

PREDIKSI HARGA MINYAK KELAPA SAWIT MENTAH (CRUDE PALM OIL) MENGGUNAKAN METODE *DECISION TREE*

Didi Ciswanto¹, Samsoni²

¹Program Studi Teknik Informatika, Univesity Pamulalng, JL Puspitek Tangerang Selatan,
Indonesia, 15310.

e-mail:¹didiciswanto321@gmail.com

^{2,3}Program Studi Teknik Informatika, Univesity Pamulalng, JL Puspitek Tangerang Selatan,
Indonesia, 15310.

e-mail:²dosen0038@unpam.ac.id

Abstract

Crude palm oil (CPO) is one of Indonesia's main export commodities that has high economic value, but its price is very volatile due to various factors such as global market conditions, climate, and political policies. This research aims to build a CPO price prediction model using the Classification and Regression Tree (CART) algorithm, a form of Decision Tree method known for its ease of interpretation and high accuracy. The data used is historical CPO price data obtained from the Investing.com website within the last 10 years. In the process, the data is divided into training and testing data. The model is evaluated using metrics such as MAE (Mean Absolute Error), RMSE (Root Mean Square Error), and R² (R-Square). The results showed that the Decision Tree method with the CART algorithm was able to provide an accuracy of 93.101%, with an R² value of 0.937, RMSE of 0.062, and MAE of 0.046. This proves that the CART algorithm is suitable for predicting CPO prices and can be a decision-making tool for industry players.

Keywords: Palm Oil, Decision Tree, CART, Price Prediction, Machine Learning

Abstrak

Minyak kelapa sawit mentah (*Crude Palm Oil/CPO*) merupakan salah satu komoditas utama ekspor Indonesia yang memiliki nilai ekonomi tinggi, namun harganya sangat fluktuatif akibat berbagai faktor seperti kondisi pasar global, iklim, dan kebijakan politik. Penelitian ini bertujuan untuk membangun model prediksi harga CPO menggunakan algoritma *Classification and Regression Tree* (CART), salah satu bentuk metode Decision Tree yang terkenal karena kemudahannya dalam interpretasi dan akurasi yang tinggi. Data yang digunakan berupa data historis harga minyak kelapa sawit mentah yang diperoleh dari *website Investing.com* dalam rentang waktu 10 tahun terakhir. Dalam prosesnya, data dibagi menjadi data pelatihan dan data pengujian. Model dievaluasi menggunakan metrik seperti MAE (*Mean Absolute Error*), RMSE (*Root Mean Square Error*), dan R² (*R-Square*). Hasil penelitian menunjukkan bahwa metode *Decision Tree* dengan algoritma CART mampu memberikan akurasi sebesar 93.101%, dengan nilai R² sebesar 0.937, RMSE sebesar 0.062, dan MAE sebesar 0.046. Hal ini membuktikan bahwa algoritma CART cocok digunakan untuk memprediksi harga CPO dan dapat menjadi alat bantu pengambilan keputusan bagi pelaku industri.

Kata kunci: *Crude Palm Oil, Decision Tree, CART, Prediksi Harga, Machine Learning.*

1. PENDAHULUAN

Crude Palm Oil (CPO) adalah minyak yang dihasilkan dari proses pengolahan buah kelapa sawit yang kemudian diolah kembali untuk dibuat produk turunannya [1]. Pemanfaatan minyak kelapa sawit telah meluas ke berbagai kegunaan, di antaranya minyak masak, minyak industri, dan bahan bakar/biodiesel [2]. Minyak kelapa sawit atau *crude palm oil* (CPO) merupakan salah satu dari beberapa jenis minyak nabati yang paling banyak dikonsumsi, diproduksi dan diperdagangkan di seluruh dunia [3]. Permintaan minyak sawit diperkirakan akan terus meningkat seiring dengan pertumbuhan populasi dan ekonomi *global* [4].

Berdasarkan data United States Department of Agriculture (USDA), Indonesia dan Malaysia adalah produsen minyak sawit terbesar di dunia. Pada tahun 2020, produksi minyak kelapa sawit Indonesia mencapai sekitar 46 juta ton, sementara Malaysia menghasilkan sekitar 19 juta ton. Kedua negara ini bersama-sama menyumbang lebih dari 85% produksi minyak kelapa sawit *global*. luas perkebunan kelapa sawit di Indonesia mencapai sekitar 16 juta hektar, sedangkan di Malaysia mencapai sekitar 5 juta hektar. [5]. Harga komoditi CPO saat ini masih mengacu pada bursa komoditi di Rotterdam dan Malaysia. Pada dasarnya harga komoditi sering berfluktuasi karena ketergantungan pada faktor-faktor yang sulit dikuasai seperti perubahan musim, bencana alam dan lain-lain maka dari itu dengan kegiatan lindung-nilai menggunakan Kontrak Berjangka, para pelaku usaha [6].

Karena meningkatnya permintaan minyak sawit *global*, produsen lain seperti Indonesia dan Nigeria telah bergabung dengan pasar di mana Malaysia adalah pemain dominan. Gelombang pendatang baru di pasar minyak sawit menyebabkan persaingan yang semakin ketat untuk komoditas ini. Akibatnya, indeks harga minyak sawit mentah bervariasi sehingga menyebabkan ketidakstabilan harga komoditas ini. Harga minyak kelapa sawit telah berfluktuasi selama tiga dekade terakhir. Harganya lebih tidak stabil di seluruh dunia dibandingkan dengan komoditas lain seperti kakao, kedelai, dan karet [7].

