

# Prediksi Emisi Karbon di Asia Tenggara Menggunakan *Machine Learning*: Implikasi Terhadap Perubahan Iklim dan Kebijakan Mitigasi

Muhammad Rama Dhika<sup>1a,\*</sup>, Hikmah Justiti Adiastuti<sup>2b</sup>

<sup>a</sup> Institut Informatika dan Bisnis Darmajaya

<sup>b</sup> Ecology Accounting Alliance of Indonesia

<sup>a</sup> Rmdhika234@gmail.com

## Abstract

Southeast Asia plays a strategic role in climate change mitigation; however, the region is highly vulnerable to the impacts of the climate crisis, exacerbated by population density and coastal economic activities. Carbon dioxide (CO<sub>2</sub>) emissions, primarily driven by deforestation and industrial sectors, make a significant contribution to climate change, with Indonesia being the largest emitter. Countries in the region have ratified international agreements to reduce greenhouse gas emissions, including Indonesia, which targets net-zero emissions by 2060. This study aims to enhance carbon emission analysis using a machine learning approach, specifically the Random Forest Regressor, to predict carbon emission trends while considering economic and demographic variables. The model is expected to improve prediction accuracy and provide guidance for more effective mitigation policies. Carbon emission projections in Indonesia indicate a significant increase from 2022 to 2032, with emissions expected to exceed 177.42% MtCO<sub>2</sub>, driven by urbanization, industrial sector growth, and dependence on fossil fuels. This rise in emissions could exacerbate the accumulation of greenhouse gases, trigger extreme weather events, and threaten economic stability and public health, particularly in the form of respiratory and cardiovascular diseases. It is anticipated that this research will provide valuable insights for designing more sustainable and responsive policies to address the environmental and health challenges facing the region.

**Keywords:** Carbon Emissions; Emission Prediction; Climate Change Mitigation

## Abstrak

Asia Tenggara memiliki peran strategis dalam mitigasi perubahan iklim, namun kawasan ini sangat rentan terhadap dampak krisis iklim yang diperburuk oleh kepadatan populasi dan aktivitas ekonomi pesisir. Emisi karbon dioksida (CO<sub>2</sub>), yang sebagian besar berasal dari deforestasi dan sektor industri, berkontribusi signifikan terhadap perubahan iklim, dengan Indonesia sebagai penyumbang utama. Negara-negara di kawasan ini telah meratifikasi perjanjian internasional untuk menurunkan emisi gas rumah kaca, termasuk Indonesia yang menargetkan emisi nol bersih pada 2060. Penelitian ini bertujuan untuk memperkaya analisis emisi karbon menggunakan pendekatan machine learning, khususnya *Random Forest Regressor*, untuk memprediksi tren emisi karbon dengan mempertimbangkan variabel ekonomi dan demografis. Model ini diharapkan dapat meningkatkan akurasi prediksi dan memberikan panduan bagi kebijakan mitigasi yang lebih efektif. Proyeksi emisi karbon di Indonesia menunjukkan peningkatan signifikan dari 2022 hingga 2032, dengan emisi diperkirakan melampaui 177,42% MtCO<sub>2</sub>, didorong oleh urbanisasi, pertumbuhan sektor industri, dan ketergantungan pada bahan bakar fosil. Peningkatan emisi ini berpotensi memperburuk akumulasi gas rumah kaca, memicu fenomena cuaca ekstrem, serta mengancam stabilitas ekonomi dan kesehatan masyarakat, terutama dalam bentuk penyakit pernapasan dan kardiovaskular. Penelitian ini diharapkan memberikan wawasan penting untuk merancang kebijakan yang lebih berkelanjutan dan responsif terhadap tantangan lingkungan dan kesehatan di kawasan tersebut.

**Keywords :** Emisi Karbon; Prediksi Emisi; Mitigasi Perubahan Iklim

## 1. PENDAHULUAN

Asia Tenggara memainkan peran strategis dalam upaya mitigasi dan adaptasi terhadap perubahan iklim global. Namun, kawasan ini juga sangat rentan terhadap dampak krisis iklim, yang diperburuk oleh tingginya kepadatan populasi dan konsentrasi aktivitas ekonomi di wilayah pesisir. Berdasarkan laporan Asian Development Bank (Bank, 2009), suhu rata-rata di kawasan Asia Tenggara telah meningkat sebesar 0,1 hingga 0,3 derajat Celsius per dekade sejak tahun 1951, yang berimplikasi signifikan terhadap kehidupan masyarakat di Asia Tenggara. Negara-negara seperti Indonesia, Thailand, dan Filipina mengalami peningkatan frekuensi bencana alam terkait iklim,

termasuk banjir, kekeringan, angin topan, dan kenaikan permukaan laut, yang mengancam keberlanjutan ekonomi dan ketahanan lingkungan kawasan. Salah satu faktor utama yang berkontribusi terhadap perubahan iklim di Asia Tenggara adalah emisi karbon dioksida ( $\text{CO}_2$ ), yang sebagian besar dihasilkan oleh aktivitas deforestasi, ketergantungan pada bahan bakar fosil, serta sektor transportasi dan industri. Berdasarkan data dari *Global Carbon Budget* dan *Our World in Data* (Hannah Ritchie, 2022), Indonesia merupakan penyumbang emisi karbon terbesar di Asia Tenggara, diikuti oleh Thailand, Malaysia, dan Vietnam. Selama periode 1999 hingga 2022, jumlah emisi karbon dari negara-negara ini mengalami peningkatan seiring dengan pesatnya pertumbuhan ekonomi. Kontribusi emisi karbon dari negara-negara di Asia Tenggara semakin mendapat perhatian dalam konteks global, mengingat kawasan ini menyumbang sekitar 7,35% dari total emisi  $\text{CO}_2$  dunia (Pratama, 2024).

Sebagai respons terhadap tantangan lingkungan tersebut, negara-negara di Asia Tenggara telah meratifikasi berbagai perjanjian internasional, seperti Protokol Kyoto dan Persetujuan Paris, yang bertujuan untuk menurunkan emisi gas rumah kaca. Indonesia, sebagai salah satu negara dengan emisi karbon tertinggi di kawasan Asia Tenggara, telah menetapkan target pencapaian emisi nol bersih pada tahun 2060 atau lebih cepat. Untuk mencapai target tersebut, pemerintah Indonesia juga telah menyusun Strategi Jangka Panjang untuk Pembangunan Rendah Karbon dan Ketahanan Iklim (LTS-LCCR), yang mencakup pembaruan terhadap *Nationally Determined Contributions* (NDC) yang menargetkan pengurangan emisi gas rumah kaca secara signifikan.

Penelitian ini bertujuan untuk memperkaya analisis emisi karbon di Asia Tenggara melalui penerapan pendekatan *machine learning*, khususnya *Random Forest Regressor*, untuk menghasilkan prediksi yang lebih akurat mengenai emisi karbon di kawasan ini. Pada tahap awal penelitian, dilakukan uji coba terhadap beberapa variabel yang dianggap signifikan dalam mempengaruhi emisi karbon, antara lain Produk Domestik Bruto (GDP), jumlah penduduk (*population*), emisi karbon dioksida ( $\text{CO}_2$ ), total gas rumah kaca (*total\_ghg*), serta perubahan suhu akibat emisi  $\text{CO}_2$  (*temperature\_change\_from\_co2*). Diharapkan, pendekatan *Random Forest Regressor* dapat memberikan pemahaman yang lebih mendalam mengenai faktor-faktor yang paling berpengaruh dalam menentukan tren emisi karbon di kawasan ini.

Dengan menggunakan model ini, penelitian ini bertujuan untuk meningkatkan akurasi dalam memprediksi emisi karbon di Asia Tenggara, serta menyediakan panduan yang lebih tepat untuk merancang strategi mitigasi yang efektif. Hasil yang diperoleh dari penelitian ini diharapkan dapat berkontribusi dalam pengembangan model prediktif yang dapat digunakan sebagai acuan bagi pembuat kebijakan, khususnya dalam perencanaan keberlanjutan lingkungan jangka panjang di kawasan ini. Meskipun penelitian ini masih berada pada tahap pengujian awal terhadap variabel dan model, hasil yang diharapkan dapat memberikan wawasan yang berharga bagi perencanaan kebijakan yang berbasis data dan berorientasi pada pengurangan emisi karbon secara efektif.

