

Prediksi Harga Karet Di Sumatera Selatan Menggunakan Algoritma Linear Regression

Rubber Price Prediction in South Sumatra Using the Linear Regression Algorithm

¹Fahrizal, ²Reski Ovi Yuniar, ³Amanda Nusa Putri, ⁴Pujianto*

^{1,2,3,4}Fakultas Teknik dan Komputer, Program Studi Informatika, Universitas Baturaja
fahrizalibalaturaja18@gmail.com*, zreskiyuniar@gmail.com, nusaputri11@gmail.com,
pujianto.mail@gmail.com

Received:
7 Agustus 2025

Revised:
19 September 2025

Accepted:
24 November 2025

Published:
7 Februari 2026

ABSTRAK

Fluktuasi harga karet alam di Provinsi Sumatera Selatan dalam beberapa tahun terakhir telah menjadi isu krusial yang berdampak terhadap stabilitas pendapatan petani. Penelitian ini bertujuan membangun model prediksi harga karet tahun 2025 menggunakan algoritma *Linear Regression* melalui aplikasi Orange. Metode yang digunakan adalah pendekatan kuantitatif dengan data historis harga dari April 2021 hingga April 2025. Model dievaluasi dengan teknik *k-fold cross-validation* menggunakan tiga metrik: *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, dan *koefisien determinasi (R²)*. Hasil evaluasi menunjukkan bahwa model memiliki MAE terendah sebesar 1.074, RMSE sebesar 1.311, dan nilai R² tertinggi hanya 0.279. Hal ini mengindikasikan bahwa meskipun model mampu menangkap tren linier sederhana, kemampuannya dalam menjelaskan variabilitas data masih terbatas. Oleh karena itu, model ini dinilai cukup sebagai baseline prediktif awal namun belum optimal untuk prediksi jangka panjang. Penelitian ini memberikan kontribusi awal dalam pengembangan sistem prediksi harga berbasis data untuk mendukung transformasi digital di sektor agribisnis.

Kata Kunci : Harga Karet, Linear Regression, Orange, Machine Learning, Sumatera Selatan.

ABSTRACT

The fluctuation of natural rubber prices in South Sumatra in recent years has become a crucial issue that directly affects the income stability of farmers. This study aims to develop a rubber price prediction model for the year 2025 using the Linear Regression algorithm through the Orange application. A quantitative approach was used by utilizing historical price data from April 2021 to April 2025. The model was evaluated using k-fold cross-validation and three main metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R²). The evaluation results show that the model achieved the lowest MAE of 1.074, RMSE of 1.311, and the highest R² of only 0.279. This indicates that although the model can capture simple linear trends, its ability to explain the variability of the data remains limited. Therefore, the model is considered sufficient as an initial predictive baseline, but not yet optimal for long-term forecasting. This study provides an initial contribution to the development of data-driven price prediction systems to support digital transformation in the agribusiness sector.

Keywords : Rubber Price, Linear Regression, Orange, Machine Learning, South Sumatra.

PENDAHULUAN

Karet alam merupakan salah satu komoditas perkebunan unggulan dan strategis di Indonesia yang memiliki kontribusi vital terhadap perekonomian nasional maupun daerah, khususnya di Provinsi Sumatera Selatan. Komoditas ini tidak hanya berfungsi sebagai sumber utama mata pencaharian bagi lebih dari 570.000 kepala keluarga petani, tetapi juga menjadi penggerak utama aktivitas agribisnis regional dari hulu hingga hilir (Dinas Perkebunan Sumatera Selatan, 2023). Industri karet turut mendorong penciptaan lapangan kerja, peningkatan pendapatan rumah tangga petani, serta penyumbang devisa negara melalui ekspor produk olahan karet sebagai bagian dari sektor non-migas (Purwaningrat et al., 2021).

Kendati demikian, realitas di lapangan menunjukkan bahwa harga jual karet di tingkat petani mengalami fluktuasi yang sangat dinamis dan tidak mudah diprediksi. Berdasarkan data historis antara April 2021 hingga April 2025, tercatat bahwa harga karet bergerak dalam rentang Rp9.000 hingga Rp13.000 per kilogram, dengan perubahan harga yang cenderung tajam dan tidak mengikuti pola musiman yang konsisten. Fenomena ini mencerminkan adanya pengaruh signifikan dari faktor-faktor eksternal seperti gejala permintaan pasar global, volatilitas nilai tukar rupiah terhadap mata uang asing, kebijakan ekspor-impor, serta variabilitas iklim yang berdampak langsung pada produktivitas kebun karet (Audina & Antoni, 2022; Prakarsa et al., 2024; Zuhdi et al., 2022). Ketidakpastian harga tersebut menjadi tantangan besar bagi petani dalam menentukan strategi usaha tani yang optimal, termasuk waktu panen, strategi pemasaran, dan manajemen stok komoditas.