Berdasarkan fenomena yang terjadi, perlu dilakukannya peramalan harga kelapa sawit untuk mengetahui fluktuasi harga di pasar Internasional. Diantara memprediksi harga minyak kelapa sawit mulai dari algoritma *Linear Regression* dan *Random forest* [8]. *Autoregressive Integrated Moving Average* (ARIMA) [9]. *Decision Tree* juga bagus untuk memprediksi [10]. Metode-metode ini melatih dan menguji data kemudian dinilai dengan metrik RMSE, MAE, dan R2.

Metode yang digunakan dalam penelitian ini adalah Metode *Decision Tree* dengan menggunakan algoritma *Classification and Regression Tree* (CART) dalam memprediksi harga minyak kelapa sawit mentah. Metode *Decision Tree* merupakan salah satu cara data *processing* dalam memprediksi masa depan dengan cara membangun klasifikasi dan regresi model dalam bentuk struktur pohon. *Decision Tree* adalah model visual untuk memungkinkan memahami proses pembuatan keputusan yang terstruktur, bertahap dan secara rasional [11]. Metode ini terkenal karena kemudahan pemahaman dan interpretasinya, karena hasilnya disajikan dalam bentuk visual yang sederhana. Selain itu, pohon keputusan dapat menangani data numerik dan kategori, serta mampu menghadapi *missing values*, sehingga menghasilkan model yang *robust*. [12].

2. METODE

Decision Tree

Decision Trees merupakan metode pembelajaran yang serbaguna dan nonparametrik yang diawasi yang digunakan secara luas dalam pembelajaran mesin untuk tugas klasifikasi dan regresi. Keahlian mereka terletak pada pembuatan model prediktif yang didasarkan pada aturan keputusan yang tidak rumit yang diekstrak dari data atribut, dengan demikian memastikan interpretasi yang jelas dan meniadakan kebutuhan untuk penskalaan fitur. *Decision tree* dengan mahir mengkategorikan data dan memperkirakan dalam segmen yang berbeda, membuktikan keampuannya dalam berbagai skenario [13].

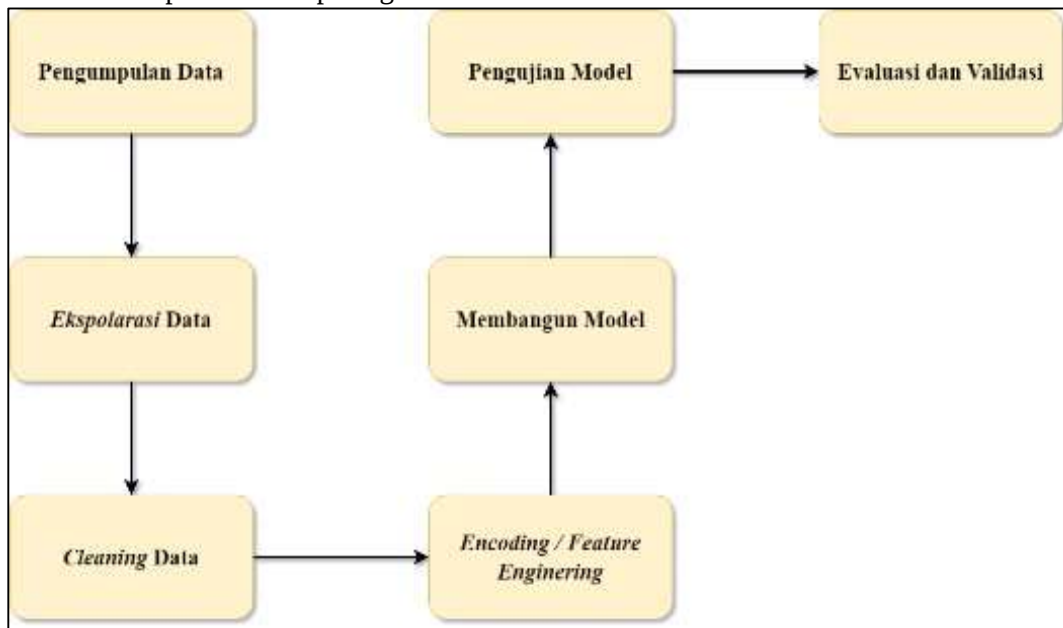
CART (Classification and Regression Tree)

Classification and regression trees adalah strategi klasifikasi berbasis data yang menggunakan data historis untuk membuat *Decision Trees* [14]. CART (*Classification and Regression Trees*) adalah teknik pembelajaran mesin yang digunakan untuk membangun pohon keputusan baik untuk klasifikasi maupun regresi. Tahapan metode CART meliputi pemilihan atribut dan kriteria pemisahan, pembangunan pohon

keputusan, pemangkasan pohon untuk menghindari overfitting, dan evaluasi serta validasi hasil. Proses ini memastikan bahwa pohon keputusan yang dihasilkan efektif dalam klasifikasi atau regresi dengan performa yang baik pada data yang tidak terlihat sebelumnya [15]

Dalam penelitian ini metode yang digunakan adalah menerapkan algoritma Decision Tree untuk memprediksi harga Minyak Sawit Mentah dengan menggunakan algoritma CART dimana data yang diambil merupakan dataset data Minyak Sawit berjangka yang diambil dari website investing.com. Dengan alamaturl : (<https://id.investing.com/commodities/palmoil-historical-data>).