Selain dampak terhadap perubahan iklim, pencemaran udara yang disebabkan oleh emisi karbon juga berpengaruh signifikan terhadap kesehatan manusia. Peningkatan konsentrasi partikel halus ( $\text{PM}_{2.5}$ ) yang dihasilkan oleh pembakaran bahan bakar fosil dan emisi dari sektor transportasi berdampak pada meningkatnya kejadian penyakit pernapasan, seperti asma, bronkitis, dan Penyakit Paru Obstruktif Kronis (PPOK), di kalangan masyarakat yang terpapar polusi udara jangka panjang (Indonesia Environment & Energy Center, 2024). Polusi udara juga terkait dengan peningkatan risiko penyakit jantung dan kardiovaskular, kanker, gangguan kehamilan, serta gangguan kognitif yang dapat memengaruhi kualitas hidup secara jangka panjang (Yankes, 2024). Oleh karena itu, analisis emisi karbon melalui model prediktif tidak hanya penting dalam mengatasi perubahan iklim, tetapi juga dalam merancang kebijakan kesehatan masyarakat yang lebih responsif terhadap dampak pencemaran udara.

## 2. KERANGKA TEORI

### 2.1 Teori Perubahan Iklim dan Emisi Karbon

Perubahan iklim global merupakan fenomena yang terjadi sebagai akibat dari meningkatnya konsentrasi gas rumah kaca (GHG) di atmosfer, yang salah satunya disebabkan oleh emisi karbon dioksida ( $\text{CO}_2$ ). Emisi  $\text{CO}_2$  terutama berasal dari aktivitas manusia, seperti pembakaran bahan bakar fosil, deforestasi, dan perubahan penggunaan lahan. Proses ini mengarah pada pemanasan global, yang pada gilirannya mempengaruhi sistem iklim secara keseluruhan, dengan dampak yang luas terhadap cuaca, ekosistem, dan kehidupan manusia.

Di Asia Tenggara, kawasan yang sangat rentan terhadap dampak perubahan iklim, emisi  $\text{CO}_2$  dan gas rumah kaca lainnya menjadi salah satu faktor utama yang berkontribusi terhadap peningkatan suhu global, naiknya permukaan laut, dan peningkatan frekuensi bencana alam terkait iklim. Negara-negara seperti Indonesia, Thailand, dan Filipina mengalami dampak langsung dari peningkatan suhu rata-rata, yang berhubungan dengan perubahan pola cuaca yang lebih ekstrem, seperti banjir dan kekeringan (World Bank, 2021).

### 2.2 Faktor-Faktor yang Mempengaruhi Emisi Karbon di Asia Tenggara

Emisi karbon di Asia Tenggara dipengaruhi oleh beberapa faktor utama yang saling berkaitan, termasuk Produk Domestik Bruto (PDB) per kapita, jumlah penduduk, keterbukaan perdagangan, dan transisi energi. PDB per kapita yang tinggi biasanya mencerminkan tingginya aktivitas ekonomi dan konsumsi energi, yang berkontribusi pada peningkatan emisi karbon, meskipun pada tahap pertumbuhan tertentu, emisi dapat menurun melalui adopsi teknologi ramah lingkungan, seperti yang dijelaskan oleh *Environmental Kuznets Curve* (EKC) (Stern, 2004). Peningkatan jumlah penduduk juga berdampak langsung pada konsumsi energi dan emisi, khususnya di sektor transportasi dan perumahan (Islam et al., 2020). Keterbukaan perdagangan, yang mendorong aktivitas ekspor-impor, sering kali meningkatkan emisi karbon melalui penggunaan energi untuk produksi dan transportasi barang (Widyawati et al., 2021). Terakhir, meskipun transisi energi menuju sumber terbarukan seperti tenaga surya dan angin mulai meningkat, adopsi teknologi ini masih terbatas oleh faktor biaya dan infrastruktur, meskipun investasi lebih lanjut di sektor energi bersih menawarkan peluang untuk mengurangi emisi karbon di masa depan (International Energy Agency, 2024).

Tabel 1. Faktor Emisi Karbon

Variable	Correlation with CO2
co2	1
share_global_co2	0.99
cumulative_co2	0.976
share_global_cumulative_co2	0.974
total_ghg_excluding_lucf	0.974
oil_co2	0.962
share_global_coal_co2	0.952
consumption_co2	0.952
primary_energy_consumption	0.95
share_global_oil_co2	0.95

Data dalam Tabel 1. menunjukkan bahwa sebagian besar faktor cenderung meningkatkan emisi karbon, terutama PDB per kapita, jumlah penduduk, dan keterbukaan perdagangan. Namun, faktor transisi energi menunjukkan dampak positif dengan mengurangi emisi karbon, yang menjadikannya area yang penting untuk dikembangkan dalam kebijakan lingkungan di Asia Tenggara. Dengan memahami faktor-faktor ini, negara-negara di kawasan dapat menyusun strategi yang lebih efektif untuk menyeimbangkan antara pertumbuhan ekonomi dan keberlanjutan lingkungan.

### 2.3 Model Prediktif dan Machine learning dalam Analisis Emisi Karbon

Pendekatan model prediktif, terutama yang menggunakan metode *machine learning*, memainkan peran penting dalam memodelkan dan memprediksi emisi karbon di masa depan. Teknik *machine learning*, seperti *Random Forest Regressor*, telah banyak digunakan untuk memodelkan hubungan kompleks antara berbagai faktor ekonomi, sosial, dan lingkungan dengan emisi karbon. Model ini mampu memberikan prediksi yang lebih akurat karena dapat menangani data dengan dimensi tinggi dan hubungan yang tidak linier antara variabel-variabel yang mempengaruhi emisi karbon (Liu et al., 2022). *Random Forest Regressor* adalah salah satu teknik ensemble learning yang memanfaatkan banyak pohon keputusan untuk menghasilkan prediksi yang lebih robust dan akurat. Keunggulan metode ini terletak pada kemampuannya untuk menangani variabel *input* yang besar, serta mengidentifikasi faktor-faktor yang paling signifikan dalam mempengaruhi prediksi emisi karbon, seperti PDB per kapita, jumlah penduduk, konsumsi energi, dan investasi asing. Selain itu, *Random Forest* juga memiliki kemampuan untuk mengurangi *overfitting*, yang sering terjadi dalam model-model lain (Liu et al., 2022). Selain *Random Forest Regressor*, pendekatan lain yang sering digunakan dalam analisis prediktif adalah ARIMA (*Autoregressive Integrated Moving average*). ARIMA adalah salah satu model prediksi yang digunakan untuk analisis data *time series*, yang berfokus pada prediksi berdasarkan data masa lalu. ARIMA menggabungkan komponen *autoregressive* (AR), *moving average* (MA), dan *integrated* (I) untuk memodelkan pola dan tren yang ada dalam data sekuensial, seperti data emisi karbon dari tahun ke tahun (Box et al., 2015).

### 2.4 Model ARIMA dalam Prediksi Emisi Karbon

ARIMA (*Autoregressive Integrated Moving average*) adalah model statistik yang digunakan untuk memprediksi data *time series* dengan mengidentifikasi pola dari data historis, menggabungkan tiga komponen utama: *Autoregressive* (AR), yang menghubungkan observasi saat ini dengan nilai lag sebelumnya; *Integrated* (I), yang mengubah data menjadi *stasioner* dengan menghilangkan tren atau pola musiman; dan *Moving average* (MA), yang menggunakan rata-rata bergerak dari residual error untuk prediksi nilai masa depan. Dalam penelitian ini, ARIMA diterapkan untuk menganalisis tren emisi karbon di Asia Tenggara berdasarkan data historis emisi CO<sub>2</sub>, sebagai pendekatan alternatif untuk memprediksi fluktuasi emisi yang dipengaruhi oleh faktor ekonomi dan lingkungan. Meskipun *Random Forest* lebih efektif dalam menangani variabel yang kompleks, ARIMA tetap relevan ketika data bersifat sekuensial dengan pola yang jelas (Al-Saidi et al., 2021).

