

Data-Driven Model for Predicting Essential Commodity Demand: A Comparative Study of Weekend–Weekday Patterns in Chili Commodities Using Decision Tree and Random Forest Regression

Mayang Adiani Erissafilla ^{1*}, Marcelinus Yosep Teguh Sulistyono ^{2*}

* Sistem Informasi, Universitas Dian Nuswantoro, Semarang

112202206894@mhs.dinus.ac.id ¹, teguh.sulistyono@dsn.dinus.ac.id ²

Article Info

Article history:

Received 2026-01-28

Revised 2026-03-02

Accepted 2026-04-08

Keyword:

Chili Price Forecasting, Decision Tree Regression, Random Forest Regression, Weekday-weekend Pattern, Machine Learning.

ABSTRACT

Price fluctuations of chili commodities as essential food products directly affect public purchasing power and regional economic stability. Daily price movements are often influenced by temporal demand dynamics between weekdays and weekends; however, predictive approaches based on such patterns remain relatively limited in previous studies. This study aims to predict the prices of Red Bird's Eye Chili and Curly Red Chili in Central Java Province using a data-driven approach by comparing the performance of Decision Tree Regression and Random Forest Regression based on weekday–weekend classification patterns. The dataset consists of daily price data from January 2023 to December 2025 obtained from the Provincial Industry and Trade Office of Central Java. The preprocessing stage includes missing value imputation, outlier detection using the Interquartile Range (IQR) method, Min–Max normalization, and categorical variable encoding. Pearson correlation analysis indicates a strong positive relationship between the two commodities with a coefficient of 0.64, suggesting interconnected price movements although not perfectly correlated. Model evaluation was conducted using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). The results show that Random Forest Regression produces relatively more stable predictive performance compared to Decision Tree Regression. For Red Bird's Eye Chili, the model achieved an MAE of 0.2481, RMSE of 0.2741, and R^2 of 0.0038. For Curly Red Chili, the results obtained were an MAE of 0.1751, RMSE of 0.2204, and R^2 of -0.0135 . Although the R^2 values indicate limited explanatory power in capturing price variability, the ensemble learning approach provides better prediction consistency in modeling volatile agricultural commodities. These findings contribute empirically to the development of machine learning-based price forecasting models at the provincial level by incorporating weekly temporal patterns.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Stabilitas harga kebutuhan pokok merupakan salah satu indikator penting dalam menjaga kesejahteraan masyarakat dan kestabilan ekonomi daerah [1]. Komoditas hortikultura, khususnya cabai, memiliki peran strategis dalam struktur konsumsi rumah tangga di Indonesia karena digunakan secara luas dalam berbagai produk pangan serta menjadi komponen penyumbang inflasi bahan makanan. Tingginya sensitivitas harga cabai terhadap gangguan pasokan menjadikannya salah

satu komoditas dengan volatilitas paling tinggi di pasar domestik. Fluktuasi harga cabai dipengaruhi oleh berbagai faktor, antara lain musim panen, kondisi iklim, gangguan distribusi, biaya logistik, serta perubahan pola konsumsi masyarakat [2]. Karakteristik produksi yang bersifat musiman dan ketergantungan pada kondisi cuaca menyebabkan ketidakseimbangan antara permintaan dan penawaran, sehingga harga sering mengalami lonjakan tajam dalam periode tertentu. Kondisi ini tidak hanya berdampak pada daya beli konsumen dan pendapatan petani, tetapi juga

berkontribusi terhadap tekanan inflasi daerah serta instabilitas pasar pangan [3].

Di tingkat daerah, Dinas Perindustrian dan Perdagangan (Disperindag) berperan dalam memantau harga dan menjaga stabilitas pasokan melalui pengawasan distribusi serta pelaksanaan operasi pasar [4]. Namun, kebijakan stabilisasi harga masih cenderung bersifat reaktif karena keterbatasan sistem prediksi yang mampu mengantisipasi dinamika harga secara cepat dan akurat. Oleh karena itu, diperlukan pendekatan prediktif berbasis data yang mampu memodelkan pola historis sekaligus menangkap karakteristik fluktuasi jangka pendek dalam siklus waktu tertentu.

Berbagai penelitian sebelumnya telah menggunakan metode statistik tradisional seperti *ARIMA* dan *Fuzzy Time Series* dalam memprediksi harga komoditas pertanian [5]. Meskipun metode tersebut efektif dalam menangkap pola linier dan tren musiman, pendekatan tersebut memiliki keterbatasan dalam merepresentasikan hubungan non-linier serta interaksi kompleks antarvariabel pada data harga harian yang bersifat volatil [6]. Selain itu, berbagai penelitian sebelumnya cenderung berfokus pada analisis berbasis tren bulanan atau musiman. Di sisi lain, studi mengenai *day-of-the-week effect* menunjukkan bahwa harga atau volatilitas dapat berbeda secara sistematis berdasarkan hari dalam satu minggu [7]. Namun demikian, penerapan pendekatan ini secara spesifik pada komoditas hortikultura seperti cabai masih relatif terbatas. Padahal, dinamika permintaan pasar sering kali dipengaruhi oleh siklus mingguan yang berpotensi menciptakan variasi harga sistematis.

Algoritma berbasis pohon keputusan seperti *Decision Tree Regression* serta metode *ensemble* seperti *Random Forest Regression* telah digunakan dalam berbagai studi prediksi harga komoditas [8]. Sementara itu, pada konteks harga cabai, pendekatan lain seperti model probabilistik Rantai Markov juga telah diterapkan dalam penelitian sebelumnya [9]. Penggunaan algoritma *Decision Tree*, *Random Forest*, dan *LSTM* telah diterapkan dalam berbagai studi prediksi harga komoditas di Indonesia [10]. Namun demikian, berdasarkan penelusuran literatur yang dilakukan, integrasi variabel temporal berbasis klasifikasi *weekday-weekend* dalam kerangka machine learning untuk pemodelan harga cabai pada tingkat provinsi masih relatif terbatas.

Penelitian ini bertujuan membangun dan membandingkan kinerja model *Decision Tree Regression* dan *Random Forest Regression* dalam memprediksi harga cabai berdasarkan pola waktu mingguan. Hasil penelitian diharapkan memberikan kontribusi empiris terhadap pengembangan model prediksi harga kebutuhan pokok serta menjadi dasar pertimbangan bagi Disperindag dalam merumuskan strategi harga yang lebih proaktif dan berbasis data.

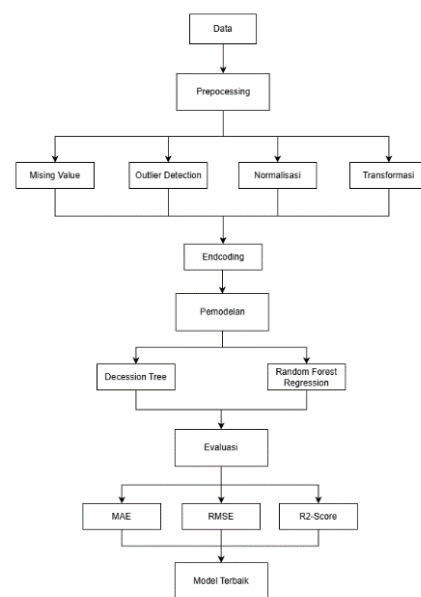
Berdasarkan celah penelitian tersebut, kebaruan (*novelty*) penelitian ini terletak pada: (1) integrasi variabel temporal *weekday-weekend* dalam pemodelan harga cabai berbasis *machine learning*, (2) analisis komparatif performa *Decision Tree Regression* dan *Random Forest Regression* pada data harga harian tingkat provinsi yang bersifat volatil, serta (3)

evaluasi implikasi hasil model terhadap strategi stabilisasi harga daerah. Dengan demikian, penelitian ini tidak hanya berkontribusi pada pengembangan metodologi prediksi harga komoditas berbasis data, tetapi juga memberikan dasar empiris bagi perumusan kebijakan harga yang lebih proaktif dan adaptif.