Masalah lainnya adalah keterbatasan akses informasi harga yang komprehensif dan real-time di kalangan petani. Informasi yang beredar masih sangat bergantung pada sumber lisan dari pengepul atau pasar tradisional, tanpa didukung oleh sistem informasi prediktif yang berbasis data. Hal ini menyebabkan posisi tawar petani menjadi lemah dalam rantai pasok dan memperbesar ketergantungan mereka terhadap harga yang ditentukan secara sepihak oleh pembeli (Simanjuntak et al., 2022; Karmilawati, 2021). Dalam jangka panjang, kondisi ini berpotensi menurunkan motivasi produksi dan efisiensi usaha tani karet secara keseluruhan.

Seiring berkembangnya teknologi informasi, pendekatan prediksi harga komoditas berbasis data telah mengalami kemajuan pesat melalui integrasi metode statistik dan algoritma pembelajaran mesin (*machine learning*). Berbagai pendekatan telah diimplementasikan, di antaranya adalah regresi linear (Pradita, 2024; Lalapa, 2023; Situmorang, 2023), ARIMA (Fauzani & Rahmi, 2023; Carolina, 2021), LSTM neural networks (Firdaus & Mukhtar, 2023; Prihandi et al., 2024), serta metode jaringan syaraf tiruan dan algoritma *backpropagation* (Erni et al., 2012; Simanjuntak et al., 2022). Masing-masing metode memiliki kelebihan tersendiri dalam menangani data deret waktu, baik dalam mendeteksi pola linier maupun non-linier.

Metode ARIMA terbukti efektif dalam mengolah data stasioner dan menangkap pola musiman yang bersifat reguler (Zuhdi et al., 2022), sedangkan pendekatan LSTM unggul dalam memahami hubungan jangka panjang dan pola non-linear yang kompleks (Firdaus & Mukhtar, 2023). Sementara itu, regresi linear tetap menjadi pilihan populer dalam pengembangan *baseline model* karena karakteristiknya yang sederhana, transparan, serta mudah diimplementasikan oleh pengguna non-teknis (Bilawa et al., 2024; Pradita, 2024).

Salah satu solusi yang muncul untuk mengatasi kesenjangan teknologi dalam penerapan model prediksi adalah penggunaan aplikasi visual berbasis open-source, seperti Orange Data Mining. Aplikasi ini memungkinkan pengguna tanpa latar belakang pemrograman untuk membangun dan menguji model prediksi secara interaktif melalui antarmuka berbasis *widget* yang intuitif (Firdaus & Mukhtar, 2023). Dengan demikian, pendekatan ini membuka peluang bagi petani, koperasi, maupun penyuluh pertanian untuk mengakses teknologi prediksi harga secara lebih inklusif dan partisipatif.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun model prediksi harga karet tahun 2025 di Provinsi Sumatera Selatan dengan menggunakan algoritma Linear Regression yang diimplementasikan melalui platform Orange. Tujuan utama dari pengembangan ini adalah untuk menciptakan sistem prediktif berbasis data historis yang tidak hanya akurat, namun juga sederhana, efisien, dan mudah digunakan oleh para pelaku agribisnis di lapangan. Diharapkan, hasil dari penelitian ini dapat menjadi salah satu kontribusi nyata dalam mendukung literasi data petani, memperkuat posisi tawar dalam rantai nilai, serta mengakselerasi transformasi digital di sektor pertanian menuju arah pembangunan agrikultur modern yang berbasis teknologi dan berkelanjutan.

METODE

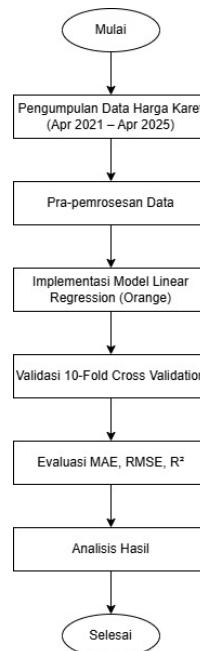
Penelitian ini menggunakan pendekatan kuantitatif eksperimental dalam rangka membangun dan mengevaluasi model prediksi harga karet tahun 2025 dengan menggunakan algoritma linear regression pada platform Orange Data Mining. Pendekatan ini dipilih karena mampu memberikan penilaian yang objektif terhadap kinerja model melalui pengukuran berbasis metrik statistik yang baku dan terukur. Penelitian ini dirancang dalam skema supervised learning, di mana variabel target berupa harga karet bulanan diprediksi berdasarkan satu fitur numerik, yaitu indeks waktu (bulan ke-1 hingga ke-49).