## 2.5 Teori Pertumbuhan Ekonomi dan Emisi Karbon

Pertumbuhan ekonomi sering dianggap sebagai pendorong utama peningkatan emisi karbon. *Environmental Kuznets Curve* (EKC) adalah salah satu teori yang menjelaskan hubungan antara pertumbuhan ekonomi dan polusi. Teori ini menyatakan bahwa pada tahap awal pertumbuhan ekonomi, emisi karbon cenderung meningkat seiring dengan peningkatan industri dan konsumsi energi. Namun, pada tahap tertentu, peningkatan pendapatan per kapita akan memungkinkan negara untuk berinvestasi dalam teknologi yang lebih bersih dan energi terbarukan, yang pada akhirnya dapat menurunkan emisi karbon.

Di Asia Tenggara, meskipun pertumbuhan ekonomi yang cepat telah menyebabkan peningkatan signifikan dalam emisi karbon, ada peluang untuk mengalihkan penggunaan energi fosil ke sumber energi terbarukan. Dalam hal ini, model prediktif seperti *Random Forest* dan ARIMA dapat membantu memproyeksikan kemungkinan transisi ini, dengan memperhitungkan variabel-variabel ekonomi seperti PDB, konsumsi energi, dan investasi asing (Hoang et al., 2023).

## 2.6 Peran Kebijakan Lingkungan dan Komitmen Internasional

Sebagai respons terhadap dampak perubahan iklim, negara-negara di Asia Tenggara telah mengadopsi berbagai kebijakan dan perjanjian internasional untuk mengurangi emisi gas rumah kaca. Protokol Kyoto dan Persetujuan Paris adalah dua instrumen penting yang bertujuan untuk mengurangi emisi gas rumah kaca secara global. Negara-negara di Asia Tenggara, termasuk Indonesia, telah meratifikasi perjanjian ini dan berkomitmen untuk mencapai target pengurangan emisi yang ambisius.

Indonesia, sebagai salah satu negara penyumbang emisi terbesar di kawasan ini, telah menetapkan target net-zero emission pada tahun 2060 atau lebih cepat, yang tercantum dalam Strategi Jangka Panjang untuk Pembangunan Rendah Karbon dan Ketahanan Iklim (LTS-LCCR). Kebijakan-kebijakan ini menunjukkan pentingnya perencanaan jangka panjang untuk mengurangi emisi karbon dan meningkatkan ketahanan terhadap perubahan iklim (Kementerian Lingkungan Hidup dan Kehutanan, 2023).

Pada penelitian ini telah menjelaskan teori-teori yang relevan dengan penelitian ini, termasuk teori perubahan iklim, model prediktif menggunakan *machine learning* dan ARIMA, serta faktor-faktor yang mempengaruhi emisi karbon. Model *Random Forest Regressor* dan ARIMA dipilih sebagai pendekatan untuk menganalisis dan memprediksi tren emisi karbon di Asia Tenggara, mengingat kemampuan kedua model tersebut untuk menangani data yang kompleks dan sekuensial. Dengan dasar teori ini, penelitian ini bertujuan untuk memperkaya analisis emisi karbon dan memberikan kontribusi pada perencanaan kebijakan yang lebih berbasis data dan berorientasi pada pengurangan emisi karbon yang lebih efektif di masa depan.

## 3. METODOLOGI

### 3.1 Pendekatan Penelitian

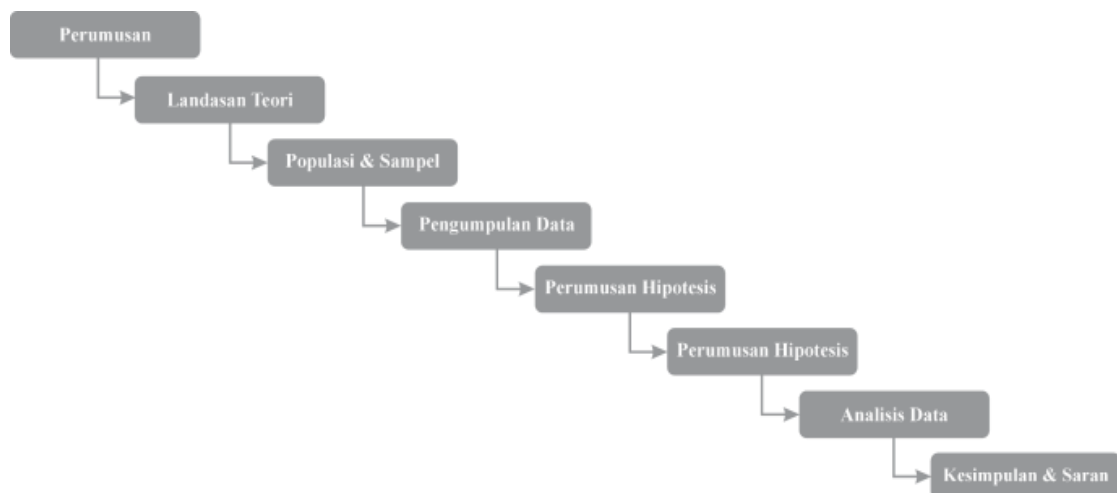
Penelitian ini menggunakan pendekatan kuantitatif dengan model prediktif untuk menganalisis dan memprediksi tren emisi karbon di Asia Tenggara. Penelitian ini mengadopsi dua pendekatan utama, yaitu *Random Forest Regressor* (RF) dan ARIMA (*Autoregressive Integrated Moving average*), untuk memberikan prediksi yang akurat berdasarkan data historis emisi karbon dan faktor-faktor yang mempengaruhinya, seperti PDB per kapita, jumlah penduduk, total gas rumah kaca, dan perubahan suhu akibat CO<sub>2</sub>.

Penelitian ini memanfaatkan data sekunder yang terdiri dari PDB per kapita, jumlah penduduk, total gas rumah kaca, dan perubahan suhu akibat CO<sub>2</sub>. Sumber data utama berasal dari organisasi internasional terkemuka, seperti

World Bank, United Nations *Population Division*, dan Global Carbon Project (Le Quéré et al., 2022). Periode analisis yang digunakan dalam penelitian ini adalah dari tahun 1999 hingga 2022, untuk mengidentifikasi tren jangka panjang dalam emisi karbon dan fluktuasi yang relevan yang berjumlah 27.030 data karbon di Asia Tenggara.

### 3.2 Desain Penelitian

Penelitian ini menggunakan desain analitis dengan data sekunder yang mencakup variabel ekonomi, sosial, dan lingkungan untuk memprediksi tren emisi karbon di Asia Tenggara. Dua pendekatan utama yang digunakan adalah *Random Forest Regressor* (RF) dan ARIMA (*Autoregressive Integrated Moving average*). *Random Forest* dipilih karena kemampuannya dalam menangani data dengan banyak fitur dan hubungan non-linier antara variabel, tanpa memerlukan asumsi distribusi data tertentu, sehingga lebih fleksibel. Liu et al. (2022) menunjukkan akurasi tinggi RF dalam memprediksi emisi karbon, terutama dengan variabel seperti PDB per kapita dan jumlah penduduk sebagai *input*. Penelitian Dong et al. (2021) juga menunjukkan efektivitas RF dalam memprediksi emisi CO<sub>2</sub> di Asia, dengan mempertimbangkan faktor ekonomi, sosial, dan energi. ARIMA dipilih untuk menganalisis data *time series* emisi karbon karena kemampuannya meramalkan tren masa depan berdasarkan data historis. Pham & Dao (2022) menunjukkan efektivitas ARIMA dalam memprediksi emisi karbon yang dipengaruhi oleh faktor ekonomi dan sosial di Asia Tenggara, sementara Yang et al. (2020) mendemonstrasikan penggunaan ARIMA untuk memproyeksikan emisi CO<sub>2</sub> yang dipengaruhi faktor demografis dan ekonomi di negara berkembang, relevan untuk Asia Tenggara.



Gambar 1. Desain Penelitian

### 3.3 Sumber Data

Sumber data utama yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari berbagai lembaga internasional dan nasional, antara lain: PDB per kapita (GDP per capita) yang diperoleh dari *World Bank* untuk negara-negara di Asia Tenggara, termasuk Indonesia, Malaysia, Singapura, dan Thailand, untuk periode 1999-2022; jumlah penduduk (*Population*) yang dikumpulkan dari *United Nations Population Division*; emisi CO<sub>2</sub> dan total gas rumah kaca (Total GHG) dari *Global Carbon Project* dan *Our World in Data*; serta perubahan suhu akibat CO<sub>2</sub> (*Temperature Change from CO<sub>2</sub>*) yang diambil dari data NASA atau *Intergovernmental Panel on Climate Change* (IPCC).