Penelitian ini bertujuan untuk membangun dan membandingkan kinerja model *Decision Tree Regression* dan *Random Forest Regression* dalam memprediksi harga cabai berdasarkan pola waktu mingguan, serta menganalisis pengaruh perbedaan *weekday* dan *weekend* terhadap dinamika harga. Hasil penelitian diharapkan dapat memperkaya literatur terkait pemodelan harga komoditas hortikultura serta mendukung pengambilan keputusan berbasis data di tingkat daerah.

II. METODE

Penelitian ini menggunakan pendekatan kuantitatif berbasis data-driven untuk membandingkan kinerja algoritma *Decision Tree Regression* dan *Random Forest Regression* dalam memprediksi harga cabai berdasarkan pola weekday-weekend. Tahapan penelitian meliputi : (1) pengumpulan data, (2) *preprocessing*, (3) *encoding*, (4) pemodelan, (5) evaluasi, dan (6) pemilihan model terbaik, di tunjukkan pada gambar 1.



Gambar 1. Tahapan Penelitian

A. Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Dinas Perindustrian dan Perdagangan Provinsi Jawa Tengah melalui Survei Pemantauan Harga Barang Kebutuhan Pokok (SP2KP). Data yang dianalisis berupa harga harian komoditas cabai rawit merah dan cabai merah keriting pada periode Januari 2023 hingga Desember 2025. Variabel yang digunakan meliputi tanggal, bulan, tahun, komoditas, dan harga. Seluruh data

kemudian diintegrasikan ke dalam satu tabel untuk memudahkan proses pengolahan data dan pemodelan pada tahap selanjutnya, Sebagaimana ditunjukkan pada Tabel I.

TABEL I
DATASET HARGA KOMODITAS CABAI TAHUN 2023-2025

Bulan	Tahun	Komoditas	Tanggal	Harga (Rp/Kg)
1	2023	Cabai Rawit Merah	02/01/23	62.758
1	2023	Cabai Merah Keriting	02/01/23	40.419
1	2023	Cabai Rawit Merah	03/01/23	62.447
.....
.....
12	2025	Cabai Merah Keriting	31/12/25	37.905
12	2025	Cabai Rawit Merah	31/12/25	47.767

B. Pre Processing

1. Missing value

Missing Value merupakan kondisi ketika suatu atribut atau variabel dalam dataset tidak memiliki nilai (kosong atau tidak tercatat). Keberadaan *missing value* dapat memengaruhi kualitas data dan menurunkan kinerja model, sehingga perlu dilakukan penanganan sebelum proses pemodelan.

Pada penelitian ini, penanganan dilakukan menggunakan metode mean imputation, yaitu mengganti nilai yang hilang dengan nilai rata-rata atribut tersebut.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

Persamaan 1 menunjukkan perhitungan nilai rata-rata \bar{x} dari suatu data. Pada persamaan tersebut, x_i merupakan nilai data ke- i , dan n menyatakan jumlah total data yang tidak mengandung missing value [11].

2. Outlier Detection

Outlier detection merupakan proses untuk mengidentifikasi data yang memiliki nilai ekstrem dan menyimpang jauh dari pola data lainnya. *Outlier* dapat muncul akibat kesalahan pencatatan atau kondisi yang tidak umum dan berpotensi memengaruhi hasil analisis serta akurasi model.

Metode yang digunakan adalah *Interquartile Range (IQR)* :

$$IQR = Q_3 - Q_1 \quad (2)$$

$$\text{Batas Bawah} = Q_1 - 1.5 \times IQR$$

$$\text{Batas Atas} = Q_3 + 1.5 \times IQR$$

Persamaan 2 menunjukkan Q_1 adalah kuartil pertama (25% data terendah), Q_3 adalah kuartil ketiga (75% data terendah), dan IQR menunjukkan rentang sebaran data utama [12].

3. Normalisasi

Normalisasi dilakukan untuk menyetarakan skala antaratribut agar tidak terjadi dominasi variable tertentu dalam proses pembelajaran model.

Rumus Min-Max Normalization :

$$x^i = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3)$$

Persamaan 3 menunjukkan x^i adalah nilai hasil normalisasi, x nilai data asli, x_{min} nilai minimum pada suatu fitur dan x_{max} nilai maximum pada suatu fitur [13].

4. Transformasi

Transformasi data dilakukan untuk mengurangi tingkat skewness dan meningkatkan kesesuaian data terhadap model.

a. Transformasi Logaritmik

$$x^1 = \log(x + c) \quad (4)$$

Persamaan 4 menunjukkan *transformasi logaritmik*, di mana x^1 merupakan nilai data setelah transformasi x adalah nilai data sebelum transformasi, dan c merupakan konstanta kecil (biasanya 1) yang ditambahkan untuk menghindari nilai $\log(0)$.

b. Transformasi Akar Kuadrat

$$x^1 = \sqrt{x} \quad (5)$$

Persamaan 5 menunjukkan transformasi akar kuadrat, di mana x^1 merupakan nilai data setelah transformasi dan x adalah nilai data sebelum transformasi [14].

5. Encoding

Encoding merupakan proses mengubah data kategorikal menjadi bentuk numerik agar dapat diproses oleh algoritma *machine learning*.

$$x_i \rightarrow k, k \in \{0, 1, 2, \dots, n\} \quad (6)$$

Persamaan 6 menunjukkan proses *encoding*, di mana x_1 merupakan kategori ke- i pada data kategorikal yang kemudian direpresentasikan ke dalam bentuk numerik k . Nilai k berada dalam himpunan $\{0, 1, 2, \dots, n - 1\}$, dengan n menyatakan jumlah kategori dalam variable tersebut [15].

C. Pemodelan

Pemodelan prediksi merupakan proses membangun model matematis untuk memperkirakan nilai masa depan berdasarkan pola data historis. Tahapan ini meliputi pemilihan algoritma, pelatihan model, dan evaluasi kinerja [16].

1. Decision Tree Regression

Merupakan algoritma *machine learning* berbasis struktur pohon yang digunakan untuk memprediksi nilai kontinu. Model ini bekerja dengan membagi data ke dalam beberapa subset berdasarkan atribut tertentu sehingga membentuk node akhir (leaf). Pada setiap leaf node, nilai prediksi ditentukan berdasarkan rata-rata nilai target pada node tersebut [17][18].

$$MSE_t = \frac{1}{n} = \sum_{i=1}^n (y_i - \bar{y}_t)^2 \quad (7)$$

Persamaan 7 menunjukkan perhitungan *Mean Squared Error* (MSE) pada node ke- t . Pada persamaan tersebut, n merupakan jumlah data pada node yang sedang dievaluasi, y_i adalah nilai aktual harga cabai pada data ke- t . Komponen $(y_i - \bar{y}_t)$ menunjukkan deviasi nilai aktual terhadap rata-rata pada node tersebut, sedangkan $(y_i - \bar{y}_t)^2$ merupakan kuadrat deviasi yang digunakan untuk memastikan seluruh nilai kesalahan bernilai positif dan mengukur besar variasi dalam node. Notasi $\sum_{i=1}^n$ menyatakan penjumlahan seluruh kuadrat deviasi dari data ke-1 hingga data ke- n dan $\frac{1}{n}$ menunjukkan rata-rata kuadrat deviasi pada node tersebut.

2. Random Forest Regression

Random Forest Regression merupakan metode *ensemble learning* yang menggabungkan beberapa model *Decision Tree* untuk meningkatkan akurasi dan stabilitas prediksi. Setiap pohon dibangun menggunakan subset data dan fitur yang dipilih secara acak, sehingga dapat mengurangi *overfitting* [19].

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T \hat{y}_t \quad (8)$$

Persamaan 8 menunjukkan bahwa nilai prediksi akhir \hat{y} diperoleh dengan menghitung rata-rata dari seluruh hasil prediksi setiap pohon keputusan dalam model. Di mana T menyatakan jumlah total pohon keputusan dan \hat{y}_t merupakan prediksi yang dihasil oleh pohon ke- t .