Ada beberapa tahapan yang dibutuhkan dalam penelitian ini yaitu: Pengolahan data, Eksplorasi data, *Cleaning data*, *Feature engineering*, membangun model, kemudian Pengujian Model dan terakhir evaluasi dan validasi hasil seperti terlihat pada gambar di bawah ini.



Gambar 3. 1 Perancangan Penelitian

- Pengolahan Data**
Data minyak sawit yang diolah yaitu di ambil dari website *Investing.com* dengan rentan waktu 10 tahun dari tahun 2014 -2024 data tersebut dengan format berbentuk CSV
- Eksplorasi Data**
Eksplorasi data untuk memahami pola, tren, dan hubungan antar variabel. Pada tahap ini visualisasi data seperti grafik *time-series* atau *scatter plot* sering digunakan untuk melihat bagaimana variabel-variabel berubah seiring waktu dan apakah terdapat *outlier* atau anomali dalam data.
- Cleaning Data**
Cleaning data atau pembersihan data untuk mengatasi masalah seperti data *missing values* dan *outlier*. Data yang kosong dapat diisi dengan nilai rata-rata (*imputasi*) atau dihapus jika tidak relevan.
- Encoding (Feature Engineering)**
Feature engineering yaitu menciptakan dan memilih variabel yang paling relevan untuk meningkatkan performa model. Pada proses ini dilakukan untuk mengubah data kategorikal menjadi bentuk *numerik* yang dapat di proses oleh algoritma pembelajaran atau model *statistic*.
- Membangun Model**
Membangun model dengan algoritma *Decision Tree* diimplementasikan dengan menggunakan library *scikit-learn*. *Scikit-Learn* atau *Sklearn* adalah library berbasis *Python* untuk membangun model pembelajaran mesin. Dan menyediakan banyak algoritma pembelajaran untuk regresi, pengelompokan, dan klasifikasi. *Sklearn* kompatibel dengan *NumPy* dan *SciPy*.
- Pengujian Model**
Pengujian model menggunakan data uji untuk melihat akurasi dan kesalahan prediksi. Beberapa metrik evaluasi yang umum digunakan adalah *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *R-squared* (R^2).
- Evaluasi dan Validasi**

Evaluasi dan validasi untuk memastikan model bekerja dengan baik pada data baru dan tidak hanya pada data latih. Jika hasilnya belum optimal, tuning parameter atau metode *ensemble* dapat dicoba untuk meningkatkan akurasi.

3. HASIL

Membangun Model

Membangun model mesin learning merupakan sebuah langkah untuk membuat prediksi, dataset yang saya gunakan berjangka 10 tahun sebagai berikut.

Tabel 4. 1 Membangun Model

No	Dataset yang Digunakan	Dimensi
1	Data Historis Minyak Sawit Berjangka (FCPOc1)	2541 Baris Data & 11 Kolom

a. *Splitting Data*

Data *Splitting* atau pemisahan data adalah metode membagi data menjadi dua bagian atau lebih yang membentuk sub himpunan data. Umumnya *splitting* data yaitu membagi data menjadi 2 bagian, data historis yaitu menguji coba data historis 10 tahun kebelakang dan data uji untuk menguji data latih selama setahun kebelakang.

Tabel 4. 2 *Splitting Data*

Dataset	Jumlah
Data Historis	1779 Baris Data & 7 Kolom
Data Uji	762 Baris Data & 7 Kolom

a. *Training Model*

Training Model atau Pelatihan model adalah proses memasukkan data rekayasa ke alghoritma pembelajaran mesin *learning* untuk menghasilkan model dengan parameter optimal yang dapat dilatih dan di pelajari . Pada penelitian saya ini menggunakan model *Decision Tree* untuk menganalisis dan mengambil Keputusan berdasarkan data yang ada, model *decision tree* adalah metode pengambilan Keputusan yang menggambarkan alur logika dalam bentuk pohon. Seperti gambar berikut ini gambar pohon keputusan :



Gambar 4. 1 Training Model Decision Tree

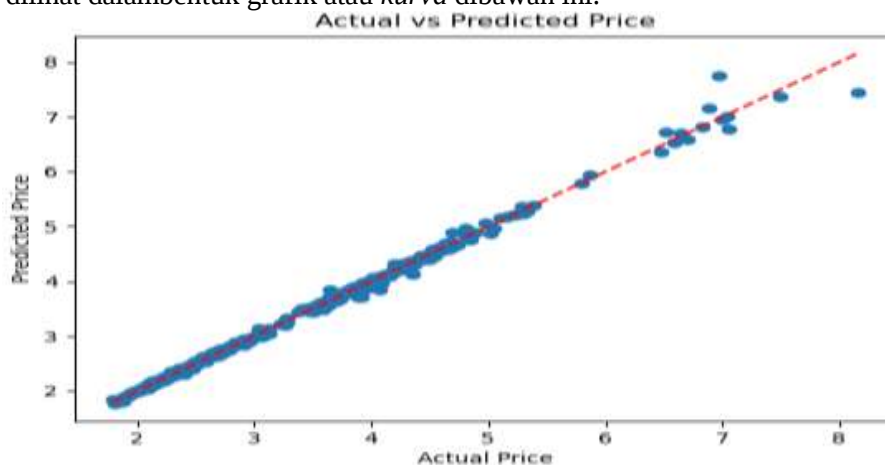
Pengujian Model

Pengunjian model ini dilakukan untuk memastikan bahwa model dapat memberikan prediksi yang akurat dan dapat diandalkan dalam mengklasifikasi dengan data test. Dalam pengujian ini, beberapa metrik evaluasi seperti *Mean Squared Errorr* (MSE), *Mean Absolute Errorr* (MAE), *R-Square* digunakan untuk mengukur sejumlah model dapat mengklasifikasikan data dengan benar.