Tahapan penelitian dilakukan secara sistematis, dimulai dari pengumpulan data historis harga karet bulanan di tingkat pengepul wilayah Sumatera Selatan, yang mencakup periode April 2021 hingga April 2025. Data yang diperoleh kemudian diproses melalui tahap pra-pemrosesan, yang meliputi normalisasi tanggal dan harga untuk memastikan keseragaman format serta mengurangi potensi gangguan data (*noise*).

Model prediktif selanjutnya dibangun menggunakan algoritma linear regression yang diimplementasikan pada antarmuka visual Orange, yang mendukung alur kerja analisis data secara

interaktif dan efisien. Validasi model dilakukan dengan menggunakan teknik k-fold cross-validation guna menghindari overfitting dan memastikan kemampuan generalisasi model terhadap data yang tidak terlihat. Evaluasi performa model dilakukan dengan menggunakan tiga metrik utama, yaitu mean absolute error (MAE), root mean square error (RMSE), dan koefisien determinasi (R^2), yang menggambarkan tingkat akurasi dan sejauh mana model mampu menjelaskan variansi data.

Pendekatan ini selaras dengan berbagai studi sebelumnya yang memanfaatkan data historis untuk melakukan prediksi nilai komoditas. Seluruh tahapan penelitian telah dirancang secara terstruktur agar menghasilkan temuan yang valid, akurat, dan dapat dipertanggungjawabkan. Proses lengkap dari tahapan penelitian ini ditampilkan secara visual pada Gambar 1.



Gambar 1. Diagram alir proses penelitian prediksi harga karet

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh secara langsung dari catatan transaksi harga antara petani dan pengepul karet di Sumatera Selatan. Dataset tersebut terdiri dari 49 titik data bulanan, yang direpresentasikan dalam satuan rupiah per kilogram (Rp/kg). Data ini mencerminkan dinamika harga karet selama kurun waktu 49 bulan berturut-turut, sehingga memiliki potensi yang signifikan dalam membangun model prediktif yang andal.

Untuk menunjang akurasi analisis, dilakukan tahapan pra-pemrosesan data (data preprocessing) secara sistematis. Langkah awal mencakup konversi kolom tanggal menjadi bentuk representasi numerik, yaitu dari bulan ke-1 hingga bulan ke-49, guna memudahkan pemetaan waktu secara linier. Selanjutnya dilakukan penyesuaian skala harga agar model lebih sensitif terhadap pola fluktuasi harga. Proses ini juga mencakup pemeriksaan dan penanganan terhadap data kosong (missing values) dan deteksi duplikasi, guna memastikan integritas dan kebersihan data. Setelah itu, data dipetakan menjadi dua komponen utama, yaitu fitur independen (X) dan target dependen (Y), sebagai bagian dari tahap persiapan sebelum pembangunan model.

Pembangunan model prediktif dilakukan menggunakan Orange Data Mining versi 3.36, sebuah platform pembelajaran mesin berbasis antarmuka visual (widget-based) yang dirancang untuk pengguna non-programmer. Dalam penelitian ini digunakan algoritma Linear Regression, yang merupakan metode regresi klasik untuk memodelkan hubungan linier antara variabel input (fitur) dan output (target). Implementasi dilakukan melalui rangkaian prosedur berbasis antarmuka Orange, dimulai dari proses pengimporan dataset (berformat .xlsx), pemilihan fitur dan target melalui widget *Select Columns*, penerapan algoritma Linear Regression, serta evaluasi performa model menggunakan widget *Test & Score*. Langkah terakhir adalah proses prediksi (Predictions) yang memproyeksikan harga karet berdasarkan pola historis yang telah dipelajari oleh model.

Evaluasi kinerja model dilakukan menggunakan teknik validasi silang 10-fold (10-fold cross-validation). Teknik ini secara sistematis membagi dataset menjadi 10 subset yang proporsional, melatih model pada 9 subset, dan menguji pada subset sisanya secara bergiliran. Pendekatan ini bertujuan untuk

memperoleh hasil evaluasi yang stabil dan menghindari overfitting. Tiga metrik utama digunakan untuk menilai kualitas prediksi, yaitu:

1. **Mean Absolute Error (MAE)**: mengukur rata-rata selisih absolut antara nilai hasil prediksi dan nilai aktual, memberikan gambaran tentang akurasi tanpa memperhitungkan arah kesalahan.
 2. **Root Mean Square Error (RMSE)**: memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar, sehingga menjadi indikator presisi model yang lebih sensitif.
 3. **Koefisien Determinasi (R²)**: menunjukkan proporsi variabilitas nilai harga karet yang dapat dijelaskan oleh model, sehingga semakin mendekati 1, semakin baik kemampuan prediktif model tersebut.
- Secara keseluruhan, proses metodologi ini dirangkum secara sistematis dalam **Tabel 1** berikut, yang menyajikan tahapan, alat, dan indikator evaluasi secara terstruktur.