D. Evaluasi

1. Mean Absolute Error (MAE)

MAE mengukur rata-rata besarnya kesalahan absolut antara nilai aktual dan nilai prediksi. Semakin kecil nilai MAE, semakin baik kinerja model [20].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

Persamaan 9 menunjukkan perhitungan *Mean Absolute Error* (MAE) yang digunakan untuk mengukur rata-rata besarnya kesalahan absolut antara nilai actual dan nilai prediksi. Pada persamaan tersebut, n menyatakan jumlah total data pengamatan, y_i merupakan nilai aktual (nilai sebenarnya) pada pengamatan ke- i , dan \hat{y}_i adalah nilai prediksi yang dihasilkan oleh model pada pengamatan ke- i . Komponen $|y_i - \hat{y}_i|$ menunjukkan selisih absolut antara nilai aktual dan nilai prediksi. Semakin kecil nilai MAE, maka semakin baik kinerja model dalam melakukan prediksi.

2. Root Mean Square Error (RMSE)

RMSE memberikan penalti yang lebih besar terhadap kesalahan prediksi yang bernilai besar.

$$RMSE = \frac{1}{n} \sum (y_i - \hat{y}_i)^2 \quad (10)$$

Persamaan 10 menunjukkan n adalah jumlah total data observasi yang digunakan dalam pengujian model, y_i adalah nilai aktual harga cabai pada data ke- i , \hat{y}_i merupakan nilai hasil prediksi harga cabai yang dihasilkan oleh model pada data ke- i , Selisih $(y_i - \hat{y}_i)$ merupakan kesalahan nilai prediksi (error). $(y_i - \hat{y}_i)^2$ adalah kuadrat error yang digunakan untuk memastikan nilai kesalahan selalu bernilai positif serta memberikan penalti lebih besar pada kesalahan prediksi yang besar. $\sum_{i=1}^n$ adalah penjumlahan seluruh kuadrat error dari data ke-1 hingga data ke- n dan $\frac{1}{n}$ merupakan rata-rata dari keseluruhan kuadrat error, dan tanda akar kuadrat ($\sqrt{\quad}$) digunakan untuk mengembalikan satuan kesalahan ke satuan asli data harga cabai.

3. Koefisien Determinasi (R^2)

R^2 mengukur kemampuan model dalam menjelaskan variasi data. Nilai R^2 berada pada rentang 0-1. Semakin mendekati 1, Semakin baik model dalam menjelaskan variabilitas data [21].

$$R^2 = 1 - \frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{\sum_{i=1}^n |y_i - \bar{y}|^2} \quad (11)$$

Persamaan 10 menunjukkan R^2 atau *koefisien determinasi* yang digunakan untuk mengukur kemampuan model dalam menjelaskan variasi harga cabai. Pada persamaan tersebut, n menyatakan jumlah total data observasi yang digunakan dalam pengujian model, y_i merupakan nilai aktual harga cabai pada data ke- i , dan \hat{y}_i adalah nilai prediksi yang dihasilkan oleh model pada data ke- i , \bar{y} menyatakan nilai rata-rata dari seluruh data harga cabai aktual, Komponen $(y_i - \hat{y}_i)^2$ merupakan kuadrat selisih antara nilai aktual dan nilai prediksi (*residual sum of squares*), yang digunakan untuk mengukur kesalahan prediksi model, sedangkan $(y_i - \bar{y})^2$ merupakan kuadrat selisih antara nilai aktual dan nilai rata-rata data (*total sum of squares*), yang menggambarkan total variasi data harga cabai. Noatasi $\sum_{i=1}^n$ menunjukkan proses penjumlahan dari data ke-1 hingga ke- n .

III. HASIL DAN PEMBAHASAN

A. Statistik Deskriptif dan Hasil Prediksi

Tabel berikut menyajikan ringkasan statistik deskriptif harga komoditas cabai berdasarkan data historis yang digunakan dalam penelitian ini. Statistik deskriptif yang ditampilkan meliputi nilai minimum, maksimum, rata-rata (*mean*), dan standar deviasi untuk masing-masing komoditas. Penyajian ini bertujuan untuk memberikan gambaran awal mengenai pola sebaran dan tingkat variasi data sebagai dasar analisis sebelum dilakukan proses pemodelan prediktif.

TABEL II
HASIL PREDIKSI CABAI RAWIT MERAH

Komoditas Cabai Rawit Merah		
Aktual	Decision Tree Regression	Random Forest Regression
0,6727	0,4487	0,4501
0,1860	0,4351	0,4388
0,9810	0,4487	0,4501
.....
.....
0,1554	0,4081	0,4130
0,0697	0,4341	0,4348
0,7329	0,4491	0,4516

TABEL III
HASIL PREDIKSI CABAI MERAH KERITING

Komoditas Cabai Merah Keriting		
Aktual	Decision Tree Regression	Random Forest Regression
0,8385	0,5121	0,5131
0,6132	0,5132	0,5150
1,0000	0,5121	0,5131
.....
.....
0,6736	0,4472	0,4475
0,6736	0,4813	0,4812
0,5543	0,5148	0,5163

Berdasarkan Tabel II dan Tabel III, disajikan hasil perbandingan antara nilai aktual dan nilai prediksi harga cabai menggunakan metode *Decision Tree Regression* dan *Random Forest Regression*. Jumlah data yang digunakan dalam pengujian sebanyak 738 observasi dan seluruh data telah melalui proses normalisasi ke dalam rentang 0–1. Pada komoditas Cabai Rawit Merah, nilai aktual memiliki rata-rata sebesar 0,4369 dengan nilai minimum 0,0191 dan maksimum 0,9871. Rentang nilai yang cukup lebar tersebut menunjukkan adanya volatilitas harga yang tinggi selama periode pengamatan. Hasil prediksi menggunakan *Decision Tree Regression* menghasilkan rata-rata sebesar 0,4427 dengan standar deviasi 0,0177, sedangkan *Random Forest Regression* memiliki rata-rata sebesar 0,4443 dengan standar deviasi 0,0178. Nilai standar deviasi yang relatif kecil dibandingkan variasi data aktual menunjukkan bahwa kedua model menghasilkan prediksi yang lebih terkonsentrasi di sekitar nilai rata-rata, sehingga model cenderung stabil namun kurang responsif terhadap lonjakan harga yang bersifat ekstrem. Secara kuantitatif, *Random Forest Regression* menunjukkan hasil yang sedikit lebih stabil dibandingkan *Decision Tree Regression* berdasarkan nilai rata-rata dan standar deviasi yang diperoleh.

Pada komoditas Cabai Merah Keriting, nilai aktual memiliki rata-rata sebesar 0,5168 dengan nilai minimum 0,0122 dan maksimum 1,0000. Rentang nilai yang luas ini kembali mencerminkan fluktuasi harga yang signifikan. Prediksi menggunakan *Decision Tree Regression* menghasilkan rata-rata sebesar 0,5014 dengan standar deviasi 0,0199, sedangkan *Random Forest Regression* memiliki rata-rata sebesar 0,5026 dengan standar deviasi 0,0205. Dibandingkan dengan variasi data aktual, standar deviasi hasil prediksi kedua model relatif lebih kecil, yang

menunjukkan bahwa model belum sepenuhnya mampu menangkap dinamika volatilitas harga secara ekstrem. Secara umum, pendekatan berbasis pohon mampu menghasilkan prediksi yang relatif konsisten dan tidak terlalu menyimpang dari nilai rata-rata, namun masih memiliki keterbatasan dalam merepresentasikan fluktuasi harga yang tajam. Pada komoditas ini, *Random Forest Regression* juga memperlihatkan kecenderungan performa yang lebih stabil dibandingkan *Decision Tree Regression*.

B. Modeling dan Evaluasi

Evaluasi kinerja model dilakukan menggunakan tiga metrik, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan Koefisien Determinasi (R^2). Hasil evaluasi *Decision Tree Regression* disajikan pada Tabel IV, sedangkan hasil *Random Forest Regression* disajikan pada Tabel V.