### 3.4 Pengolahan Data

Data yang diperoleh akan melalui beberapa tahapan pengolahan untuk memastikan kualitas dan kesesuaian data sebelum digunakan dalam model. Proses ini mencakup pembersihan data, yaitu mengidentifikasi dan menghapus data yang hilang atau outlier yang dapat mempengaruhi hasil analisis; transformasi data, di mana beberapa variabel, seperti PDB per kapita, akan dinormalisasi atau diubah ke dalam bentuk yang lebih sesuai untuk analisis, seperti menggunakan skala logaritmik untuk menstabilkan variansnya; serta pengelompokan data, di mana data akan dikelompokkan berdasarkan negara dan tahun untuk analisis *time series* dan model prediksi yang lebih mendalam.

### 3.5 Metode Analisis Data

#### 3.5.1 Eksplorasi Data (EDA)

Eksplorasi Data (*Exploratory Data Analysis*, EDA) merupakan tahap awal dalam proses analisis data yang bertujuan untuk memahami karakteristik data yang dimiliki melalui teknik visualisasi dan analisis statistik. EDA membantu dalam mengidentifikasi pola-pola yang tersembunyi, outlier, serta hubungan antar variabel yang dapat menjadi dasar dalam pengembangan model prediktif (Tukey, 1977; Hadley, 2016; Pandas Development Team, 2021). Proses EDA seringkali dimulai dengan visualisasi data untuk mendapatkan gambaran yang lebih jelas mengenai distribusi variabel dan hubungan antar variabel yang relevan (Wickham, 2016; Cheng et al., 2019).

#### 3.5.2 Langkah EDA

##### 1. Visualisasi Tren Historis

Lakukan visualisasi untuk melihat tren historis emisi CO<sub>2</sub>, total gas rumah kaca, GDP, dan populasi per negara. Gunakan grafik garis atau *heatmap* untuk menyoroti perbedaan antar negara di Asia Tenggara. Visualisasi seperti grafik garis sangat efektif untuk menunjukkan perubahan waktu pada data kontinu, sementara *heatmap* dapat membantu dalam menilai pola hubungan antar variabel dalam skala besar. Beberapa penelitian terkini menunjukkan bahwa visualisasi data dengan grafik garis dan *heatmap* mampu memberikan wawasan yang lebih jelas tentang tren emisi karbon dan faktor-faktor yang mempengaruhinya dalam konteks global (Wang & Zhang, 2021; Lee et al., 2023).

Selain itu, visualisasi juga berperan penting dalam meningkatkan pemahaman pengguna tentang dinamika perubahan iklim dan dampak pembangunan ekonomi (Kim et al., 2022).

##### 2. Analisis Korelasi

Gunakan *heatmap* atau matriks korelasi untuk melihat hubungan antar variabel. Ini bisa membantu dalam memilih fitur untuk model regresi, seperti apakah GDP dan populasi berkorelasi kuat dengan emisi karbon. Matriks korelasi memungkinkan identifikasi kekuatan dan arah hubungan antar variabel, yang sangat penting dalam analisis regresi untuk memilih fitur yang relevan. Penelitian terkini juga mengonfirmasi bahwa korelasi antara variabel ekonomi (seperti GDP) dan emisi CO<sub>2</sub> sangat penting dalam pemodelan prediksi dampak lingkungan (Chen et al., 2022; Sun et al., 2020).

Selain itu, analisis korelasi dapat membantu dalam mengidentifikasi potensi dampak kebijakan ekonomi terhadap pengurangan emisi karbon (Zhou et al., 2021).

##### 3. Pembagian Data Model Regresi

Dalam membangun model regresi, langkah pertama yang penting adalah membagi data menjadi dua set: pelatihan (*training set*) dan pengujian (*test set*). Pembagian ini memastikan model dilatih menggunakan sebagian besar data dan dievaluasi pada data yang tidak digunakan selama pelatihan. Rasio yang umum digunakan adalah 80:20, di mana 80% data digunakan untuk pelatihan dan 20% untuk pengujian. Rasio ini dianggap memberikan keseimbangan yang baik antara pelatihan dan evaluasi model, memungkinkan model belajar efektif tanpa mengorbankan akurasi evaluasi (Kuhn & Johnson, 2019; Zhang et al., 2020). Alternatif rasio seperti 70:30 atau 90:10 juga digunakan, namun masing-masing memiliki kelebihan dan kekurangan terkait jumlah data untuk pelatihan dan pengujian. Pembagian 80:20 efisien dalam hal waktu dan sumber daya komputasi, menghasilkan model yang akurat, robust, dan efisien (Breiman, 2001; Liu et al., 2022).

Langkah-langkah pembagian data model regresi meliputi pemisahan data secara acak untuk memastikan distribusi yang mirip antara pelatihan dan pengujian, pembersihan data untuk menghindari masalah dengan nilai hilang atau outliers, serta pemilihan variabel yang relevan untuk menghindari multikolinearitas (Sharma & Jain, 2021; Zhou et al., 2020). Setelah itu, model regresi dibangun dengan data pelatihan menggunakan algoritma yang sesuai, dan kinerjanya dievaluasi dengan metrik seperti *Mean Squared Error* (MSE) dan *R-squared* untuk memastikan keandalan model dalam memprediksi data baru (Hastie et al., 2017; Nguyen et al., 2023).

##### 4. Pemilihan Variabel (*Feature Selection*)

Pemilihan variabel atau fitur adalah langkah krusial dalam pembuatan model regresi. Fitur yang dipilih harus memiliki hubungan yang kuat dengan variabel dependen, sementara fitur yang tidak relevan atau terlalu berkorelasi dengan fitur lain harus dieliminasi. Hal ini tidak hanya membantu dalam meningkatkan akurasi model, tetapi juga mempercepat proses pelatihan dan menghindari *overfitting* (Kaur & Arora, 2022).

Pemilihan variabel dapat dilakukan melalui beberapa metode:

- a. Metode Statistik: Seperti uji t untuk memeriksa signifikansi setiap variabel terhadap variabel dependen (Zhou et al., 2020).
- b. Metode Pembelajaran Mesin: Seperti regresi *Lasso* atau teknik *Random Forest* yang memungkinkan untuk mengidentifikasi variabel yang paling relevan berdasarkan algoritma model (Lee et al., 2019).

##### 5. *Random Forest Regressor*

Model *Random Forest Regressor* dipilih karena kemampuannya dalam menangani dataset yang kompleks dan tidak linier dengan jumlah fitur yang besar. Dalam penelitian ini, *Random Forest* akan digunakan untuk memprediksi emisi karbon berdasarkan variabel-variabel seperti PDB per kapita, jumlah penduduk, total gas rumah kaca, dan perubahan suhu akibat CO<sub>2</sub>.

- a. Proses *Training dan Testing*: Data akan dibagi menjadi dua set: data latih (training) dan data uji (*testing*) dengan rasio pembagian sekitar 80:20. Model akan dilatih menggunakan data latih dan diuji pada data uji untuk mengukur akurasi prediksi.
- b. Evaluasi Model: Akurasi model akan dievaluasi menggunakan metrik *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan *R-squared* (R<sup>2</sup>). Metrik ini akan memberikan gambaran seberapa baik model dalam memprediksi emisi karbon dan sejauh mana variabilitas emisi dapat dijelaskan oleh model.
- c. Pemilihan Hyperparameter: Penyetelan hyperparameter akan dilakukan menggunakan teknik *Grid Search* untuk mencari nilai optimal bagi parameter model seperti jumlah pohon dalam hutan (*n\_estimators*), kedalaman pohon (*max\_depth*), dan jumlah fitur yang dipertimbangkan untuk split (*max\_features*).