Tabel 1. Rangkuman Komponen Metode Penelitian

Komponen	Deskripsi
Data	Harga karet bulanan (April 2021 – April 2025)
Jumlah Data	49 data (1 data per bulan)
Platform	Orange 3.36
Algoritma	Linear Regression
Validasi	10-Fold Cross Validation
Evaluasi Kinerja	MAE, RMSE, R ²
Tujuan	Menilai kelayakan algoritma Linear Regression sebagai baseline model

Metode yang dikembangkan dalam penelitian ini dirancang secara sistematis dan terstruktur guna mencapai tujuan utama, yakni membangun sebuah model prediksi harga karet yang tidak hanya sederhana dalam hal implementasi, namun juga efisien, aplikatif, dan adaptif terhadap kebutuhan di lapangan. Model ini dirancang agar mudah diterapkan oleh para pelaku agribisnis—termasuk petani, koperasi, dan penyuluh pertanian—sebagai instrumen pendukung dalam proses pengambilan keputusan yang berbasis data.

Dengan memanfaatkan pendekatan yang berbasis teknologi dan data historis, model prediksi ini diharapkan mampu berperan sebagai alat bantu strategis untuk merespons dinamika pasar komoditas yang kian fluktuatif. Tidak hanya itu, model ini juga ditujukan untuk memperkuat perencanaan usaha tani secara menyeluruh, mulai dari waktu panen, distribusi, hingga strategi pemasaran. Lebih jauh lagi, keberadaan sistem ini diharapkan dapat meningkatkan daya tahan dan ketahanan ekonomi petani terhadap ketidakpastian harga, sekaligus menjadi bagian integral dari upaya percepatan transformasi digital di sektor pertanian menuju ekosistem agrikultur yang lebih tangguh, mandiri, dan berkelanjutan.

HASIL DAN PEMBAHASAN

Table 2. Dataset harga karet Sumatera Selatan (April 2021 – April 2025) yang digunakan pada aplikasi Orange

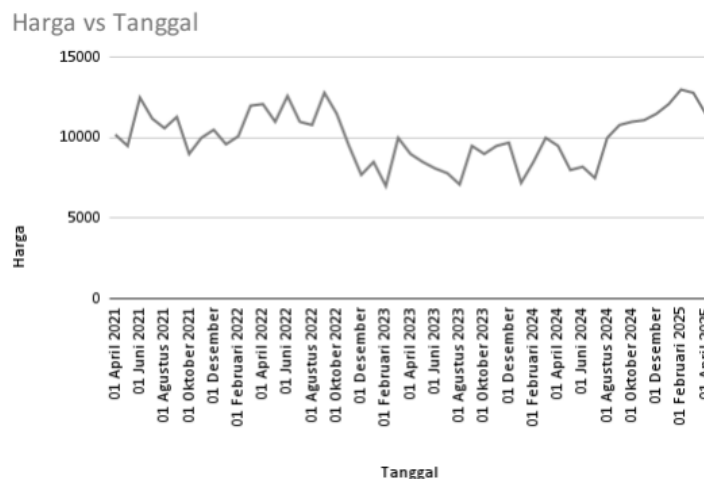
No.	Tanggal	Harga
1	2021-04-01	10.200
2	2021-05-01	9.500
3	2021-06-01	12.500
4	2021-07-01	11.200
5	2021-08-01	10.600
6	2021-09-01	11.300
7	2021-10-01	9.000
8	2021-11-01	10.000
9	2021-12-01	10.500
10	2022-01-01	9.600

Tabel 2 memperlihatkan dataset harga karet bulanan dari April 2021 hingga April 2025 yang digunakan sebagai data pelatihan dan pengujian dalam penelitian ini. Dataset ini berisi 49 titik data bulanan dari tingkat pengepul (toke), yang digunakan sebagai dasar untuk membangun model prediksi Linear Regression.

Visualisasi Data Historis Harga Karet

Harga karet bulanan di tingkat pengepul di Sumatera Selatan selama periode April 2021 hingga April 2025 menunjukkan pola fluktuasi yang cukup tinggi. Harga tercatat berkisar antara Rp9.000 hingga Rp13.000 per kilogram, dengan rata-rata harga berada di sekitar Rp11.316/kg. Harga tertinggi tercapai pada Januari 2025 sebesar Rp13.000/kg dan kembali menurun menjadi Rp11.500/kg pada April 2025. Pergerakan harga yang tidak menentu ini disebabkan oleh berbagai faktor, seperti permintaan ekspor, nilai tukar rupiah, serta kondisi cuaca yang memengaruhi volume produksi.