TABEL IV
HASIL EVALUASI MODEL DECISION TREE REGRESSION

Evaluasi	Cabai Rawit Merah	Cabai Merah Keriting
MAE	0,2480	0,1751
RMSE	0,2741	0,2205
R^2	0,0036	-0,0139

TABEL V
HASIL EVALUASI MODEL RANDOM FOREST REGRESSION

Evaluasi	Cabai Rawit Merah	Cabai Merah Keriting
MAE	0,2481	0,1751
RMSE	0,2741	0,2204
R^2	0,0038	-0,0135

Pada *Decision Tree Regression*, komoditas cabai rawit merah menghasilkan MAE sebesar 0,2480 dan RMSE sebesar 0,2741, dengan R^2 sebesar 0,0036. Untuk cabai merah keriting diperoleh MAE sebesar 0,1751 dan RMSE sebesar 0,2205, dengan R^2 sebesar -0,0139. Nilai kesalahan yang lebih rendah pada cabai merah keriting menunjukkan bahwa model memiliki tingkat akurasi yang relatif lebih baik pada komoditas tersebut. Namun demikian, nilai R^2 yang mendekati nol bahkan negatif menunjukkan bahwa kemampuan model dalam menjelaskan variasi harga masih sangat terbatas. Secara statistik, R^2 negatif mengindikasikan bahwa performa model lebih rendah dibandingkan pendekatan sederhana yang menggunakan nilai rata-rata sebagai prediktor.

Hasil *Random Forest Regression* menunjukkan nilai evaluasi yang relatif serupa. Pada cabai rawit merah diperoleh MAE sebesar 0,2481, RMSE sebesar 0,2741, dan R^2 sebesar 0,0038, sedangkan pada cabai merah keriting diperoleh MAE sebesar 0,1751, RMSE sebesar 0,2204, dan R^2 sebesar -0,0135. Dibandingkan dengan *Decision Tree Regression*, *Random Forest Regression* menunjukkan peningkatan performa yang sangat kecil, ditandai oleh nilai R^2 yang sedikit lebih tinggi serta RMSE yang sedikit lebih rendah. Secara komparatif, perbedaan performa kedua model tidak signifikan.

Secara keseluruhan, kedua model masih memiliki keterbatasan dalam menangkap volatilitas harga cabai, sebagaimana tercermin dari nilai R^2 yang sangat rendah. Hal ini mengindikasikan bahwa variasi harga kemungkinan dipengaruhi oleh faktor eksternal lain yang belum dimasukkan dalam model. Oleh karena itu, pengembangan model dengan penambahan variabel eksternal atau pendekatan pemodelan yang lebih komprehensif dapat dipertimbangkan pada penelitian selanjutnya.

Berdasarkan Tabel VI dan Tabel VII, *Random Forest Regression* menunjukkan performa yang sedikit lebih baik dibandingkan *Decision Tree Regression* pada kedua komoditas, meskipun perbedaannya relatif kecil.

TABEL VI
PERBANDINGAN PERFORMA MODEL

Komoditas Cabai Rawit Merah			
MODEL	MAE	RMSE	R^2
<i>Decision Tree</i>	0.2480	0.2741	0.0036
<i>Random Forest Regression</i>	0.2481	0.2741	0.0038

TABEL VII
PERBANDINGAN PERFORMA MODEL

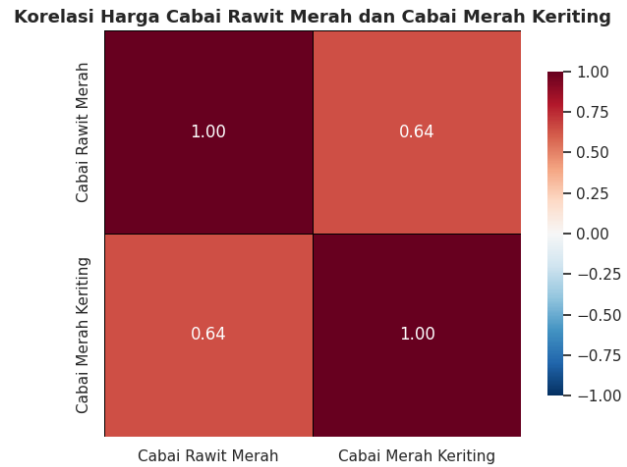
Komoditas Merah Keriting			
MODEL	MAE	RMSE	R^2
<i>Decision Tree</i>	0.1751	0.2205	-0.0139
<i>Random Forest Regression</i>	0.1751	0.2204	-0.0135

Pada komoditas cabai rawit merah, *Random Forest Regression* menghasilkan nilai R^2 sebesar 0,0038, sedikit lebih tinggi dibandingkan *Decision Tree Regression* sebesar 0,0036. Namun, nilai MAE dan RMSE kedua model hampir identik, masing-masing sebesar 0,2480 – 0,2481 untuk MAE dan 0,2741 untuk RMSE. Hal ini menunjukkan bahwa peningkatan performa melalui pendekatan *ensemble* belum signifikan secara numerik pada komoditas ini.

Pada komoditas cabai merah keriting, *Random Forest Regression* menghasilkan R^2 sebesar -0,0135, lebih mendekati nol dibandingkan *Decision Tree Regression* sebesar -0,0139. Nilai MAE kedua model sama, yaitu 0,1751, sedangkan RMSE *Random Forest Regression* sedikit lebih rendah 0,2204 dibandingkan *Decision Tree Regression* 0,2205. Meskipun peningkatan tersebut sangat kecil, hasil ini menunjukkan adanya perbaikan stabilitas prediksi pada model *Random Forest Regression*.

Secara keseluruhan, perbedaan nilai MAE dan RMSE antar model relatif kecil, sehingga peningkatan kinerja melalui pendekatan *ensemble learning* belum signifikan secara kuantitatif. Namun demikian, secara teoritis *Random Forest Regression* memiliki keunggulan dalam mengurangi variansi model dan meningkatkan stabilitas prediksi melalui penggabungan beberapa pohon keputusan. Dengan demikian, meskipun kedua model masih memiliki keterbatasan dalam menjelaskan fluktuasi harga yang volatil, *Random Forest*

Regression menunjukkan konsistensi dan performa yang sedikit lebih unggul dibandingkan *Decision Tree Regression*.

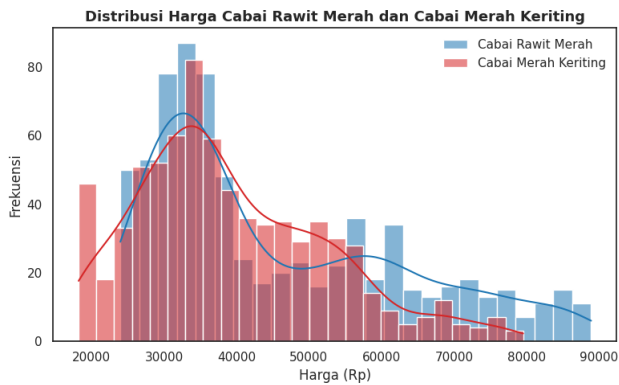


Gambar 2 Heatmap Matriks Korelasi Pearson Antara Harga Cabai Rawit Merah dan Cabai Merah Keriting

Gambar 2 menampilkan heatmap matriks korelasi Pearson antara harga Cabai Rawit Merah dan Cabai Merah Keriting. Nilai koefisien korelasi sebesar 0,64 menunjukkan adanya hubungan positif dengan tingkat kekuatan sedang hingga kuat. Hal ini mengindikasikan bahwa kenaikan harga Cabai Rawit Merah cenderung diikuti oleh kenaikan harga Cabai Merah Keriting dalam arah yang sama, meskipun hubungan tersebut tidak bersifat sempurna.

Nilai 1 pada diagonal matriks merepresentasikan korelasi sempurna suatu variabel terhadap dirinya sendiri, yang merupakan karakteristik umum dalam matriks korelasi. Nilai korelasi sebesar 0,64 menunjukkan bahwa kedua komoditas memiliki keterkaitan pergerakan harga yang cukup kuat, namun tetap dipengaruhi oleh faktor lain sehingga fluktuasinya tidak sepenuhnya identik.