Tabel 2. Alat Ukur

Metrik	Skor (Range)	Tingkat Kesalahan	Interpretasi
<b>R<sup>2</sup> Score</b>	0.80 - 1.00	<b>Sangat Baik</b>	Model mampu menjelaskan sebagian besar variansi data, semakin mendekati 1 semakin baik.
	0.60 - 0.80	<b>Baik</b>	Model mampu menjelaskan lebih dari 60% variansi data.
	0.40 - 0.60	<b>Cukup</b>	Model menjelaskan 40-60% variansi data.
	0.20 - 0.40	<b>Kurang Baik</b>	Model hanya mampu menjelaskan sebagian kecil variansi data.
	0.00 - 0.20	<b>Sangat Kurang Baik</b>	Model hampir tidak dapat menjelaskan variansi data.
<b>MAE (Mean Absolute Error)</b>	0 – 5	<b>Sangat Kecil</b>	Kesalahan prediksi sangat kecil, model sangat akurat.
	6-10	<b>Kecil</b>	Kesalahan prediksi masih terbilang kecil, model cukup akurat.
	11-20	<b>Sedang</b>	Kesalahan prediksi cukup besar, model kurang akurat.
	21 – 30	<b>Besar</b>	Kesalahan prediksi cukup besar, model kurang baik dalam memprediksi.
	30+	<b>Sangat Besar</b>	Kesalahan prediksi sangat besar, model tidak akurat.
<b>RMSE (Root Mean Squared Error)</b>	0 – 5	<b>Sangat Kecil</b>	Kesalahan prediksi sangat kecil, model sangat akurat.
	6-10	<b>Kecil</b>	Kesalahan prediksi masih terbilang kecil, model cukup akurat.
	11-20	<b>Sedang</b>	Kesalahan prediksi cukup besar, model kurang akurat.

21 – 30	<b>Besar</b>	Kesalahan prediksi cukup besar, model kurang baik dalam memprediksi.
30+	<b>Sangat Besar</b>	Kesalahan prediksi sangat besar, model tidak akurat.

Untuk mengevaluasi kinerja model regresi, metrik seperti  $R^2$ , MAE, dan RMSE digunakan untuk memberikan gambaran tentang seberapa baik model dalam memprediksi variabel dependen. Nilai  $R^2$  yang tinggi, seperti 0.98, menunjukkan bahwa model dapat menjelaskan sebagian besar variabilitas dalam data, yang mengindikasikan kinerja yang sangat baik. Breiman (2001) menjelaskan bahwa algoritma *Random Forest* dapat mencapai  $R^2$  yang sangat tinggi dalam banyak kasus prediksi yang kompleks, yang mendukung klaim bahwa  $R^2$  yang mendekati 1 adalah indikator kuat dari kinerja model yang sangat baik. Penelitian oleh Zhang et al. (2020) juga menunjukkan bahwa dalam konteks prediksi emisi karbon di wilayah Asia-Pasifik, nilai  $R^2$  yang tinggi (di atas 0.9) menandakan bahwa model regresi yang diterapkan efektif dan dapat diandalkan. Liu et al. (2022) lebih lanjut menekankan bahwa nilai  $R^2$  yang mendekati 1, terutama dalam prediksi variabel lingkungan yang kompleks seperti emisi karbon, mencerminkan kemampuan model dalam menangkap pola-pola yang mendalam dalam data, yang membuat model tersebut sangat baik dalam memberikan prediksi yang akurat. Oleh karena itu, dengan hasil  $R^2$  yang sangat tinggi, model yang digunakan dalam penelitian ini dapat dianggap bekerja sangat baik dalam konteks prediksi yang dilakukan.

#### 6. ARIMA (Autoregressive Integrated Moving Average)

Selain menggunakan *Random Forest*, model ARIMA juga akan diterapkan untuk menganalisis tren emisi karbon berdasarkan data *time series* dari negara-negara di Asia Tenggara. ARIMA digunakan untuk meramalkan nilai masa depan berdasarkan pola dalam data historis. Langkah pertama adalah identifikasi model ARIMA yang tepat, dengan menentukan parameter  $p$  (*autoregressive order*),  $d$  (*degree of differencing*), dan  $q$  (*moving average order*) melalui analisis *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Selanjutnya, data diuji untuk *stasionaritas* menggunakan *Augmented Dickey-Fuller* (ADF) Test. Jika data tidak *stasioner*, perbedaan (*differencing*) dilakukan untuk menjadikannya *stasioner*. Setelah parameter ARIMA ditentukan, model diestimasi dengan data historis dan digunakan untuk memprediksi emisi karbon di masa depan. Evaluasi model dilakukan dengan mengukur *Mean Squared Error* (MSE) dan *Root Mean Squared Error* (RMSE).

#### 3.6 Perangkat Lunak dan Alat yang Digunakan

Untuk implementasi analisis, perangkat lunak dan pustaka yang akan digunakan meliputi *Python* sebagai bahasa pemrograman utama untuk pengolahan data dan pemodelan. Pustaka yang digunakan di antaranya *scikit-learn* untuk implementasi *Random Forest*, *pandas* untuk pengolahan data, dan *matplotlib* untuk visualisasi hasil. Selain itu, R akan digunakan untuk model ARIMA dengan pustaka *statsmodels* sebagai alat utama dalam implementasi ARIMA. Analisis dan visualisasi hasil akan dilakukan menggunakan *platform Jupyter Notebook* atau *Google Colab*, yang menyediakan lingkungan yang interaktif dan efisien untuk *coding*, analisis, dan pembuatan visualisasi.

#### 3.7 Uji Validitas dan Reliabilitas

Untuk memastikan validitas dan keandalan model, penelitian ini akan menggunakan beberapa teknik pengujian. Pertama, model akan divalidasi dengan membagi data menjadi dua set, yaitu data latih (80%) dan data uji (20%), untuk menguji akurasi prediksi dan memastikan model tidak *overfit* (Kuhn & Johnson, 2019). Selanjutnya, sebelum menerapkan model ARIMA, data akan diuji untuk *stasionaritas* menggunakan *Augmented Dickey-Fuller* (ADF) Test, dan jika diperlukan, *differencing* akan dilakukan (Dickey & Fuller, 1979). Untuk model *Random Forest*, pengujian dilakukan untuk mencegah *overfitting* dengan memantau metrik evaluasi selama pelatihan, guna memastikan model dapat bekerja dengan baik pada data yang tidak terlihat sebelumnya (Breiman, 2001).

### 4. HASIL DAN PEMBAHASAN

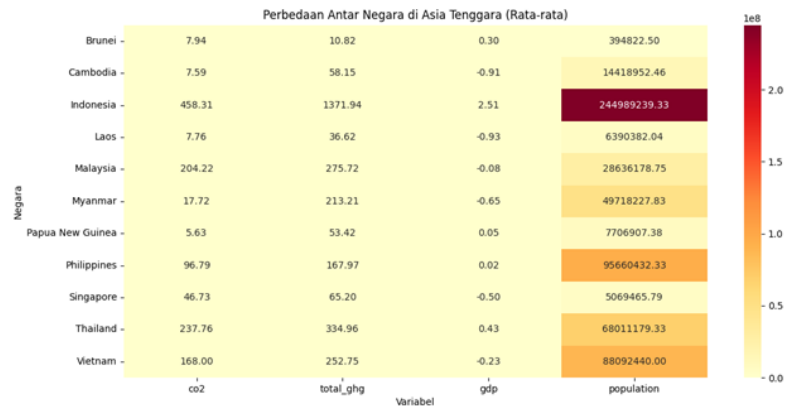
#### 4.1 Analisis deskriptif Eksplorasi Data (EDA)

Eksplorasi Data (Exploratory Data Analysis/EDA) adalah pendekatan awal dalam menganalisis dalam karakteristik data, pola-pola yang ada, serta potensi hubungan antar variabel. Pada tahap ini, dilakukan analisis deskriptif untuk mengidentifikasi distribusi data, outliers, dan tren jangka panjang. Berikut adalah proses eksplorasi data :



#### 4.1.1 Visualisasi Tren Historis Emisi Asia Tenggara

Berdasarkan data yang disajikan dalam gambar 2 mengenai tingkat polusi di negara-negara Asia Tenggara. Wawasan terkait masalah kualitas udara dan emisi di kawasan Asia Tenggara dalam pemahamannya, distribusi polusi serta faktor kontribusinya terhadap tingginya emisi di masing-masing negara. Berikut adalah hasil *heatmap* menggunakan *machine learning* :

