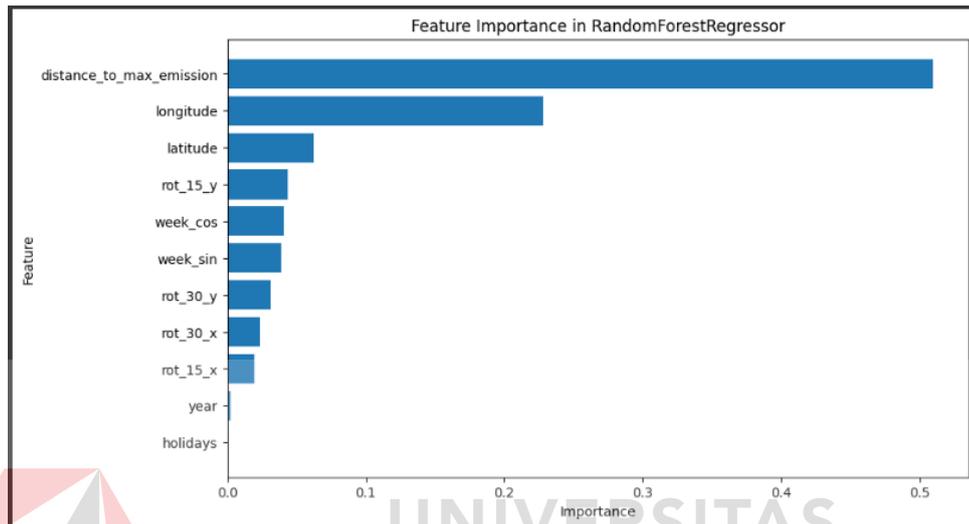
```
[39] ✓ 1.5s
... RMSE Score: 18.80903035592466
    R2 Score: 0.982048296680112
```

Gambar 4.24 *Evaluation Model*

- 1) Pada gambar 4.24, hasil dari *model RandomForest* yang telah menjalani *hyperparameter tuning* menunjukkan bahwa *model* ini memiliki kinerja prediksi yang sangat baik. Hal ini dibuktikan oleh nilai RMSE (*Root Mean Square Error*) yang rendah, yang menunjukkan bahwa kesalahan prediksi *model* minimal dan hasil prediksinya sangat mendekati nilai aktual. Selain itu, nilai R^2 (Koefisien Determinasi) yang hampir sempurna, mendekati angka 1, menandakan bahwa *model* mampu menjelaskan sebagian besar variabilitas dalam *data*, sehingga menunjukkan kecocokan yang baik terhadap *data* yang ada. Kombinasi dari nilai RMSE yang rendah dan R^2 yang tinggi mengindikasikan bahwa *model RandomForest* yang telah dioptimalkan ini memiliki tingkat akurasi dan presisi yang tinggi, menjadikannya alat prediktif yang andal untuk analisis *data* yang kompleks. Proses *hyperparameter tuning* yang dilakukan berhasil meningkatkan kinerja *model* secara signifikan,

sehingga *model* ini dapat diandalkan untuk aplikasi prediksi dengan tingkat akurasi yang tinggi.

- 2) Pada gambar 4.25, terdapat fitur-fitur yang penting untuk pemodelan *RandomForest*. Top 3 *feature importance*, yaitu :



Gambar 4.25 *Feature Importance of RandomForest*

- a) *Distance_to_max_emission* : Menunjukkan bahwa jarak ke lokasi dengan emisi tertinggi adalah fitur yang paling berpengaruh dalam model.
- b) *Longitude* : Menunjukkan bahwa koordinat geografis memainkan peran penting dalam prediksi.
- c) *Latitude* : Memiliki *importance* yang lebih rendah dibandingkan *longitude* tetapi masih cukup signifikan.

4.7 Deployment

Deployment merupakan tahapan terakhir dalam metodologi CRISP-DM. pada *Project* ini, tahapan *deployment* dilakukan menggunakan *Streamlit* untuk mendemokan prediksi emisi karbon.