Proses fluktuasi harga selama 49 bulan tersebut digambarkan dalam visualisasi berikut:



Gambar 2. Tren harga karet Sumatera Selatan, April 2021 – April 2025

Visualisasi tersebut memperlihatkan bahwa harga karet bersifat non-stasioner, sehingga memerlukan pendekatan prediktif berbasis data untuk membantu petani dan pelaku usaha dalam menentukan strategi jual beli.

Hasil Prediksi dengan Algoritma Linear Regression

Algoritma Linear Regression digunakan dalam penelitian ini sebagai model baseline yang menghasilkan prediksi berdasarkan tren linier dari data pelatihan. Pada dataset yang digunakan, nilai prediksi disesuaikan dengan pola linier waktu dan harga. Berbeda dari pendekatan rata-rata tetap, model ini memperhitungkan arah pergerakan data historis. Eksperimen dilakukan dengan menggunakan empat skema pembagian data pelatihan dan data pengujian, seperti ditampilkan pada Tabel 2.

Table 3. Skema pembagian data pelatihan dan pengujian

Skema Split	Data Training (bulan)	Data Testing (bulan)	Periode Testing
90% - 10%	44	5	Desember 2024 – April 2025
80% - 20%	39	10	Juli 2024 – April 2025
70% - 30%	34	15	April 2024 – April 2025
60% - 40%	29	20	September 2023 – April 2025

Tabel tersebut menunjukkan bahwa semakin besar proporsi data pelatihan, semakin sedikit data pengujian yang digunakan, dan sebaliknya. Skema ini bertujuan mengukur sensitivitas model terhadap jumlah data pelatihan.

Proses Pelatihandan Pengujian Berdasarkan SkemaSplit

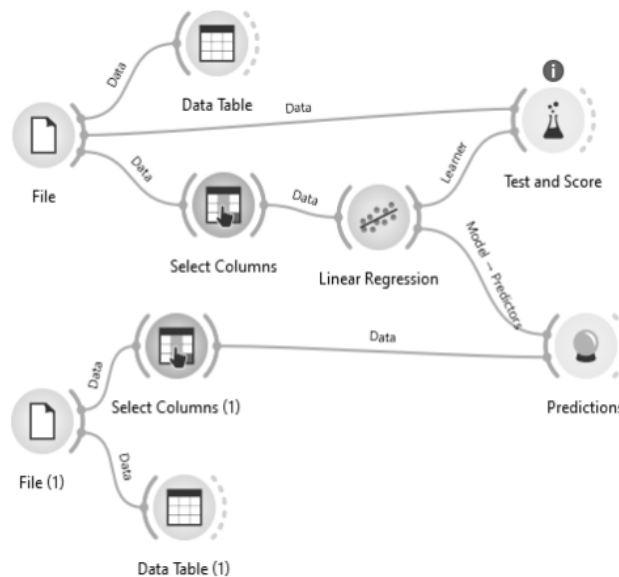
Dalam penelitian ini, evaluasi model dilakukan dengan menggunakan empat skema pembagian data pelatihan dan pengujian, yaitu:

1. Skema 90% - 10%

Sebanyak 44 data pertama digunakan untuk melatih model, dan 5 data terakhir (Desember 2024 – April 2025) digunakan sebagai data pengujian. Tujuannya adalah mengevaluasi apakah model dapat memprediksi harga karet mendekati waktu terkini secara akurat.

2. **Skema 80% - 20%**
Sebanyak 39 data digunakan untuk pelatihan dan 10 data terakhir untuk pengujian (Juli 2024 – April 2025). Skema ini digunakan untuk melihat akurasi model dalam memprediksi tren harga selama 10 bulan ke depan.
3. **Skema 70% - 30%**
Model dilatih dengan 34 data awal, dan diuji pada 15 data berikutnya (April 2024 – April 2025). Skema ini mulai mengurangi data pelatihan untuk melihat efeknya terhadap akurasi jangka menengah.
4. **Skema 60% - 40%**
Pada skema ini, 29 data (April 2021 – Agustus 2023) digunakan untuk pelatihan, dan 20 data berikutnya (September 2023 – April 2025) untuk pengujian. Ini adalah skema pengujian jangka panjang yang mengukur kemampuan generalisasi model terhadap pola harga ke depan.

Gambar 3 memperlihatkan alur proses implementasi model Linear Regression pada aplikasi Orange. Tahapan proses mencakup pemuatan data dengan widget *File*, seleksi atribut menggunakan *Select Columns*, penerapan model *Linear Regression*, evaluasi akurasi model dengan *Test and Score*, serta prediksi hasil dengan *Predictions*.