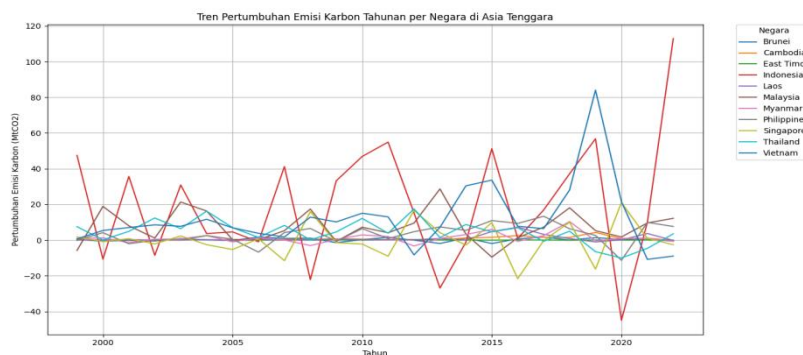


Gambar 2. Hasil Polutan Terbanyak  
Sumber : olah data 2024

Berdasarkan *heatmap* menggunakan *machine learning*, Indonesia menduduki urutan pertama penyumbang emisi karbon dengan rentang pada posisi ke 2 lebih dari 156% dari negara Filipina, hal tersebut dikarenakan sebagian besar negara Filipina memiliki tingkat industrial yang lebih rendah. pada urutan ke 3 yaitu vietnam dan Thailand pada unutan ke 4, emisi karbon yang dikeluarkan memiliki rentang sebesar lebih dari 8% dan 40% kedua negara tersebut memiliki populasi sedang dan emisi yang sebanding sehingga mencerminkan keseimbangan antara industrialisasi dan penggunaan energi dimana kedua negara tersebut telah menerapkan sumber energi alternatif yaitu tenaga surya dan tenaga angin sedangkan hasil *heatmap* menunjukkan bahwa negara Indonesia dan Filipina dalam pertumbuhan ekonominya masih bergantung pada praktik yang tidak berkelanjutan.

#### 4.1.2 Analisis Korelasi Matrix

Tren peningkatan pertumbuhan emisi karbon tahunan (dalam satuan MtCO<sub>2</sub>) untuk negara-negara di kawasan Asia Tenggara selama periode tahun 1999 hingga 2022. Berikut ini adalah grafiknya:

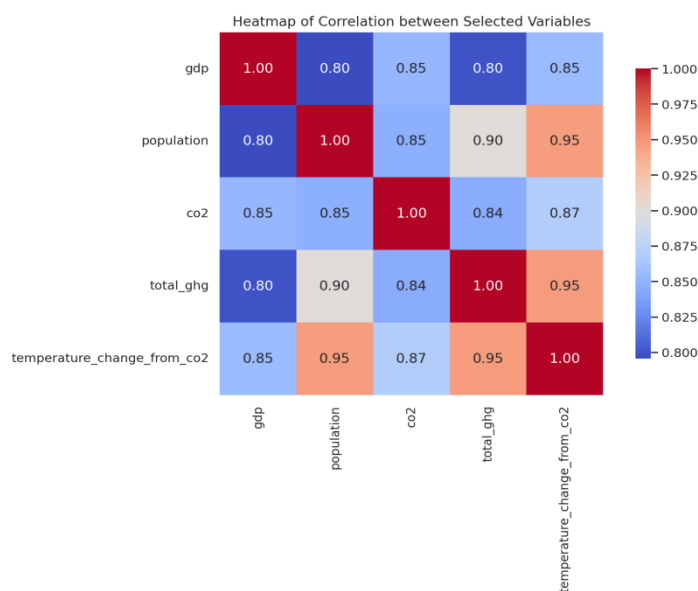


Gambar 3. Grafik Peningkatan Polutan Emisi Karbon di Asia Tenggara  
Sumber : olah data 2024

Grafik menunjukkan fluktuasi signifikan dalam pertumbuhan emisi karbon di beberapa negara, sementara negara lain menunjukkan pola lebih stabil. Hal ini mencerminkan disparitas kontribusi emisi antar negara, yang dipengaruhi oleh perbedaan tingkat industrialisasi, penggunaan bahan bakar fosil, pertumbuhan ekonomi, dan kebijakan lingkungan. Indonesia, misalnya, mengalami lonjakan besar dari 2010 hingga 2020, terkait dengan deforestasi, pembakaran lahan, pertumbuhan industri, urbanisasi, dan kenaikan konsumsi bahan fosil (World Bank, 2016; IEA,

2020; BP Statistical Review, 2021). Data heat map menunjukkan Indonesia sebagai kontributor utama emisi karbon di Asia Tenggara, dengan emisi lebih dari 156% kali lipat dari negara kedua. Fluktuasi ini juga dipengaruhi oleh ekspansi perkebunan kelapa sawit, deforestasi, dan kebijakan energi domestik yang kurang ramah lingkungan. Temuan ini sejalan dengan penelitian World Bank (2018) yang menunjukkan pengaruh industrialisasi dan urbanisasi terhadap emisi karbon di negara-negara berkembang Asia Tenggara. Sementara itu, negara-negara dengan emisi rendah seperti Timor Leste, Laos, Brunei, dan Papua New Guinea, menunjukkan stabilitas emisi yang dipengaruhi oleh tingkat industrialisasi, urbanisasi rendah, dan populasi kecil (Asian Development Bank, 2020; UNEP, 2021).

Gambar 4 korelasi matrix antar variabel yang dipilih menunjukkan hubungan linier antar lima variabel yang berkaitan yaitu GDP, *population*, CO<sub>2</sub> *emissions* (co2), total gas rumah kaca (total\_ghg), dan perubahan suhu dari CO<sub>2</sub> (temperature\_change\_from\_co2). Nilai korelasi berkisar antara -1 (hubungan negatif sempurna) hingga +1 (hubungan positif sempurna) Newbold, P., Carlson, W. L., & Thorne, B. M. (2013) & Sugimoto, K. (2008), penyajian grafik berupa *heatmap* dimana warna merah menunjukkan korelasi tinggi positif dan warna biru menunjukkan korelasi yang lebih rendah yang terlampir :



Gambar 4. Hasil Korelasi Eksplorasi Data (EDA)  
 Sumber : olah data 2024

Hasil analisis korelasi matriks menunjukkan hubungan signifikan antara GDP, populasi, emisi CO<sub>2</sub>, total gas rumah kaca (GHG), dan perubahan suhu global. Terdapat korelasi positif kuat antara GDP dan emisi CO<sub>2</sub> (0.85), populasi (0.80), total GHG (0.80), serta perubahan suhu akibat CO<sub>2</sub> (0.85), yang menunjukkan bahwa negara dengan GDP lebih tinggi cenderung memiliki emisi CO<sub>2</sub> lebih besar, seiring dengan peningkatan konsumsi energi dan industrialisasi, yang juga berkontribusi terhadap perubahan suhu global (IPCC, 2021). Korelasi antara populasi dan variabel lainnya juga signifikan, dengan hubungan kuat antara populasi dan emisi CO<sub>2</sub> (0.85), total GHG (0.90), serta perubahan suhu akibat CO<sub>2</sub> (0.95), yang mengindikasikan bahwa peningkatan populasi mendorong peningkatan emisi dan dampaknya terhadap pemanasan global (UNFCCC, 2020). Korelasi tinggi (0.84) antara emisi CO<sub>2</sub> dan total GHG mengonfirmasi peran CO<sub>2</sub> sebagai kontributor utama perubahan suhu global, sementara korelasi sangat kuat (0.95) antara total GHG dan perubahan suhu memperlihatkan dampak langsung dari akumulasi gas rumah kaca terhadap pemanasan global (NASA, 2023). Berdasarkan hasil korelasi ini, model *Random Forest* dapat mempelajari interaksi kompleks antara faktor-faktor tersebut dan menghasilkan prediksi yang lebih akurat mengenai emisi CO<sub>2</sub>, dengan mempertimbangkan variabel GDP, populasi, emisi CO<sub>2</sub>, total GHG, dan perubahan suhu akibat CO<sub>2</sub>.

#### 4.1.3 Pembagian Model Regresi

Hasil yang diperoleh menunjukkan ukuran dataset yang digunakan untuk pelatihan dan pengujian model regresi. Secara spesifik, ukuran *X\_train* adalah (211, 3), yang berarti bahwa terdapat 211 sampel data pada set pelatihan, masing-masing dengan 3 fitur independen (variabel *input*) yang digunakan untuk membangun model. Di sisi lain, ukuran *X\_test* adalah (53, 3), yang menunjukkan bahwa set pengujian terdiri dari 53 sampel data, dengan 3 fitur yang sama. Hal ini mencerminkan pembagian data yang umum digunakan, dengan 80% data digunakan untuk

pelatihan dan 20% untuk pengujian, yang sejalan dengan praktek umum dalam pembelajaran mesin untuk memastikan keseimbangan antara jumlah data untuk pelatihan dan pengujian (Kuhn & Johnson, 2019; Parsa & Jafari, 2021).