Gambar 4.26 Streamlit bagian Deskripsi *Dataset*

- 1) Seperti pada gambar 4.26, dapat dilihat bahwa *dashboard Streamlit* terbagi menjadi tiga bagian utama, yaitu Deskripsi *Dataset*, EDA (*Exploratory Data Analysis*), dan Prediksi Emisi CO2. Bagian pertama dari *dashboard* ini, yaitu Deskripsi *Dataset*, memberikan gambaran menyeluruh tentang *dataset* yang digunakan dalam proyek ini. Deskripsi ini mencakup informasi *detail* mengenai kolom-kolom yang terdapat dalam *dataset*, seperti nama kolom, jenis data, serta penjelasan singkat tentang makna setiap kolom. Deskripsi yang komprehensif ini bertujuan untuk membantu pengguna memahami konteks data sebelum melanjutkan ke tahap analisis yang lebih mendalam. Deskripsi *dataset* memastikan bahwa pengguna memiliki *base* yang kuat untuk melakukan analisis lebih lanjut dan memahami hasil prediksi emisi karbon yang akan ditampilkan pada bagian akhir *dashboard.dataset*.



Gambar 4.27 Streamlit bagian EDA

2) Pada gambar 4.27, bagian EDA (*Exploratory Data Analysis*) dari dashboard menampilkan berbagai informasi data dan visualisasi dari data emisi karbon. Bagian ini dirancang untuk memberikan pengguna atau *user* wawasan mendalam tentang pola dan tren dalam *dataset*. Visualisasi yang disertakan mencakup grafik distribusi atau *histogram plot*, grafik *box* emisi CO2 per tahun, dan grafik garis yang membantu dalam memahami hubungan antara variabel-variabel dalam data. Misalnya, histogram dapat digunakan untuk menunjukkan distribusi emisi karbon, sedangkan *scatter plot* dapat menggambarkan hubungan antara emisi karbon dan faktor-faktor lain seperti *latitude*. Bagian EDA membantu pengguna untuk mengidentifikasi pola data yang signifikan, mengenali anomali, dan mendapatkan pemahaman yang lebih baik sebelum melanjutkan ke tahap prediksi emisi CO2.

PREDIKSI EMISI CO2 DI RWANDA

KELOMPOK SEKTE ULTI NOLAN

Deskripsi Dataset EDA Prediksi Emisi CO2

Prediksi Emisi CO2

Latitude-0.51 s/d -3.299
-3,30

Longitude 29.29 s/d 30.301
30,30

Year
2022

Week Number
7

Predict

Prediksi Emisi CO2: 159.18

Gambar 4.28 Streamlit bagian Prediksi Emisi CO2

- 3) Pada gambar 4.28, merupakan bagian menu utama yang menampilkan prediksi emisi karbon (CO₂) dengan memasukkan beberapa *inputan* dari *user*, seperti: *latitude*, *longitude*, *year*, dan *week number*. Pada gambar di atas, diprediksi bahwa terdapat 159,18 tCO₂ emisi CO₂ yang dilepaskan ke atmosfer. Berdasarkan data pada satelit Sentinel-5P, jumlah emisi karbon 159,18 tCO₂ dapat dikategorikan rendah, merujuk pada rentang dari nilai emisi pada data historis tersebut. Prediksi ini memberikan pengguna wawasan tentang tingkat emisi karbon di lokasi dan periode waktu tertentu, memungkinkan mereka untuk mengambil tindakan yang tepat untuk mengurangi dampak lingkungan.

BAB V

PENUTUP

5.1 Kesimpulan

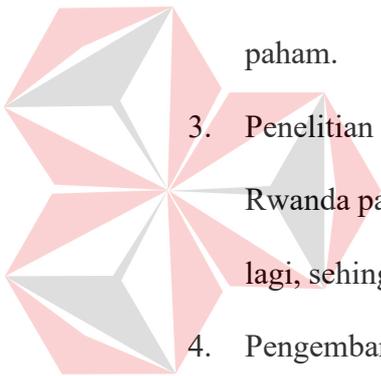
Berdasarkan hasil dari penerapan *model RandomForest Regressor* pada prediksi emisi karbon, maka dapat diambil kesimpulan sebagai berikut:

1. Model dapat melakukan prediksi emisi karbon dengan baik, ditinjau dari *evaluation metrics* RMSE 18% dan R^2 Score 0.98.
2. Informasi prediksi emisi karbon dapat ditampilkan dalam bentuk informasi dengan Streamlit.
3. Integrasi model prediksi dan Streamlit, pengguna dapat mengakses, memahami, dan menganalisis data emisi karbon. Hal ini memungkinkan pemantauan yang lebih efektif dan pengambilan keputusan yang lebih baik dalam upaya mengurangi dampak emisi karbon terhadap lingkungan.
4. Kombinasi dari model *RandomForest Regressor* dan Streamlit sebagai alat visualisasi memberikan solusi yang komprehensif untuk analisis data emisi karbon, mulai dari pemrosesan data hingga penyajian hasil secara interaktif, menjadikan keseluruhan sistem ini sangat bermanfaat bagi para peneliti dan praktisi di bidang lingkungan.

5.2 Saran

Dalam pembuatan *machine learning* untuk prediksi emisi karbon di Rwanda, penulis menyadari bahwa hasil akhir ini tidaklah sempurna. Maka saran yang dapat diberikan oleh penulis untuk pengembangan ini dapat dijelaskan sebagai berikut:

1. *Tuning hyperparameter* dengan model *RandomForest* dapat dikembangkan dan ditingkatkan lagi, sehingga hasilnya dapat semakin lebih baik. Hal tersebut dapat dilakukan dengan optimasi *hyperparameter* yang lebih mendalam, diharapkan *evaluation metrics* seperti RMSE semakin mendekati 0 dan R^2 semakin mendekati 1, yang akan meningkatkan akurasi dan keandalan model prediksi.
2. Tampilan *dashboard* streamlit dapat dibuat lebih menarik lagi serta lebih informatif, agar pengguna awam juga dapat menggunakan *dashboard* tersebut tanpa kesusahan dan kebingungan. Contohnya dapat ditambahkan Map Plot sehingga tidak perlu lagi ada *inputan* angka, sehingga *user* dapat langsung paham.
3. Penelitian ini hanya dapat digunakan untuk memprediksi emisi karbon di Rwanda pada tahun 2022 saja. Selanjutnya dapat dikembangkan lebih dinamis lagi, sehingga dapat digunakan untuk memprediksi tahun 2023 dan seterusnya.
4. Pengembangan lebih lanjut juga dapat mempertimbangkan integrasi dengan sumber data eksternal lainnya, seperti data cuaca atau aktivitas industri, untuk meningkatkan akurasi prediksi. Selain itu, penerapan model yang lebih kompleks, seperti *deep learning*, juga dapat dieksplorasi untuk mendapatkan hasil yang lebih baik.
5. Terakhir, uji coba dan validasi model pada berbagai skenario dan kondisi yang berbeda di Rwanda akan sangat penting untuk memastikan bahwa model tetap *robust* dan memberikan hasil yang akurat di berbagai situasi.