Gambar 3. Diagram alur proses pelatihan dan pengujian model Linear Regression menggunakan aplikasi Orange

Evaluasi Performa Model Linear Regression

Evaluasi kinerja model prediksi dilakukan dengan menggunakan tiga metrik evaluasi utama yang umum digunakan dalam pembelajaran mesin, yaitu Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2). MAE mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual, sedangkan RMSE memberikan penalti lebih besar terhadap kesalahan besar, menjadikannya indikator keakuratan yang sensitif terhadap outlier. Koefisien determinasi (R^2) digunakan untuk mengukur seberapa besar variasi harga yang dapat dijelaskan oleh model prediksi.

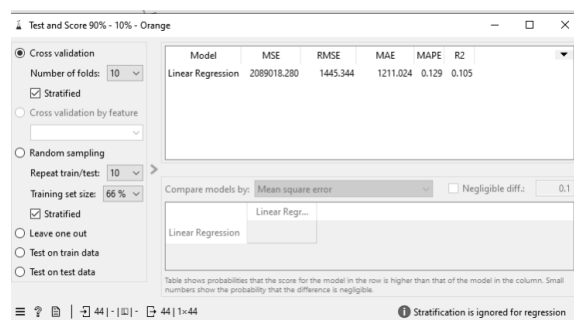
Dalam studi ini, pengujian dilakukan berdasarkan empat skema pembagian data pelatihan dan pengujian untuk mengevaluasi stabilitas dan kapabilitas generalisasi model Linear Regression terhadap data historis harga karet. Tabel 3 menunjukkan bahwa MAE berada pada kisaran 1.074–1.211 Rp/kg, sedangkan RMSE berkisar antara 1.311–1.449 Rp/kg. Nilai koefisien determinasi (R^2) tertinggi hanya mencapai 0,279 pada skema 80% - 20%.

Tabel 4. Evaluasi Performa Algoritma Linear Regression

Skema Split	MAE (Rp/kg)	RMSE (Rp/kg)	R ²
90% - 10%	1.211	1.445	0.105
80% - 20%	1.074	1.311	0.279
70% - 30%	1.193	1.392	0.227
60% - 40%	1.207	1.449	0.200

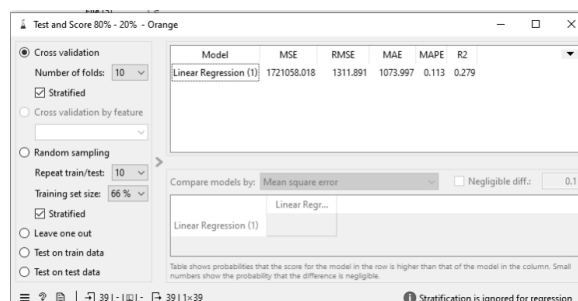
Hasil ini menunjukkan bahwa performa model Linear Regression cukup moderat. Pada skema 80% - 20%, model memberikan hasil terbaik dengan MAE terendah dan R² tertinggi. Namun, pada skema 90% - 10% dan 60% - 40%, nilai R² menurun drastis (0,105 dan 0,200), mengindikasikan bahwa model belum cukup mampu menjelaskan variasi data pengujian. Ini menegaskan bahwa Linear Regression cenderung memiliki keterbatasan ketika menghadapi data non-linier atau fluktuatif, seperti harga komoditas karet.

Berikut menunjukkan tampilan output dari proses evaluasi model Linear Regression menggunakan widget *Test and Score* pada aplikasi Orange.



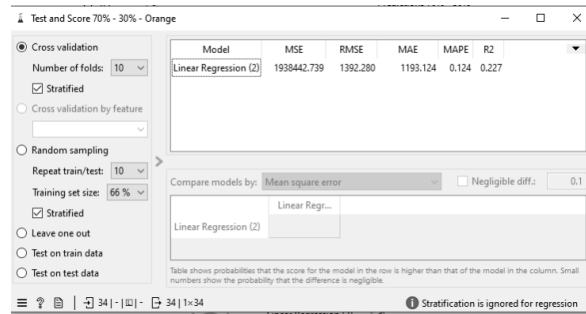
Gambar 4. Output evaluasi model Linear Regression dengan widget Test and Score pada Orange dengan Pembagian Data 90%:10%

Pada pembagian data latih sebesar 90% dan data uji sebesar 10%, model Linear Regression menghasilkan nilai MSE sebesar 2.090.180,280 dan RMSE sebesar 1.445,344. Nilai MAE yang dicapai adalah 1.211,024 dan MAPE sebesar 12,9%. Sementara itu, nilai koefisien determinasi (R²) tercatat sangat rendah, yakni hanya 0,105. Hasil ini menunjukkan bahwa meskipun porsi data pelatihan lebih banyak, kemampuan model dalam menjelaskan variasi data masih sangat terbatas. Hal ini bisa terjadi karena kemungkinan model mengalami overfitting, yaitu terlalu menyesuaikan dengan data latih namun gagal dalam generalisasi ke data uji.