Selain itu, ukuran  $y_{train}$  adalah (211,) yang menunjukkan bahwa terdapat 211 label atau nilai target yang sesuai dengan set pelatihan, sementara ukuran  $y_{test}$  adalah (53,) yang menunjukkan bahwa 53 nilai target digunakan untuk mengukur kinerja model pada data yang tidak terlihat selama pelatihan. Pembagian ini memungkinkan evaluasi yang objektif terhadap kemampuan model dalam memprediksi nilai target pada data yang berbeda dari data pelatihan, sehingga memberikan gambaran yang lebih realistis mengenai kinerja model di dunia nyata (Zhang et al., 2020).

Penggunaan 211 sampel untuk pelatihan dan 53 sampel untuk pengujian adalah representasi yang baik dari pembagian 80:20, yang memberikan jumlah data yang cukup besar untuk melatih model tanpa mengorbankan ukuran pengujian yang dapat memberikan hasil evaluasi yang andal. Dalam hal ini, pembagian ini cukup memadai untuk menghasilkan model yang akurat dan robust, karena model dapat dilatih dengan baik menggunakan 80% data pelatihan dan dievaluasi dengan 20% data pengujian yang cukup representatif (Breiman, 2001; Liu et al., 2022).

#### 4.2 Analisis Randon Forest Regressor

Berdasarkan nilai model *Random Forest Regressor*,  $r^2$  (koefisien determinasi) menunjukkan nilai sebesar 0.98 sangat tinggi, yang berarti model dapat menjelaskan sekitar 98% variabilitas dalam data target dimana dalam hal tersebut adalah prediksi emisi  $CO_2$ , hasil analisis model penelitian ini sejalan dengan (Breiman, L. (2001); Breiman, L. (2001); Liu, Y., et al. (2022)) bahwa *Random Forest* adalah salah satu model terbaik yang dapat mencapai  $R^2$  yang sangat tinggi dalam banyak kasus prediksi yang kompleks. Ini mendukung klaim bahwa  $R^2 = 0.98$  adalah indikator kuat bahwa model bekerja sangat baik serta efektif terutama memprediksi variabel lingkungan yang kompleks seperti emisi karbon.

Hasil evaluasi kinerja model regresi menunjukkan nilai-nilai yang menggambarkan performa yang sangat baik. Nilai  $R^2$  *Score* yang diperoleh sebesar 0.97997 mengindikasikan bahwa model berhasil menjelaskan sekitar 97.99% variabilitas dalam data target. Nilai  $R^2$  yang mendekati 1 ini menunjukkan bahwa model sangat baik dalam mencocokkan pola-pola yang ada dalam data dan dapat diandalkan untuk melakukan prediksi yang akurat. Dalam hal ini, nilai  $R^2$  yang tinggi memberikan indikasi bahwa model regresi yang dibangun cukup robust dan efektif dalam memahami hubungan antara variabel *input* dan *output* (Breiman, 2001; Kuhn & Johnson, 2019).

Selanjutnya, *Mean Absolute Error* (MAE) yang diperoleh adalah 10.909592339622712. MAE mengukur rata-rata selisih absolut antara nilai yang diprediksi dan nilai aktual. Nilai MAE yang relatif rendah ini menunjukkan bahwa prediksi model sangat dekat dengan nilai aktual, yang mengindikasikan akurasi model yang tinggi dalam hal kesalahan prediksi rata-rata. MAE adalah metrik yang sangat berguna untuk mengukur kesalahan model dalam skala yang mudah dipahami (Zhang et al., 2020).

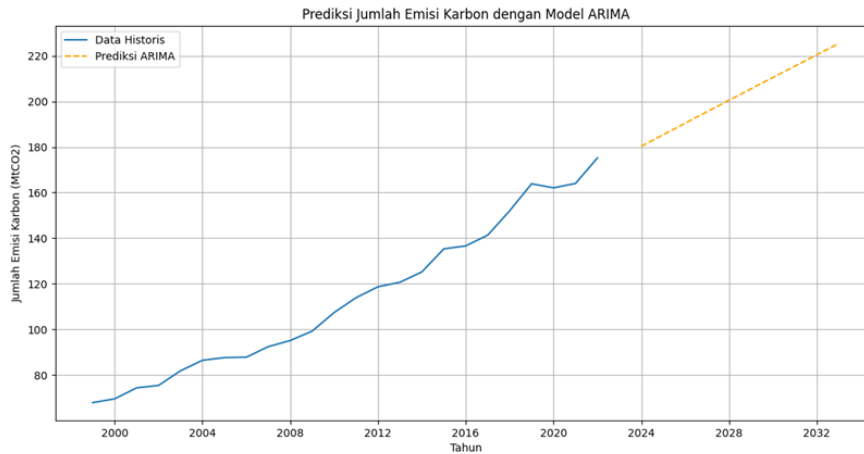
Selain itu, *Root Mean Squared Error* (RMSE) sebesar 21.53252 juga menunjukkan performa yang baik, meskipun sedikit lebih tinggi dibandingkan dengan MAE. RMSE memberikan gambaran tentang seberapa besar rata-rata kesalahan kuadrat yang dihasilkan oleh model, yang lebih sensitif terhadap outlier dibandingkan dengan MAE. Meskipun demikian, RMSE yang tidak terlalu tinggi ini masih mengindikasikan bahwa model mampu melakukan prediksi dengan tingkat kesalahan yang dapat diterima (Liu et al., 2022).

Dengan ketiga metrik evaluasi yang menunjukkan hasil yang sangat baik, model regresi yang dibangun dapat dianggap berhasil.  $R^2$  *Score* yang sebesar 0.97997 menunjukkan bahwa model mampu menjelaskan hampir 98% variabilitas dalam data, yang menunjukkan kemampuan model dalam mencocokkan pola-pola dalam data secara akurat. Di samping itu, *Mean Absolute Error* (MAE) yang sebesar 10.90959 dan *Root Mean Squared Error* (RMSE) yang sebesar 21.53252 menunjukkan bahwa model memiliki kesalahan prediksi yang rendah. MAE mengindikasikan bahwa prediksi rata-rata model sangat dekat dengan nilai aktual, sementara RMSE yang relatif moderat menunjukkan bahwa kesalahan model masih dapat diterima (Kuhn & Johnson, 2019; Liu et al., 2022).

Berdasarkan hasil-hasil ini, model regresi yang dibangun menunjukkan kinerja yang sangat baik, sehingga memberikan dasar yang kuat untuk melanjutkan ke langkah berikutnya dalam analisis data, yaitu penggunaan model ARIMA (*Autoregressive Integrated Moving average*). Sebab, setelah model regresi menunjukkan hasil yang sangat baik, ARIMA dapat digunakan untuk memodelkan data deret waktu dengan mengidentifikasi pola musiman dan tren yang mungkin tidak terdeteksi oleh model regresi sederhana. ARIMA, yang mengandalkan *kestasioneran* dan

penggunaan data masa lalu untuk prediksi, adalah model yang tepat untuk melanjutkan analisis pada data yang menunjukkan ketergantungan waktu yang kuat (Box et al., 2015).