DAFTAR PUSTAKA

- Bengnga, A., & Ishak, R. (2022). Implementasi Seleksi Fitur Klasifikasi Waktu Kelulusan Mahasiswa Menggunakan Correlation Matrix with Heatmap. *Jambura Journal of Electrical and Electronics Engineering*, 4(2). <https://doi.org/10.37905/jjee.v4i2.14403>
- Cano, N., Berrio, L., Carvajal, E., & Arango, S. (2023). Assessing the carbon footprint of a Colombian University Campus using the UNE-ISO 14064–1 and WRI/WBCSD GHG Protocol Corporate Standard. *Environmental Science and Pollution Research*, 30(2). <https://doi.org/10.1007/s11356-022-22119-4>
- Fawwaz Al Maki, W., & Korespondensi, P. (2022). Metode Computational Thinking Untuk Pengabdian Masyarakat Dalam Peningkatan Kemampuan Bahasa Pemrograman Python Siswa Smk (Studi Kasus: Smk Asshiddiqiyah Karangpawitan, Garut) Computational Thinking Method for the Community Service in Improving Python Pro. *Jurnal Pengabdian Masyarakat Teknologi Informasi dan Informatika (DIMASLOKA)*, 1(1).
- Geevaretnam, J. L., Megat Mohd. Zainuddin, N., Kamaruddin, N., Rusli, H., Maarop, N., & Wan Hassan, W. A. (2022). Predicting the Carbon Dioxide Emissions Using Machine Learning. *International Journal of Innovative Computing*, 12(2). <https://doi.org/10.11113/ijic.v12n2.369>
- Halim, R., & Bunyamin, H. (2023). *Analisis Penjualan di Cabang Toko Serba Ada dengan Algoritma Machine Learning* (Vol. 5).
- Huarng, K. H., Yu, T. H. K., & Huang, D. H. (2023). Artificial intelligence in healthcare. Dalam *Philosophy of Artificial Intelligence and Its Place in Society*. <https://doi.org/10.4018/978-1-6684-9591-9.ch003>
- Kurniadi, F. I., & Larasati, P. D. (2022). Light Gradient Boosting Machine untuk Deteksi Penyakit Stroke. *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 6(1). <https://doi.org/10.47970/siskom-kb.v6i1.328>
- Larasati, N. (2023). PREDIKSI EMISI KARBON KENDARAAN PRIBADI DAN REKOMENDASI KENDARAAN ALTERNATIF MENGGUNAKAN MACHINE LEARNING DENGAN MODEL NEURAL NETWORK. Dalam *UNIVERSITAS NEGERI MALANG SABTU* (Vol. 8).
- Mardani, D. A. (2023). Wakaf dan Ekonomi Hijau: Upaya Mengurangi Perubahan Iklim dan Emisi Gas Karbon. *La Zhulma*, 2(1).
- Obagbuwa, I. C., Danster, S., & Chibaya, O. C. (2023). Supervised machine learning models for depression sentiment analysis. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1230649>

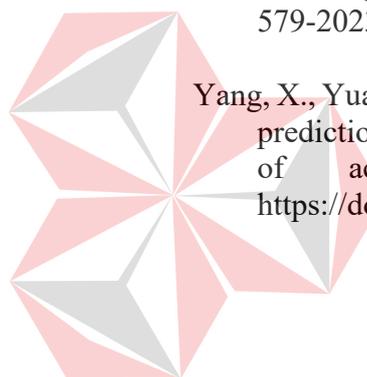
PGN LNG Indonesia. (2023). *Emisi Karbon: Pengertian, Penyebab, Dampak, dan Cara Mengurangnya*. November. Retrieved from: <https://pgnlng.co.id/berita/wawasan/emisi-karbon/>

Putranto, A., Azizah, N. L., & Astutik, I. R. I. (2023). Sistem Prediksi Penyakit Jantung Berbasis Web Menggunakan Metode Svm Dan Framework. *Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, 4(2).

Siregar, S. W., & Hasbi. (2023). Analisis Pengaruh Keterbukaan Perdagangan, Konsumsi Energi, dan Pertumbuhan Ekonomi terhadap Emisi Karbon di Negara D-8. *Jurnal Magister Ekonomi Syariah*, 2(1 Juni). <https://doi.org/10.14421/jmes.2023.021-05>

Van Der Woude, A. M., De Kok, R., Smith, N., Luijkx, I. T., Botía, S., Karstens, U., Kooijmans, L. M. J., Koren, G., Meijer, H. A. J., Steeneveld, G. J., Storm, I., Super, I., Scheeren, H. A., Vermeulen, A., & Peters, W. (2023). Near-real-time CO₂ fluxes from CarbonTracker Europe for high-resolution atmospheric modeling. *Earth System Science Data*, 15(2). <https://doi.org/10.5194/essd-15-579-2023>

Yang, X., Yuan, C., He, S., Jiang, D., Cao, B., & Wang, S. (2023). Machine learning prediction of specific capacitance in biomass derived carbon materials: Effects of activation and biochar characteristics. *Fuel*, 331. <https://doi.org/10.1016/j.fuel.2022.125718>



Dinamika