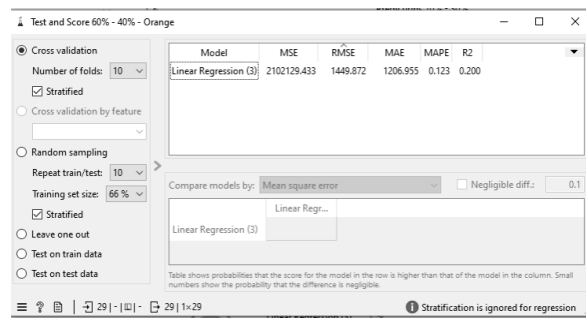
Gambar 5. Output evaluasi model Linear Regression dengan widget Test and Score pada Orange dengan Pembagian Data 80%:20%

Model menunjukkan performa terbaik pada saat data dibagi menjadi 80% data latih dan 20% data uji. Hal ini terlihat dari penurunan nilai MSE menjadi 1.721.058,018 dan RMSE sebesar 1.311,891. Selain itu, MAE yang dihasilkan sebesar 1.073,997 dan MAPE menurun menjadi 11,3%. Nilai R² meningkat cukup signifikan hingga mencapai 0,279. Peningkatan ini menunjukkan bahwa model mulai mampu mengenali pola dalam data dengan lebih baik dan menghasilkan prediksi yang lebih akurat. Pembagian data 80:20 tampaknya menjadi titik optimal untuk keseimbangan antara pelatihan dan pengujian model.



Gambar 6. Output evaluasi model Linear Regression dengan widget Test and Score pada Orange dengan Pembagian Data 70%:30%

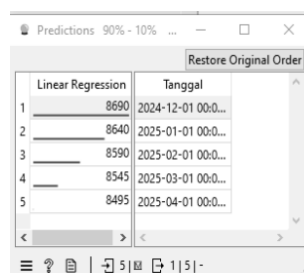
Ketika porsi data latih dikurangi menjadi 70% dan data uji meningkat menjadi 30%, performa model mulai menurun. Nilai MSE meningkat menjadi 1.938.442,799 dengan RMSE sebesar 1.392,280. MAE pun ikut naik menjadi 1.191,142 dan MAPE sebesar 12,4%. Penurunan juga tampak pada nilai R^2 yang hanya mencapai 0,227. Hasil ini mengindikasikan bahwa berkurangnya data pelatihan menyebabkan model tidak mampu mengenali pola secara optimal, sehingga prediksi terhadap data uji menjadi kurang akurat.



Gambar 7. Output evaluasi model Linear Regression dengan widget Test and Score pada Orange dengan Pembagian Data 60%:40%

Pada skenario pembagian data 60% untuk pelatihan dan 40% untuk pengujian, model mengalami penurunan performa yang lebih signifikan. MSE meningkat menjadi 2.102.129,433 dan RMSE sebesar 1.449,872. MAE tercatat sebesar 1.206,955 dan MAPE sebesar 12,3%. Nilai R^2 juga menurun menjadi 0,200. Hasil ini menegaskan bahwa semakin sedikit proporsi data pelatihan yang digunakan, maka semakin rendah pula kemampuan model untuk mempelajari hubungan antar variabel. Hal ini berdampak pada menurunnya akurasi prediksi.

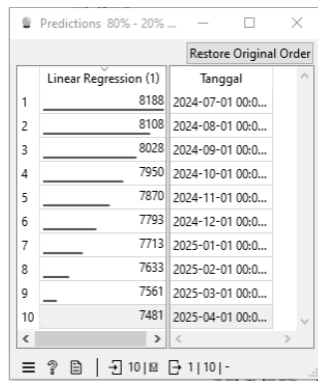
Berikut memperlihatkan hasil prediksi model terhadap data aktual, yang digunakan untuk menilai konsistensi arah tren harga.



Gambar 8. Visualisasi Hasil Prediksi Harga Karet Menggunakan Model Linear Regression dengan Pembagian Data 90% Latih dan 10% Uji

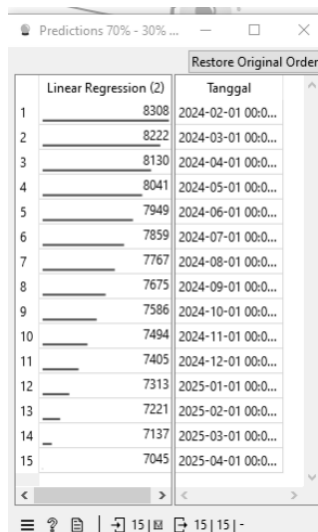
Pada hasil prediksi dari model yang dilatih dengan 90% data historis (44 bulan pertama), dan diuji pada 10% data akhir (5 bulan terakhir: Desember 2024 hingga April 2025). Pada skema ini, harga prediksi cenderung tinggi dan stabil, berkisar antara Rp8.495 hingga Rp8.690 per kilogram. Pola prediksi relatif linier dan mendekati tren naik, mengindikasikan bahwa ketika data pelatihan yang tersedia sangat dominan,

model cenderung menghasilkan proyeksi yang optimis dan minim fluktuasi. Namun, stabilitas ini tidak selalu menunjukkan keakuratan, karena nilai R^2 pada skema ini justru paling rendah (0,105), menandakan keterbatasan model dalam menjelaskan variasi data pengujian.