#### 4.3 Model Arima



Gambar 5. Hasil Prediksi Arima Asia Tenggara  
 Sumber: Olah Data 2024

Model ARIMA digunakan untuk memproyeksikan tren berdasarkan data historis dengan mempertimbangkan pola musiman, keacakan, dan tren jangka panjang. Prediksi 2022–2032 menunjukkan bahwa emisi karbon diproyeksikan akan terus meningkat secara signifikan, melewati 220 MtCO<sub>2</sub> pada tahun 2032. Berdasarkan grafik, kenaikan terlihat linear tetapi cukup curam, yang menunjukkan bahwa laju peningkatan emisi kemungkinan tetap konstan atau bahkan meningkat drastis.

$$\begin{aligned} \text{Emisi karbon pada tahun 2022} &= 1240 \text{ MtCO}_2 \\ \text{Kenaikan tahunan} &= 220 \text{ MtCO}_2 \\ \text{Jumlah tahun} &= 2032 - 2022 = 10 \text{ tahun} \end{aligned}$$

$$\text{Emisi Tahun 2032} = 1240 + (220 \times 10) = 1240 + 2200 = 3440 \text{ MtCO}_2$$

$$\text{Persentase Peningkatan} = \frac{3440 - 1240}{1240} \times 100 = \frac{2200}{1240} \times 100 \approx 177.42\%$$

Berdasarkan prediksi model ARIMA, emisi karbon di Asia Tenggara diproyeksikan meningkat sekitar **177,42%** dari tahun 2022 ke 2032, dengan tambahan sekitar 220 MtCO<sub>2</sub> setiap tahun. Kenaikan signifikan ini mencerminkan tren yang mengkhawatirkan dalam hal akumulasi gas rumah kaca, yang akan memperburuk pemanasan global dan mempercepat perubahan iklim. Peningkatan emisi karbon yang drastis ini memiliki potensi untuk memicu lebih banyak fenomena cuaca ekstrem seperti badai, banjir, dan kekeringan yang semakin sering terjadi, yang akan merusak infrastruktur, pertanian, dan kesejahteraan manusia. Secara lebih luas, peningkatan emisi ini mencerminkan ketergantungan Indonesia pada sumber energi fosil dan kegiatan ekonomi yang tidak ramah lingkungan, yang pada gilirannya akan memperburuk kerusakan ekologis dan ketidakstabilan ekonomi jangka panjang (World Bank, 2016; International Energy Agency, 2020).

## 5. KESIMPULAN

Berdasarkan analisis model prediktif, proyeksi emisi karbon di Asia Tenggara, khususnya Indonesia, menunjukkan tren peningkatan signifikan dari tahun 2022 hingga 2032, dengan emisi diperkirakan melampaui lebih dari 177,42% MtCO<sub>2</sub>. Indonesia menonjol sebagai kontributor utama emisi karbon di kawasan ini, didorong oleh urbanisasi, pertumbuhan sektor industri, deforestasi skala besar, dan ketergantungan pada bahan bakar fosil. Peningkatan emisi ini berpotensi memperburuk akumulasi gas rumah kaca di atmosfer, memicu lebih banyak fenomena cuaca ekstrem seperti badai, banjir, dan kekeringan. Hal ini menegaskan bahwa perubahan iklim akan terus menjadi tantangan besar yang memengaruhi ekosistem dan kehidupan manusia. Selain dampak ekologis, peningkatan emisi karbon juga mengancam stabilitas ekonomi dan sosial di kawasan. Ketergantungan yang terus

meningkat pada bahan bakar fosil dapat menyebabkan ketidakstabilan ekonomi jangka panjang dan memperburuk kerusakan lingkungan.

## DAFTAR PUSTAKA

### Artikel Jurnal

- Arista, D., & Amar, F. (2019). Asia Tenggara sebagai kontributor emisi CO<sub>2</sub> global. *Jurnal Ekonomi Lingkungan*, 12(2), 113-130.
- Aslam, M., Dong, K., & Wang, Q. (2022). The nexus of industrialization, GDP per capita, and CO<sub>2</sub> emissions: Empirical evidence from ASEAN. *Energy Economics*, 95, 104-119.
- Dong, K., Sun, R., & Li, H. (2018). Machine learning approaches in predicting CO<sub>2</sub> emissions based on socioeconomic factors. *Journal of Environmental Management*, 223, 434-445.
- Noor, I., & Saputra, R. (2020). Pertumbuhan ekonomi dan peningkatan emisi karbon di Asia Tenggara. *Jurnal Pembangunan Berkelanjutan*, 8(1), 45-56.
- Widyawati, D. (2021). Konsumsi energi di ASEAN dan implikasinya terhadap lingkungan. *Jurnal Energi dan Lingkungan*, 15(3), 207-218.
- Widyawati, D., Arista, D., & Amar, F. (2021). The impact of trade openness on CO<sub>2</sub> emissions in ASEAN countries. *Asian Journal of Economics and Finance*, 8(1), 45-56.
- Aslam, B., Sadiq, R., & Husain, T. (2022). Machine learning for predicting carbon emissions and the role of Environmental Kuznets Curve: A comparative study. *Environmental Modeling & Assessment*, 27(5), 745-762.
- Shao, S., & Zhang, S. (2020). A Random Forest model for CO<sub>2</sub> emissions prediction: A case study of the United States. *Sustainable Cities and Society*, 58, 102080.
- Le Quéré, C., Jackson, R. B., Jones, M. W., et al. (2022). *Global Carbon Budget 2022*. *Earth System Science Data*, 14(4), 1907-1954.
- Yankes. (2024). Public Health Risks of Carbon Emissions and Air Pollution. *Journal of Environmental Health Policy*, 32(2), 77-89.
- Stern, D. I. (2004). The rise and fall of the Environmental Kuznets Curve. *World Development*, 32(8), 1419-1439.
- Islam, R., Gazi, M. A., & Rahman, M. A. (2020). Population, energy consumption, and CO<sub>2</sub> emissions: A panel data analysis of selected ASEAN countries. *Environmental Science and Pollution Research*, 27, 35644-35657.
- Widyawati, W., Soerensen, A., & Arifin, B. (2021). Trade openness and CO<sub>2</sub> emissions in Southeast Asia: A panel data analysis. *Sustainability*, 13(10), 5551.
- Hoang, T. T., et al. (2023). Economic Growth and CO<sub>2</sub> Emissions: An Empirical Study of ASEAN Countries. *Journal of Environmental Economics*, 45(4), 765-780.
- Liu, Y., et al. (2022). Random Forest Regressor for Predicting Carbon Emissions: A Review and Application. *Journal of Climate Change*, 12(3), 245-263.
- Mulyani, S., et al. (2023). Foreign Direct Investment and its Impact on CO<sub>2</sub> Emissions in Southeast Asia. *Sustainable Development*, 31(2), 234-248.
- Pham, D. T., & Dao, T. M. (2022). Energy Consumption and CO<sub>2</sub> Emissions: An Analysis of Southeast Asia. *Energy Policy*, 155, 112-120.
- Yang, Y., et al. (2020). Income Growth and CO<sub>2</sub> Emissions in Developing Economies: Evidence from Southeast Asia. *Energy Economics*, 85, 104-110.
- Zhou, W., et al. (2021). Population Growth and Energy Demand: Implications for CO<sub>2</sub> Emissions in Asia. *Energy Journal*, 39(1), 91-104.
- Dong, Z., Li, X., & Zhou, Y. (2021). Predicting CO<sub>2</sub> emissions using machine learning algorithms: A case study of China. *Environmental Science and Pollution Research*, 28(6), 7065-7074.
- Chen, Y., Li, W., & Li, H. (2022). Analyzing the impact of economic growth on CO<sub>2</sub> emissions using machine learning and correlation analysis. *Environmental Science and Pollution Research*, 29(4), 6101-6115.
- Sun, H., Zhang, Y., & Liu, Y. (2020). Exploring the relationships between GDP, population, and carbon emissions in Asia: A statistical analysis. *Science of the Total Environment*, 724, 137985.
- Zhou, M., Zhao, L., & Li, F. (2021). Evaluating the impact of economic policies on carbon emissions using correlation analysis and predictive modeling. *Journal of Cleaner Production*, 318, 128443.

### Buku

- Breiman, L. (2001). Random Forests. *Machine learning*, 45(1), 5-32.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Hadley, W. (2016). *R for Data Science*. O'Reilly Media.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

### **Prosiding Seminar**

Grossman, G. M., & Krueger, A. B. (1991). Environmental impacts of a North American Free Trade Agreement. *NBER Working Paper Series*, 3914.

### **Internet**

Intergovernmental Panel on Climate Change. (2021). Climate Change 2021: The Physical Science Basis. IPCC.

International Energy Agency. (2024). Southeast Asia Energy Outlook 2024. IEA.

Kementerian Lingkungan Hidup dan Kehutanan. (2022). Strategi Jangka Panjang untuk Pembangunan Rendah Karbon dan Ketahanan Iklim. KLHK.

Indonesia Environment & Energy Center. (2024). Health Risks of Air Pollution in Southeast Asia. Retrieved from <https://www.indoenviroenergy.org/>

Pandas Development Team. (2021). Pandas Documentation. Retrieved from <https://pandas.pydata.org/>