Tabel 4. 3 Pengujian Model

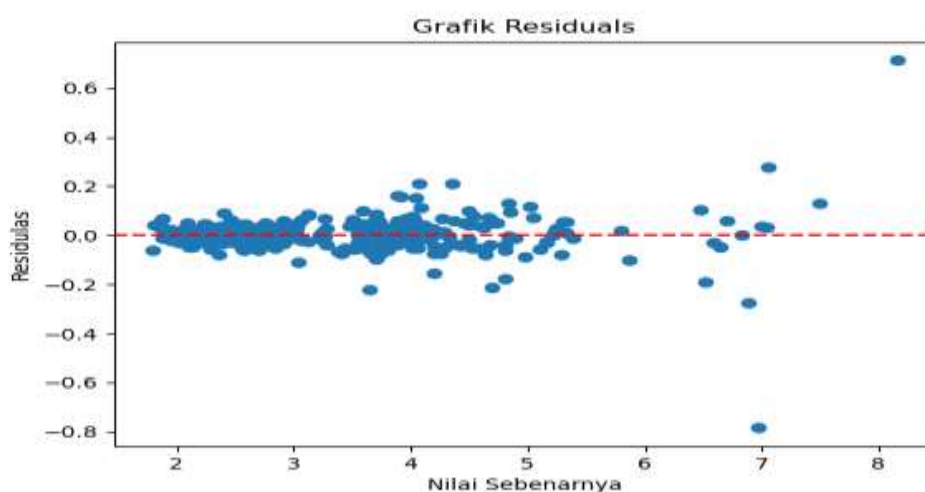
Model	Hasil
<i>Mean Squared Error (MSE)</i>	0.004
<i>Mean Absolute Error (MAE)</i>	0.032
<i>R-Square</i>	0.996

Berdasarkan tabel diatas pengujian model dari pembuatan model tersebut mendapatkan nilai akurasi *R-Square* 0.996, nilai MSE sebesar 0.032, dan nilai MAE sebesar 0.004. Hasil pengujian model mesin learning dapat dilihat dalam bentuk grafik atau *kurva* dibawah ini.



Gambar 4. 2 Hasil Pengujian Model *Machine Learning*

Berdasarkan grafik perbandingan antara harga aktual dan harga prediksi, terlihat bahwa sebagian besar titik data berada sangat dekat dengan garis merah. Hal ini menunjukkan bahwa model memiliki tingkat akurasi yang tinggi, karena prediksi harga yang dihasilkan hampir sama dengan nilai sebenarnya. Namun, terdapat beberapa penyimpangan kecil pada nilai harga yang lebih tinggi, di mana titik-titik sedikit menjauhi garis. Meskipun demikian, secara keseluruhan model telah menunjukkan performa yang sangat baik dan mampu memprediksi dengan akurasi yang tinggi pada sebagian besar data. Sedangkan grafik *Residual* dari dataset tersebut sebagai berikut :



Gambar 4. 3 Grafik Residual Model *Machine Learning*

Grafik residual menunjukkan bahwa sebagian besar titik residual berada cukup dekat dengan garis nol, khususnya pada rentang nilai sebenarnya antara 2 hingga 5. Hal ini mengindikasikan bahwa model memiliki performa prediksi yang baik dan kesalahan prediksi relatif kecil pada rentang nilai tersebut. Grafik

residual tersebut cukup baik karena semakin banyak persebaran data yang mendekati nilai 0 maka data tersebut semakin baik dengan mengikuti garis *linieritas*.

1.1 Evaluasi dan Validasi Hasil

Pada tahap ini dilakukan prosen pengujian model mesin *learning* yang telah dibuat dengan menggunakan pola data atau dataset uji yang telah disiapkan mulai dari bulan januari - Mei 2024 sebelumnya, adapun dataset yang digunakan dapat dilihat pada tabel berikut.

Tabel 4. 4 Evaluasi Data

No	Dataset yang Digunakan	Dimensi
1	Data Historis Minyak Sawit Berjangka (FCPOc1)	69 Baris Data & 7 Kolom

Pada tabel diatas merupakan dataset yang akan digunakan untuk pengujian evaluasi dan validasi hasil, Dimana data yang diperoleh dari data uji sebanyak 208 data dengan 7 kolom.

```
rmse= mean_squared_error(y_uji, y_pred2, squared=False)
print("Root Mean Squared Error:", rmse)

mse = mean_squared_error(y_uji, y_pred2)
print('Mean Squared Error:', mse)

mae = mean_absolute_error(y_uji, y_pred2)
print('Mean Absolute Error:', mae)

r2 = r2_score(y_uji, y_pred2)
print('R-Square:', r2)

r2=r2 *100
print('Accuracy:', r2, '%')

Root Mean Squared Error: 0.06209598831418828
Mean Squared Error: 0.0038559117647058716
Mean Absolute Error: 0.04858823529411761
R-Square: 0.9310121400582015
Accuracy: 93.10121400582015 %
```