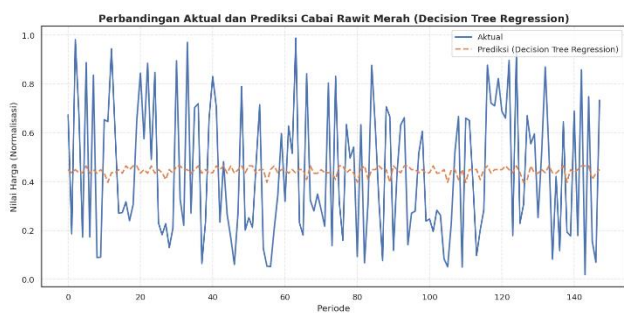
Dalam konteks pemodelan regresi, tingkat korelasi ini tidak menunjukkan adanya hubungan yang ekstrem sehingga tidak menimbulkan risiko multikolinearitas yang tinggi. Selain itu, penelitian ini menggunakan pendekatan berbasis pohon, yaitu *Decision Tree Regression* dan *Random Forest Regression*, yang relatif lebih toleran terhadap korelasi antar fitur karena proses pemisahan data dilakukan berdasarkan reduksi variansi. Dengan demikian, analisis korelasi ini menjadi dasar pendukung dalam tahap pemodelan prediksi harga selanjutnya.



Gambar 3 Distribusi Harga Cabai Rawit Merah dan Cabai Merah Keriting

Gambar 3 menampilkan distribusi harga Cabai Rawit Merah dan Cabai Merah Keriting dalam bentuk histogram yang dipadukan dengan kurva *Kernel Density Estimation* (KDE) untuk memperjelas pola sebaran data. Kedua komoditas menunjukkan konsentrasi frekuensi tertinggi pada rentang Rp 30.000–Rp 40.000/kg. Cabai Rawit Merah memiliki puncak distribusi pada kisaran Rp 32.000–Rp 38.000/kg dan memanjang hingga mendekati Rp 90.000/kg, yang menunjukkan adanya lonjakan harga pada periode tertentu serta tingkat volatilitas yang relatif tinggi.

Sebaliknya, Cabai Merah Keriting terkonsentrasi pada rentang Rp 28.000–Rp 38.000/kg dengan harga maksimum sekitar Rp 70.000–Rp 75.000/kg, yang mencerminkan fluktuasi yang relatif lebih rendah dibandingkan Cabai Rawit Merah. Kedua distribusi cenderung menceng ke kanan (*right-skewed*), yang ditandai dengan ekor distribusi yang lebih panjang pada sisi harga tinggi. Pola ini menunjukkan adanya nilai ekstrem pada periode tertentu dan sejalan dengan hasil evaluasi model sebelumnya yang memperlihatkan keterbatasan dalam menangkap lonjakan harga yang bersifat ekstrem.

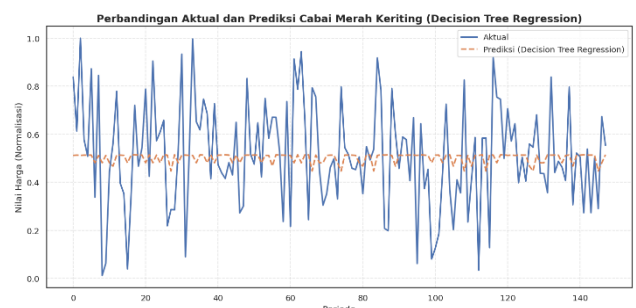


Gambar 4 Perbandingan Nilai Aktual dan Prediksi Cabai Rawit Merah Menggunakan Metode Decision Tree Regression

Gambar 4 menampilkan perbandingan antara nilai aktual dan hasil prediksi harga Cabai Rawit Merah menggunakan metode *Decision Tree Regression* pada data uji yang telah dinormalisasi. Secara visual, data aktual memperlihatkan fluktuasi yang tinggi antarperiode, yang mencerminkan volatilitas harga yang signifikan.

Sebaliknya, hasil prediksi tampak lebih stabil dan terkonsentrasi di sekitar nilai rata-rata. Rentang variasi prediksi yang lebih sempit dibandingkan data aktual mengindikasikan bahwa model belum mampu mengikuti lonjakan maupun penurunan harga yang bersifat ekstrem secara optimal. Pola ini konsisten dengan hasil evaluasi sebelumnya yang memperlihatkan nilai *koefisien determinasi* (R^2) yang sangat rendah, sehingga kemampuan model dalam menjelaskan variabilitas harga masih terbatas.

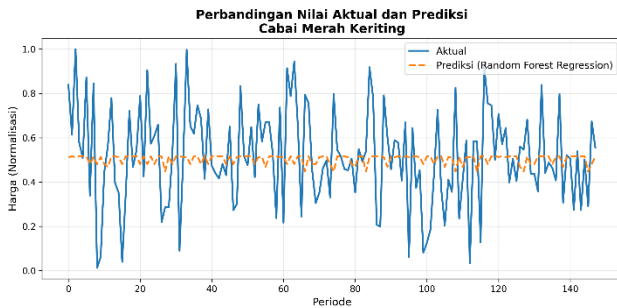
Dengan demikian, *Decision Tree Regression* pada penelitian ini mampu merepresentasikan kecenderungan umum pergerakan harga, namun belum efektif dalam menangkap dinamika volatilitas yang tinggi. Kondisi ini mengindikasikan adanya kecenderungan model untuk menyederhanakan pola data (*underfitting*) pada fluktuasi harga yang ekstrem.



Gambar 5 Perbandingan Nilai Aktual dan Prediksi Cabai Merah Keriting Menggunakan Metode Decision Tree Regression

Gambar 5 menampilkan perbandingan antara nilai aktual dan hasil prediksi harga Cabai Merah Keriting menggunakan metode *Decision Tree Regression* pada data uji yang telah dinormalisasi. Secara visual, nilai aktual menunjukkan fluktuasi harga yang cukup tajam antarperiode, dengan beberapa lonjakan dan penurunan yang signifikan.

Sebaliknya, hasil prediksi terlihat relatif lebih stabil dengan rentang variasi yang lebih sempit dibandingkan data aktual. Garis prediksi cenderung berada pada kisaran nilai tengah dan tidak sepenuhnya mengikuti pola naik-turun yang ekstrem. Kondisi ini menunjukkan bahwa model belum mampu menangkap dinamika volatilitas harga secara optimal. Perbedaan variasi yang cukup kontras antara nilai aktual dan prediksi selaras dengan nilai *koefisien determinasi* (R^2) yang rendah pada tahap evaluasi, sehingga kemampuan model dalam merepresentasikan perubahan harga yang tajam masih terbatas.

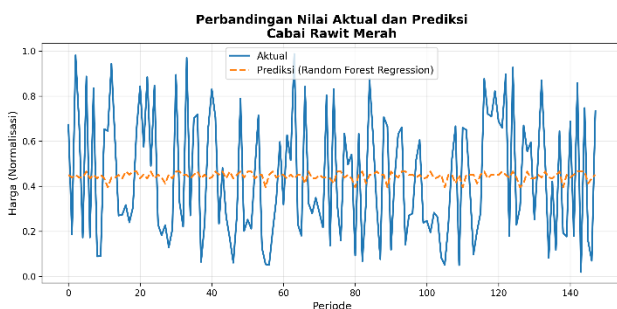


Gambar 6 Perbandingan Nilai Aktual dan Prediksi Cabai Merah Keriting Menggunakan Metode Random Forest Regression

Gambar 6, menampilkan perbandingan antara nilai aktual dan hasil prediksi harga Cabai Merah Keriting menunjukkan fluktuasi yang tinggi antarperiode, dengan beberapa lonjakan dan penurunan yang cukup tajam. Sementara itu, hasil prediksi menggunakan *Random Forest Regression* terlihat jauh lebih stabil dan cenderung bergerak dalam rentang yang sempit di sekitar nilai tengah.

Garis prediksi tidak mengikuti pola naik-turun ekstrem yang tampak pada data aktual, sehingga terdapat perbedaan variasi yang cukup jelas antara keduanya. Hal ini menunjukkan bahwa model belum mampu merepresentasikan volatilitas harga secara optimal dan lebih condong menghasilkan prediksi yang mendekati nilai rata-rata.

Kondisi tersebut sejalan dengan nilai *koefisien determinasi* (R^2) yang rendah, yang mengindikasikan bahwa kemampuan *Random Forest Regression* dalam menjelaskan variasi harga pada data uji masih terbatas. Secara keseluruhan, model mampu menggambarkan kecenderungan umum harga, namun belum responsif terhadap perubahan harga yang bersifat ekstrem.