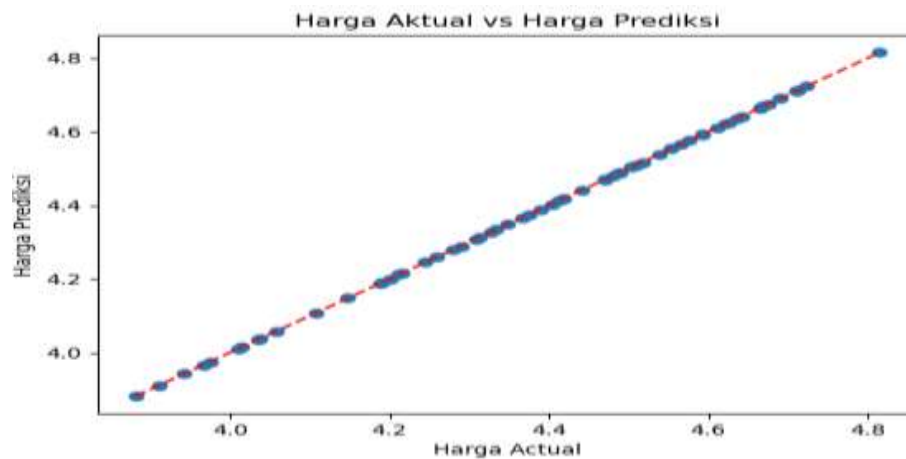
Gambar 4. 4 Validasi Hasil

Gambar diatas merupakan perhitungan dari pengujian model dan validasi hasil menggunakan data uji yang digunakan untuk menghitung parameter model uji dalam penelitian ini dan didapat kan hasil sebagai berikut :

Tabel 4. 5 Validasi Hasil

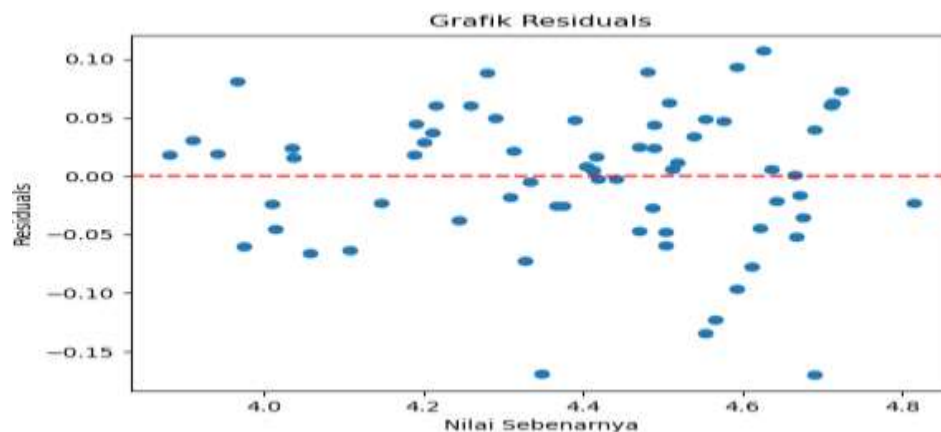
Model	Hasil
Root Mean Squared Errorr (RMSE)	0.062
Mean Squared Errorr (MSE)	0.003
Mean Absolute Errorr (MAE)	0.046
R-Square	0.937

Hasil dari model memiliki tingkat akurasi yang dengan nilai *R-Square* 0.937, nilai RMSE sebesar 0.062, nilai MSE sebesar 0.003, dan nilai MAE sebesar 0.046, Sedangkan grafik harga aktual dengan harga prediksi sebagai berikut :



Gambar 4. 5 Hasil Aktual Dengan Prediksi

Pada gambar diatas mmenunjukkan bahwa semakin dekat titik data dengan garis linier, maka semakin baik performa model dalam memprediksi harga CPO. Titik-titik biru yang merepresentasikan harga aktual vs harga prediksi tersebar sangat rapat di sekitar garis merah putus-putus, yang menunjukkan bahwa model prediksi memiliki akurasi yang sangat tinggi. Sedangkan grafik residual dari data uji tersebut sebagai berikut :



Gambar 4. 6 Grafik Residual Data Uji

Berdasarkan grafik residual yang dihasilkan, terlihat bahwa sebagian besar titik residual tersebar secara acak dan berada sangat dekat dengan garis nol. Jikka model bagus, titik-titik residual akan menyebar secara acak di sekitar nol. Jika membentuk pola tertentu, berarti model belum tepat. Hal ini menunjukkan bahwa selisih antara nilai prediksi dan nilai aktual relatif kecil, yang mengindikasikan bahwa model memiliki tingkat akurasi yang tinggi.

4. PEMBAHASAN

Persiapan Data

Data historis Minyak Sawit Berjangka (FCPOc1). dari *website* resmi *Investing.com* yang di unduh dai url: (<https://id.investing.com/commodities/palmoil-historical-data>). Seperti pada tabel di bawah ini :

Tabel 4. 1 Persiapan Dataset

No	Nama	Record	Feature	Class
1	Data historis Minyak Sawit Berjangka (FCPOc1)	2541	6	1

Dataset ini merupakan dataset *public* dengan *record* yang cukup banyak dengan berbagai tipe. Karena dataset ini merupakan dataset *realtime* yang setiap harinya akan terus bertambah dengan data baru.

Implementasi Metode

<i>Input/Atrtribute</i>	Keterangan
Tanggal	Tanggal data Minyak Sawit Berjangka (FCPOc1) diambil dan tercatat.
Terakhir	Harga penutupan indeks pada hari perdagangan tersebut.
Pembukaan	Harga pembukaan indeks pada hari perdagangan tersebut.
Tertinggi	Harga tertinggi indeks pada hari perdagangan tersebut.
Terendah	Harga terendah indeks pada hari perdagangan tersebut.
<i>Vol.</i>	Jumlah yang di perdagangan pada hari perdagangan tersebut.
Perubahan%	Persentase perubahan harga indeks dibandingkan dengan hari perdaganan sebelumnya.

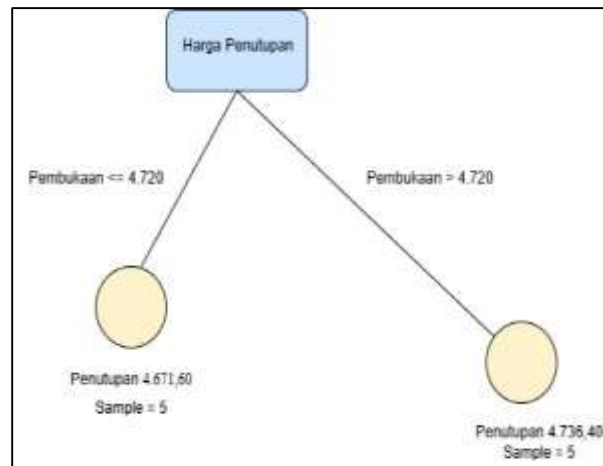
Pembuatan model pohon keputusan dilakukan pada dataset yang terdiri dari 7 atribut yang merupakan atribut dari dataset histori minyak sawit berjangka. Algoritma yang digunakan adalah CART (*Classification and Regression Tree*). Metode ini bekerja dengan membuat pohon keputusan untuk memprediksi variabel target harga CPO, perhitungannya dilakukan sebagai berikut dengan menggunakan sampel 10 data dari dataset uji yang saya gunakan

a. Pengambilan 10 sampel data di tahun 2025

Tabel 4. 2 Implementasi Metode

No	Tanggal	Harga Pembukaan	Harga Terakhir
1	02/01/2025	4.862,00	4.690,00
2	03/01/2025	4.720,00	4.723,00
3	06/01/2025	4.740,00	4.690,00
4	07/01/2024	4.659,00	4.713,00
5	08/01/2024	4.700,00	4.711,00
6	09/01/2024	4.695,00	4.635,00
7	10/01/2024	4.697.00	4.710,00
8	13/01/2024	4.800,00	4.815,00
9	14/01/2024	4.789,00	4.668,00
10	15/01/2024	4.637,00	4.675,00

Kita akan menggunakan harga pembukaan sebagai atribut prediktor (X) dan harga penutupan sebagai target (Y).



Pengolahan Data

Dataset di tampilkan untuk mengetahui tipe data yang ada di dalam dataset tersebut.

Tabel 4. 3 Keterangan Dataset Awal

Nama Atribut	Tipe data
Tanggal	Object
Terakhir	Object
Pembukaan	Object
Tertinggi	Object
Terendah	Object
Vol.	Object
Perubahan%	Object

Dari dataset awal semua data memiliki tipe data *Object*. Setelah data diketahui maka data tersebut diolah dengan dikonversi ke dalam tipe data *numerik* agar sesuai dengan kebutuhan dari model pohon keputusan.

Tabel 4. 4 Konversi Tipe data

Nama Atribut	Tipe data	Nilai
Tanggal	<i>Date time</i>	2014/08/18 - 2024/12/31
Terakhir	<i>Float</i>	1.759 – 8.163
Pembukaan	<i>Float</i>	1.751 – 8.200
Tertinggi	<i>Float</i>	1.759 – 8.757
Terendah	<i>Float</i>	1.750 – 7.780
Vol.	<i>Float</i>	0.000 – 632.000
Perubahan %	<i>Float</i>	0.000 - 1091.000

Eksplorasi Data

Pada tahap eksplorasi data ini dilakukan proses pembacaan data awal dari dataset dengan tujuan untuk mengetahui adanya data yang hilang serta mencari anomali-anomali pada dataset tersebut. Pada dataset awal di atas kita dapat mengetahui terdapat 2541 baris dan 7 kolom atribut, dengan *Missing Value* terdapat pada kolom *Vol*.

Analisa Data

Analisis data merupakan serangkaian proses pengolahan data dan penelitian yang bertujuan untuk menemukan informasi penting guna membantu pengambilan keputusan dan pemecahan masalah. Proses ini dimulai dengan mengelompokkan data berdasarkan karakteristik, membersihkan data untuk mengatasi ketidak konsistenan dan nilai yang hilang, mentransformasikan data agar dapat dipersiapkan dengan baik untuk dianalisis, dan menyediakan pola pencarian yang relevan. Proses ini melibatkan beberapa fase, dimulai dengan membangun model yang memperoleh wawasan penting dari data.

Dalam analisa data pada dataset dilakukan pemisahan terkait dengan *class* atau label yang berfungsi sebagai target atau *output* dengan *atribut* atau *feature* yang berfungsi sebagai *input*.

Tabel 4. 6 Analisa Data

Kolom	Tipe Data
Tanggal	<i>Object</i>
Terakhir	<i>Object</i>
Pembukaan	<i>Object</i>
Tertinggi	<i>Object</i>
<i>Vol</i> .	<i>Object</i>
Perubahan%	<i>Object</i>

Pada Dataset ini, setiap kolom masih dalam tipe data objek yang Dimana nantinya harus di encode agar bisa di proses.

a. *Class* atau Label

Class atau label pada dataset ini adalah kolom Terakhir yang merupakan harga penutupan atau harga terakhir dari penjualan pada hari tersebut, yang akan berfungsi sebagai *output* dari model mesin *learning* yang di buat, dari analisa *price* atau terakhir dapat dilihat pada tabel berikut.