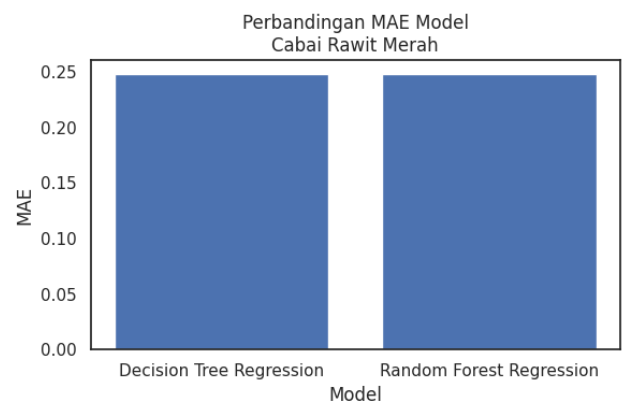


Gambar 7 Perbandingan Nilai Aktual dan Prediksi Cabai Rawit Merah Menggunakan Metode Random Forest Regression

Gambar 7 menampilkan perbandingan antara nilai aktual dan hasil prediksi harga Cabai Rawit Merah menggunakan metode *Random Forest Regression* pada data uji yang telah dinormalisasi. Secara visual, nilai aktual menunjukkan fluktuasi yang sangat tinggi antarperiode, dengan beberapa lonjakan dan penurunan yang tajam. Pola ini mencerminkan tingkat volatilitas harga yang cukup besar pada komoditas tersebut.

Sebaliknya, hasil prediksi terlihat relatif lebih stabil dan bergerak dalam rentang yang lebih sempit dibandingkan data aktual. Garis prediksi cenderung berada di sekitar nilai tengah dan tidak sepenuhnya mengikuti perubahan ekstrem yang terjadi pada data aktual. Perbedaan variasi yang cukup kontras antara nilai aktual dan prediksi mengindikasikan bahwa model belum mampu merepresentasikan dinamika volatilitas harga secara optimal.

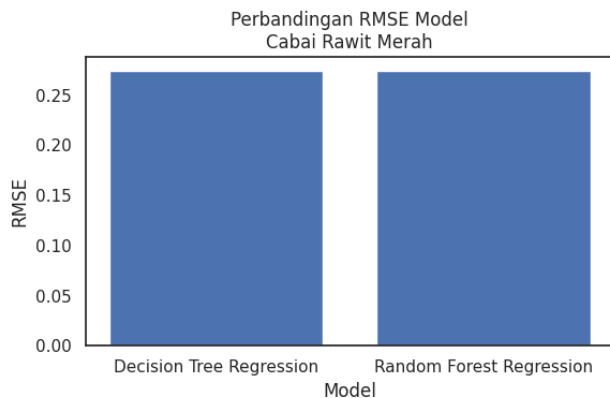
Temuan visual ini konsisten dengan hasil evaluasi yang menunjukkan nilai koefisien determinasi (R^2) yang rendah, sehingga kemampuan *Random Forest Regression* dalam menjelaskan variasi harga pada data uji masih terbatas. Secara keseluruhan, model mampu menggambarkan kecenderungan umum pergerakan harga, namun belum responsif terhadap perubahan harga yang bersifat ekstrem.



Gambar 8 Hasil Evaluasi MAE Model Prediksi Harga Cabai Rawit Merah Menggunakan Decision Tree Regression dan Random Forest Regression

Gambar 8 menunjukkan perbandingan nilai *Mean Absolute Error* (MAE) antara model *Decision Tree Regression* dan *Random Forest Regression* pada komoditas Cabai Rawit Merah. Berdasarkan grafik, kedua model menghasilkan nilai MAE yang relatif hampir sama, dengan perbedaan yang sangat kecil. Nilai MAE yang berada pada kisaran sekitar 0,24 menunjukkan bahwa rata-rata kesalahan absolut prediksi terhadap data aktual masih cukup tinggi dalam konteks data yang telah dinormalisasi. *Random Forest Regression* menunjukkan nilai MAE yang sedikit lebih rendah dibandingkan *Decision Tree Regression*, yang mengindikasikan adanya peningkatan performa, meskipun peningkatan tersebut tidak signifikan.

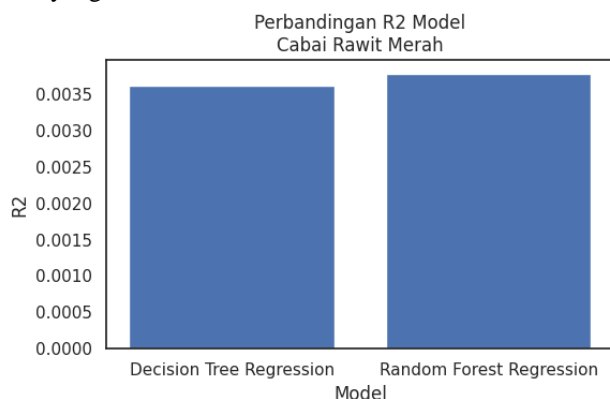
Perbedaan yang sangat tipis antara kedua model menunjukkan bahwa metode *Random Forest Regression* belum memberikan perbaikan yang substansial dibandingkan *Decision Tree* dalam memprediksi harga Cabai Rawit Merah. Hal ini mengindikasikan bahwa kompleksitas tambahan pada *Random Forest Regression* belum mampu secara efektif meningkatkan akurasi prediksi terhadap data yang memiliki tingkat volatilitas tinggi. Secara keseluruhan, kedua model masih memiliki keterbatasan dalam meminimalkan kesalahan prediksi, sehingga diperlukan pendekatan atau optimasi lanjutan untuk meningkatkan performa model.



Gambar 9 Hasil Evaluasi RMSE Model Prediksi Harga Cabai Rawit Merah Menggunakan Decision Tree Regression dan Random Forest Regression

Gambar 9 menunjukkan perbandingan nilai Root Mean Square Error (RMSE) antara model *Decision Tree Regression* dan *Random Forest Regression* pada komoditas Cabai Rawit Merah. Berdasarkan grafik, kedua model menghasilkan nilai RMSE yang hampir sama, dengan perbedaan yang sangat kecil. Nilai RMSE berada pada kisaran sekitar 0,27, yang menunjukkan bahwa rata-rata kesalahan prediksi yang telah dikuadratkan dan diakarkan masih relatif tinggi dalam konteks data yang telah dinormalisasi. *Random Forest Regression* menunjukkan nilai RMSE yang sedikit lebih tinggi dibandingkan *Decision Tree Regression*, meskipun selisihnya tidak signifikan.

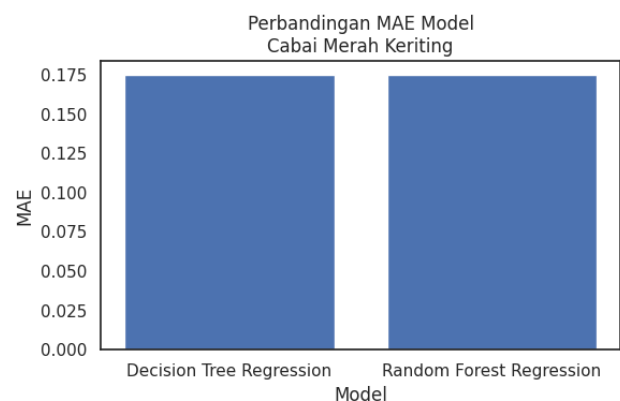
Karena RMSE memberikan penalti lebih besar terhadap kesalahan prediksi yang besar, nilai yang hampir sama pada kedua model menunjukkan bahwa keduanya memiliki kemampuan yang serupa dalam menangani kesalahan ekstrem pada data harga Cabai Rawit Merah. Hal ini mengindikasikan bahwa penambahan kompleksitas model melalui *Random Forest Regression* belum memberikan peningkatan performa yang signifikan dibandingkan *Decision Tree Regression*. Secara Keseluruhan kedua model menunjukkan performa yang relatif setara dalam meminimalkan kesalahan kuadrat rata-rata, namun belum mampu memberikan akurasi prediksi yang optimal terhadap data yang bersifat fluktuatif.