Tabel 4. 7 Analisa Data Terakhir

Keterangan	Nilai
Nilai Minimal	1.759
Nilai Maksimal	8.163
Jumlah Baris	3541 Baris Data
Keunikan Setiap Data	1534Baris Data

Atribut atau Feature

1. Tanggal

Tanggal adalah kolom yang merupakan tanggal dari penjualan minyak kelapa sawit mentah pada hari tersebut. Pada kolom tanggal ini terdapat 2541 baris data yang dimana setiap isi berbeda-beda sebanyak 2541 data. Tanggal yang di ambil dari data ini dimulai dari 18/08/2014 sampai 31/12/2024.

Tabel 4. 8 Analisa data Tanggal

Keterangan	Nilai
Jumlah Baris	2541 Baris Data
Keunikan setiap data	2541 Baris Data
Tanggal Awal	18/08/2014
Tanggal Akhir	31/12/2024

2. Pembukaan Kolom pembukaan merupakan harga pembukaan/awal dari penjualan pada hari tersebut. Pada kolom pembukaan terdapat 2541 baris data dengan jumlah keunikan 1534 baris data, data yang paling banyak muncul yaitu 2.200,00 dengan frekuensi kemunculan sebanyak 11 kali.

Tabel 4. 9 Analisa data Pembukaan

Keterangan	Nilai
Jumlah Baris	2541 Baris Data
Keunikan setiap data	1534 Baris Data
Data yang sering muncul	2.200,00
Frekuensi kemunculan	11

3. Tertinggi

Kolom Tertinggi merupakan harga tertinggi dari penjualan pada hari tersebut. Pada kolom Tertinggi terdapat 2541 baris data dengan jumlah keunikan 1461 baris data, data yang paling banyak muncul yaitu 2.150,00 dengan frekuensi kemunculan sebanyak 13 kali.

Tabel 4. 10 Analisa data Tertinggi

Keterangan	Nilai
Jumlah Baris	2541 Baris Data
Keunikan setiap data	1461 Baris Data
Data yang sering muncul	2.150,00
Frekuensi kemunculan	13

4. Terendah

Kolom Terendah merupakan harga terendah dari penjualan pada hari tersebut. Pada kolom Terendah terdapat 2541 baris data dengan jumlah keunikan 1454 baris data, data yang paling banyak muncul yaitu 2.850,00 dengan frekuensi kemunculan sebanyak 4 kali.

Tabel 4. 11 Analisa data Terendah

Keterangan	Nilai
Jumlah Baris	2541 Baris Data
Keunikan setiap data	1454 Baris Data
Data yang sering muncul	3.850,00
Frequensi kemunculan	11

5. Vol.

Kolom *Vol.* merupakan volume dari penjualan pada hari tersebut. Pada kolom *Vol.* terdapat 2392 baris data dengan jumlah keunikan 276 baris data, data yang paling banyak muncul yaitu 0,01k dengan frekuensi kemunculan sebanyak 87 kali.

Tabel 4. 12 Analisa Data Vol.

Keterangan	Nilai
Jumlah Baris	2392 Baris Data
Keunikan setiap data	276 Baris Data
Data yang sering muncul	0.01K
Frequensi kemunculan	87

6. Perubahan%

Kolom perubahan merupakan harga perubahan dari penjualan kedalam bentuk persentase pada hari tersebut. Pada kolom *Perubahan%* terdapat 2541 baris data dengan jumlah keunikan 713 baris data, data yang paling banyak muncul yaitu 0,00% dengan frekuensi kemunculan sebanyak 80 kali.

Tabel 4. 13 Analisa data Perubahan%

Keterangan	Nilai
Jumlah Baris	2541
Keunikan setiap data	713
Data yang sering muncul	0,00%
Frequensi kemunculan	80

Statistika Data

Pada tahap ini dilakukan proses pemanggilan informasi terkait dengan nilai-nilai statistik pada dataset, khususnya pada kolom-kolom yang bertipe *numerik* pada dataset ini, seperti tabel berikut ini.

Tabel 4. 14 Statistika Data

	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan%
Count	2541.000	2541.000	2541.000	2541.000	2541.000	2541.000
Mean	3.207	3.204	3.234	3.178	57.609	131.288
Min	1.759	1.751	1.759	1.750	0.000	0,000
25%	2.278	2.274	2.290	2.263	9000	47.000

50%	2.298	2.799	2.811	2.779	32.000	100.000
75%	3.929	3.930	3.962	3.895	86.000	175.000
Max	8.163	8.200	8.757	7.780	632.000	1091.000
Std	1.137	1.132	1.155	1.114	67.409	125.888

Tabel 4. 15 Keterangan Statistik Data

Nama	Keterangan
Count	digunakan untuk menghitung jumlah data elemen dalam suatu array atau struktur data.
Mean	Nilai rata-rata dari sejumlah nilai atau data.
Minimal	Nilai terkecil dari data yang ada
25%	Nilai kuartil pertama.
50%	Nilai kuartil kedua.
75%	Nilai kuartil ketiga.
Maximal	Nilai tertinggi dari data yang ada.
Standar Deviasi	Nilai statistik yang di gunakan untuk mengukur sejumlah mana data tersebar dari rata-rata dalam satu himpunan data.

5. KESIMPULAN

Kesimpulan Berdasarkan implementasi dan pengujian data yang telah dilakukan pada penelitian ini dapat disimpulkan:

- Prediksi harga minyak kelapa sawit mentah dengan metode decision tree menggunakan algoritma Classification And Regression atau CART menghasilkan akurasi yang sangat baik dengan accuracy 93.101%, R-Square 0.937, RMSE 0.062, MAE 0.046.
- Hasil dari penelitian ini terbukti bahwa algoritma Classification And Regression tree cocok untuk diterapkan memprediksi harga minyak kelapa sawit mentah (Crude Palm Oil).

6. UCAPAN TERIMA KASIH

Penulis menyadari bahwa penyusunan skripsi ini tidak lepas dari bantuan dan dukungan dari berbagai pihak. Oleh karena itu penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

- Bapak samsoni S.kom., M.kom selaku dosen pembimbing yang telah memberikan arahan menuntun saya hingga proses penelitian ini berjalan dengan baik.
- Kepada Kedua orang tua dan teman-teman yang selalu memberikan semangat dan doa serta dukungan kepada penulis agar penulis dapat menyelesaikan penelitian ini.
- Kepada semua pihak yang tidak disebutkan namanya yang telah membantu dalam penyusunan penelitian ini saya ingin mengucapkan banyak terima kasih.

DAFTAR PUSTAKA

- [1] R. Mustafa, "Pengaruh Harga CPO (Crude Palm Oil) Di Global Market Terhadap Harga Minyak Goreng Di Pasar Domestik," *SIBATIK Journal*, pp. 1565-1573, 2022.
- [2] D. Levia and M. , "Proses Produksi CPO Untuk Mengidentifikasi Faktor-Faktor Yang Mempengaruhi Kualitas Mutu CPO," *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, pp. 82-87, 2023.
- [3] L. . S. Oktaviana, A. Mayes and Darmayuda, "Pengaruh Produk Domestik Bruto China, Harga CPO International, Dan Harga Minyak Kedelai Internasional Terhadap Ekspor Crude Palm Oil (CPO) Di Provinsi Riau," *Jurnal Analisis dan Perkembangan Ekonomi (JAPE)*, pp. 21-26, 2025.
- [4] S. L. Abdullah, A. F. Akbariyah and R. Wikansari, "POTENSI EKSPOR CRUDE PALM OIL (CPO) DI INDONESIA," *Journal of Science and Social Research* , pp. 61-67, 2024.
- [5] D. Nurhayati, H. Noviarita and M. . I. Fasa, "Kerjasama Indonesia Dan Malaysia Dalam Hidrilisasi Kelapa Sawit," *Management Studies and Entrepreneurship Journal* , pp. 9136-9141 , 2023.
- [6] R. Porwati , "Analisis Pengaruh Harga CPO, Tingkat Inflasi Dan Nilai Tukar Rupiah Terhadap Harga Saham Perusahaan Sub Sektor Perkebunan Kelapa Sawit Tahun 2017-2021," *Maeswara: Jurnal Riset Ilmu Manajemen dan Kewirausahaan*, pp. 169-179, 2024.
- [7] M. Ofuoku and T. Ngriatedema, "Predicting the Price of Crude Palm oil:A Deep Learning Approach," *International Journal of Strategic Decision Sciences*, pp. 1-2, 2022.
- [8] S. Wijaya and F. , "A Comperative analysis of Llinear Regression And Random Forest Methods For Predicting Palm Oil Price Sales," *Jurnal Ilmiah Teknologi dan Rekayasa Volume*, pp. 265-274, 2023.
- [9] K. Kasturi, . N. . R. Krishnan, U.R.Hashim, M. M. I. S and M. Jarrah, "A Comparative Study on Univariate Time Series based Crude Palm Oil Price Prediction Model using Machine Learning Algorithms," *International Journal of Advanced Trends in Computer Science and Engineering*, pp. 5802-5806, 2020.
- [10] N. Nafi'iyah and N. N. Aulia, "Prediksi Harga Minyak Sayuran Data Kaggle dengan Regresi Linear Berganda dan Backpropagation," *Jurnal SISFOTENIKA*, pp. 136-145, 2022.
- [11] A. T. Nurani, T. Setiawa and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *Jurnal Sains dan Edukasi Sains*, pp. 34-43, 2023.
- [12] R. A. Saputra and A. Pratama, "implementasi decision tree untuk prediksi harga rumah di daerah tebet," *Journal of Information System Management (JOISM)*, pp. 164-170, 2025.
- [13] R. W. Romadhonia, . A. Sofro, D. Ariyanto, D. A. Maulana and J. . B. Prihanto, "Application of Decision Trees in Athlete Selection: A Cart Algorithm Approach," *Journal Of Data Science*, pp. 1-9, 2023.
- [14] H. . A. Salman , . A. Kalakech and A. Steiti, "Random Forest Algorithm Overview," *Babylonian Journal of Machine Learning*, pp. 69-79, 2024.
- [15] R. Agustin , S. Defit and Sumijan , "Komparasi Algoritma CART dan C 4.5 pada Citra Tandan Buah Sawit untuk Mengetahui Tingkat Kematangan dalam Penentuan Harga," *Jurnal KomtekInfo*, pp. 263-273, 2024.