Gambar 10 Hasil Evaluasi R2 Model Prediksi Harga Cabai Rawit Merah Menggunakan Decision Tree Regression dan Random Forest Regression

Gambar 10 menunjukkan perbandingan nilai koefisien determinasi (R^2) antara model *Decision Tree Regression* dan *Random Forest Regression* pada komoditas Cabai Rawit Merah. Berdasarkan grafik, kedua model menghasilkan nilai R^2 yang sangat rendah, yaitu berada pada kisaran sekitar 0,003–0,004. Nilai R^2 yang mendekati nol menunjukkan bahwa kemampuan model dalam menjelaskan variasi data harga Cabai Rawit Merah sangat terbatas. Artinya, proporsi variansi harga yang dapat dijelaskan oleh variabel prediktor dalam model hampir tidak signifikan. *Random Forest Regression* menunjukkan nilai R^2 yang sedikit lebih tinggi dibandingkan *Decision Tree Regression*, namun peningkatan tersebut sangat kecil dan tidak menunjukkan perbedaan performa yang berarti.

Rendahnya nilai R^2 ini mengindikasikan bahwa kedua model belum mampu menangkap pola fluktuasi harga yang bersifat dinamis dan volatil. Hal tersebut dapat disebabkan oleh tingginya variasi harga antarperiode atau kurangnya variabel independen yang relevan dalam proses pemodelan. Secara keseluruhan, meskipun kedua model dapat menghasilkan prediksi, kemampuan mereka dalam menjelaskan variasi harga masih sangat terbatas, sehingga diperlukan pengembangan model lebih lanjut atau penambahan variabel prediktor yang lebih representatif.

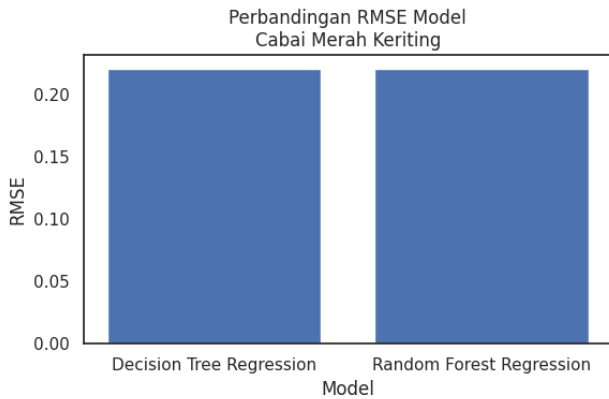


Gambar 11 Hasil Evaluasi MAE Model Prediksi Harga Cabai Merah Keriting Menggunakan Decision Tree Regression dan Random Forest Regression

Gambar 11 menunjukkan perbandingan nilai Mean Absolute Error (MAE) antara model *Decision Tree Regression* dan *Random Forest Regression* pada komoditas Cabai Merah Keriting. Berdasarkan grafik, kedua model menghasilkan nilai MAE yang hampir sama, yaitu berada pada kisaran sekitar 0,17. Nilai MAE tersebut menunjukkan bahwa rata-rata kesalahan absolut prediksi terhadap data aktual masih cukup signifikan dalam konteks data yang telah dinormalisasi. *Random Forest Regression* memiliki nilai MAE yang sedikit lebih rendah dibandingkan *Decision Tree Regression*, namun selisihnya sangat kecil dan tidak menunjukkan perbedaan performa yang signifikan.

Kesamaan nilai MAE pada kedua model mengindikasikan bahwa kompleksitas tambahan pada *Random Forest* belum memberikan peningkatan akurasi yang berarti dibandingkan

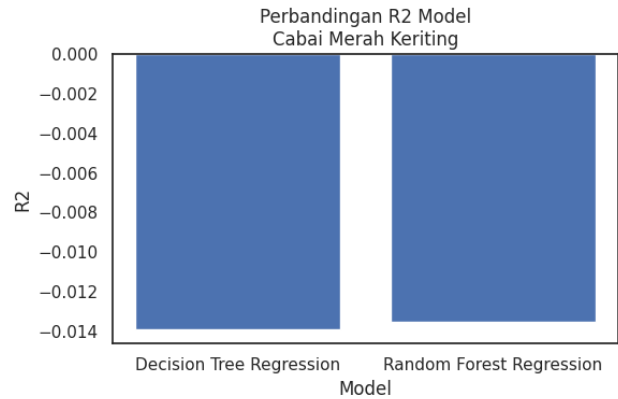
Decision Tree. Hal ini dapat disebabkan oleh karakteristik data harga Cabai Merah Keriting yang memiliki pola fluktuasi tertentu sehingga sulit ditangkap secara optimal oleh kedua model. Secara keseluruhan, kedua model menunjukkan performa yang relatif setara dalam meminimalkan kesalahan absolut, namun belum menunjukkan keunggulan yang signifikan satu sama lain.



Gambar 12 Hasil Evaluasi RMSE Model Prediksi Harga Cabai Merah Keriting Menggunakan Decision Tree Regression dan Random Forest Regression

Gambar 12 menunjukkan perbandingan nilai *Root Mean Square Error* (RMSE) antara model *Decision Tree Regression* dan *Random Forest Regression* pada komoditas Cabai Merah Keriting. Berdasarkan grafik, kedua model menghasilkan nilai RMSE yang hampir identik, yaitu berada pada kisaran sekitar 0,22. Nilai RMSE tersebut menunjukkan bahwa rata-rata kesalahan prediksi yang telah dikuadratkan dan diakarkan masih cukup signifikan, meskipun lebih rendah dibandingkan komoditas Cabai Rawit Merah. *Random Forest Regression* memiliki nilai RMSE yang sedikit lebih rendah dibandingkan *Decision Tree Regression*, namun selisihnya sangat kecil sehingga tidak menunjukkan peningkatan performa yang signifikan.

Karena RMSE memberikan penalti lebih besar terhadap kesalahan prediksi yang besar (outlier), kesamaan nilai RMSE pada kedua model mengindikasikan bahwa keduanya memiliki kemampuan yang relatif sama dalam menangani fluktuasi harga ekstrem pada Cabai Merah Keriting. Secara keseluruhan, kedua model menunjukkan performa yang setara dalam meminimalkan kesalahan kuadrat rata-rata, dan tidak terdapat perbedaan yang signifikan dalam hal akurasi prediksi.



Gambar 13 Hasil Evaluasi R2 Model Prediksi Harga Cabai Merah Keriting Menggunakan Decision Tree Regression dan Random Forest Regression

Gambar 12 menampilkan perbandingan nilai *koefisien determinasi* (R^2) antara model *Decision Tree Regression* dan *Random Forest Regression* pada komoditas Cabai Merah Keriting. Berdasarkan grafik, kedua model menghasilkan nilai R^2 yang bernilai negatif, yaitu sekitar -0,014 untuk *Decision Tree Regression* dan sekitar -0,013 untuk *Random Forest Regression*. Nilai R^2 yang negatif menunjukkan bahwa model memiliki performa yang lebih buruk dibandingkan dengan model *baseline* sederhana yang hanya menggunakan rata-rata data sebagai prediksi. Dengan kata lain, model belum mampu menjelaskan variasi data harga Cabai Merah Keriting secara memadai.

Random Forest Regression menunjukkan nilai R^2 yang sedikit lebih tinggi (kurang negatif) dibandingkan *Decision Tree Regression*, namun perbedaannya sangat kecil dan tidak signifikan secara praktis. Hal ini mengindikasikan bahwa penambahan kompleksitas model belum mampu meningkatkan kemampuan prediksi secara berarti. Secara keseluruhan, kedua model belum mampu menangkap pola variasi harga Cabai Merah Keriting dengan baik, yang kemungkinan disebabkan oleh tingginya volatilitas harga atau kurangnya variabel prediktor yang relevan dalam proses pemodelan.

IV. KESIMPULAN

Penelitian ini membangun dan membandingkan model *Decision Tree Regression* dan *Random Forest Regression* dalam memprediksi harga harian Cabai Rawit Merah dan Cabai Merah Keriting di Provinsi Jawa Tengah dengan mengintegrasikan klasifikasi temporal *weekday-weekend*. Hasil analisis korelasi Pearson menunjukkan hubungan positif yang kuat antara kedua komoditas dengan koefisien sebesar 0,64, yang mengindikasikan adanya keterkaitan pergerakan harga meskipun dipengaruhi oleh faktor yang berbeda pada masing-masing komoditas.

Evaluasi kinerja model menunjukkan bahwa *Random Forest Regression* menghasilkan performa prediksi yang relatif lebih stabil dibandingkan *Decision Tree Regression*. Pada Cabai Rawit Merah diperoleh MAE sebesar 0,2481,

RMSE sebesar 0,2741, dan R^2 sebesar 0,0038, sedangkan pada Cabai Merah Keriting diperoleh MAE sebesar 0,1751, RMSE sebesar 0,2204, dan R^2 sebesar $-0,0135$. Nilai koefisien determinasi yang rendah menunjukkan bahwa variasi harga belum sepenuhnya dapat dijelaskan oleh model berbasis variabel temporal mingguan, sehingga fluktuasi harga cabai kemungkinan dipengaruhi oleh faktor eksternal lain yang belum dimasukkan dalam pemodelan.

Secara keseluruhan, integrasi pola *weekday-weekend* dalam kerangka *machine learning* memberikan kontribusi metodologis dalam analisis harga komoditas hortikultura pada tingkat provinsi serta menunjukkan bahwa pendekatan *ensemble learning* memiliki konsistensi prediksi yang lebih baik dalam merepresentasikan kecenderungan umum harga. Penelitian selanjutnya disarankan untuk mengintegrasikan variabel eksternal yang lebih komprehensif serta mempertimbangkan pendekatan pemodelan berbasis *time series* guna meningkatkan kemampuan prediksi terhadap dinamika harga yang bersifat volatil.

DAFTAR PUSTAKA

- [1] E. Prawesti Ningrum, S. M. S. Endah Nursyamsi, and N. Siregar, "Faktor Terkait Kesenjangan Ekonomi dan Kesejahteraan," *PRIVE J. Ris. Akunt. dan Keuang.*, vol. 7, no. 2, pp. 116–126, 2024, doi: 10.36815/prive.v7i2.3480.
- [2] P. Risk, K. Variasi, and R. Harga, "Risiko Harga Cabai Pada Tingkat Petani di Kabupaten Kediri Niken," vol. 8, no. 1, pp. 181–192, 2026.
- [3] M. David, I. Cholissodin, and N. Yulistira, "Prediksi Harga Cabai menggunakan Metode Long-Short Term Memory (Case Study : Kota Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 1214–1219, 2023, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] D. D. K. Mahardika, "Peraturan Menteri Perdagangan Republik Indonesia Nomer 5 Tahun 2024," no. February, pp. 4–6, 2024.
- [5] U. Lenisa Rizki and I. Puji Astuti, "Jurnal Rekamaya Teknologi dan Komputasi," vol. 1, pp. 17–31, 2025.
- [6] I. Marina, D. Sukmawati, E. Juliana, and Z. N. Safa, "Dinamika Pasar Komoditas Pangan Strategis: Analisis Fluktuasi Harga Dan Produksi," *Paspalum J. Ilm. Pertan.*, vol. 12, no. 1, p. 160, 2024, doi: 10.35138/paspalum.v12i1.700.
- [7] L. Grebe and D. Schiereck, *Day-of-the-week effect: a meta-analysis*, vol. 14, no. 4. Springer International Publishing, 2024. doi: 10.1007/s40822-024-00293-9.
- [8] M. N. R. Ibrahim, "Price Forecasting of Shallots Using the Machine Learning Approach of Random Forest Regression Supporting Price Stabilization," *J. Keteknikan Pertan.*, vol. 13, no. 3, pp. 449–461, 2025, doi: 10.19028/jtep.013.3.449-461.
- [9] D. I. Budiarti, G. Kholijah, S. Yurinda, and B. Mardhotillah, "Prediksi Harga Cabai Rawit Hijau di Kota Jambi Menggunakan Rantai Markov," *J. Stat. Univ. Jambi*, vol. 2, no. 1, 2023, [Online]. Available: <https://online-journal.unja.ac.id/multiproximity>
- [10] S. A. Harjanto, S. Sa'adah, and G. S. Wulandari, "Export Commodity Price Forecasting in Indonesia Using Decision Tree, Random Forest, and Long Short-Term Memory," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 8, no. 4, p. 660, 2023, doi: 10.26555/jiteki.v8i4.25242.
- [11] A. Widiarti and I. Pratama, "Penanganan Missing Values dan Prediksi Data Timbunan," *RABIT J. Teknol. dan Sist. Inf. Univrab*, vol. 9, no. 2, pp. 242–251, 2024.
- [12] Y. Anzari *et al.*, "Penerapan Algoritma K-Means Untuk Klasterisasi Pola Iklim Studi Kasus : Provinsi Jambi Periode 2020-2024," vol. 6, no. 3, pp. 1037–1052, 2025, doi: 10.46576/djtechno.
- [13] P. P. Alloreng, A. Erna, M. Bagussahrir, and S. Alam, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 3, pp. 178–191, 2024, doi: 10.14421/jiska.2024.9.3.178-191.
- [14] G. Risky Pratiwi, D. Wahiddin, E. E. Awal, A. Fauzi, U. Buana, and P. Karawang, "Klasterisasi Tingkat Kemiskinan Kabupaten/Kota di Indonesia Menggunakan Algoritma K-Means dan K-Medoids," *J. Algoritma*, vol. 21, no. 2, pp. 197–208, 2024, doi: 10.33364/algoritma/v.21-2.1788.
- [15] J. W. Yuda, H. Audytra, and N. Mahmudah, "Sistem Klasifikasi Tingkat Kerusakan Kunci Motor Menggunakan Random Forest dengan Hyperparameter Tuning," *JURIKOM (Jurnal Ris. Komputer)*, vol. 12, no. 2, pp. 84–94, 2025, doi: 10.30865/jurikom.v12i2.8517.
- [16] M. Y. T. Sulistyono, E. S. Pane, E. M. Yuniarno, and M. H. Purnono, "Hybrid Significant Stroke Feature: A Novel Stroke Feature Analysis Approach for Stroke Severity Classification of EEG Signals Based on Time Domain, Frequency Domain, and Signal Decomposition Domain," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 6, pp. 1241–1267, 2024, doi: 10.22266/ijies2024.1231.91.
- [17] N. Almajid, Y. Ginting, and A. I. Ramadhan, "Penerapan Decision Tree Regression dalam Memprediksi Harga Rumah di Provinsi Jawa Barat," *J. Ris. Inform. dan Teknol. Inf.*, vol. 1, no. 3, pp. 111–115, 2024, doi: 10.58776/jriti.v1i3.64.
- [18] Y. Rokhayati, N. S. Utomo, and Sartikha, "Prediksi Kelayakan Operasional Mesin Rivet Menggunakan Regresi Linear Berganda," *J. Sustain. J. Has. Penelit. dan Ind. Terap.*, vol. 10, no. 1, pp. 10–15, 2021, doi: 10.31629/sustainable.v10i1.2336.
- [19] F. E. Penalun, A. Hermawan, and D. Avianto, "Perbandingan Random Forest Regression dan Support Vector Regression Pada Prediksi Laju Penguapan," *J. Fasilkom*, vol. 13, no. 02, pp. 104–111, 2023, doi: 10.37859/jf.v13i02.4976.
- [20] S. A. Khoiri and A. Wahid, "Jurnal Sistem dan Teknologi Informasi Indonesia Analisis Kinerja Algoritma Machine Learning dalam Prediksi Harga Cryptocurrency Performance Analysis of Machine Learning Algorithms in Cryptocurrency Price Prediction," vol. 9, no. 2, pp. 133–141, 2024.
- [21] K. D. Sanjaya, "Prediksi Harga Rumah Dengan Metode Regresi Linear Dan Support Vector Regression Di Daerah Tebet Jakarta Selatan," *J. Komput. dan Inform.*, vol. 19, no. 2, pp. 95–102, 2024, [Online]. Available: <https://journal.untar.ac.id/index.php/JKI/article/view/34588>