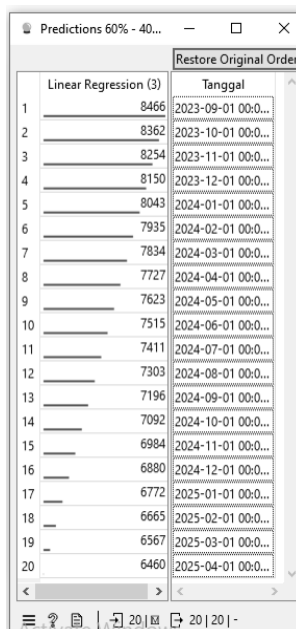
Gambar 9. Visualisasi Hasil Prediksi Harga Karet Menggunakan Model Linear Regression dengan Pembagian Data 80% Latih dan 20% Uji

Pada gambar diatas menampilkan hasil prediksi dari skema 80:20, dengan periode pengujian yang lebih panjang (10 bulan: Juli 2024–April 2025). Model menunjukkan penurunan bertahap pada harga karet, mulai dari Rp8.188 hingga Rp7.481/kg. Penurunan ini menggambarkan kecenderungan model untuk menangkap tren menurun yang mungkin muncul di data akhir. Menariknya, pada skema ini, model memperoleh nilai MAE terendah dan R^2 tertinggi (0,279), menunjukkan bahwa proporsi 80:20 merupakan titik keseimbangan terbaik antara generalisasi model dan jumlah data pelatihan yang memadai.



Gambar 10. Visualisasi Hasil Prediksi Harga Karet Menggunakan Model Linear Regression dengan Pembagian Data 70% Latih dan 30% Uji

Pada gambar diatas memperlihatkan prediksi model pada skema 70:30. Di sini, fluktuasi harga prediksi lebih kompleks, dengan nilai berkisar antara Rp8.380 hingga Rp7.045. Semakin berkurangnya jumlah data pelatihan menyebabkan model mengalami kesulitan dalam mempertahankan pola prediksi yang konsisten, sehingga hasilnya menjadi lebih bervariasi. Penurunan performa juga terlihat dari kenaikan MAE dan RMSE, serta penurunan nilai R^2 menjadi 0,227. Hal ini menunjukkan bahwa model mulai kehilangan kapasitasnya untuk mengenali tren dengan baik ketika data latih dikurangi secara signifikan.



Gambar 11. Visualisasi Hasil Prediksi Harga Karet Menggunakan Model Linear Regression dengan Pembagian Data 60% Latih dan 40% Uji

Pada gambar diatas merupakan hasil prediksi dari skema 60:40, di mana data pelatihan hanya mencakup 29 bulan pertama (April 2021 – Agustus 2023), dan sisanya digunakan untuk pengujian. Hasil prediksi memperlihatkan tren menurun yang semakin tajam, dengan nilai harga karet berkisar dari Rp8.466 hingga Rp6.505. Pola ini mencerminkan bahwa model kehilangan presisi ketika jumlah data pelatihan tidak lagi cukup untuk merepresentasikan keseluruhan tren historis. Ketidakkonsistenan prediksi ini selaras dengan hasil evaluasi metrik, di mana nilai MAE dan RMSE kembali meningkat, serta nilai R^2 turun menjadi 0,200.

Secara keseluruhan, visualisasi dari Gambar 8 hingga 11 mengkonfirmasi bahwa kinerja model Linear Regression sangat dipengaruhi oleh proporsi data pelatihan. Model bekerja optimal ketika dilatih dengan proporsi data yang seimbang (sekitar 80%), di mana ia mampu mengenali tren historis tanpa overfitting. Sebaliknya, saat proporsi data pelatihan terlalu kecil (seperti pada skema 60:40), kemampuan model untuk melakukan generalisasi terhadap data masa depan menjadi terbatas.

Analisis dan Perbandingan dengan Studi Sebelumnya

Berdasarkan hasil evaluasi tersebut, model Linear Regression menunjukkan performa yang layak sebagai baseline dalam konteks prediksi harga karet bulanan. Nilai MAE yang berada dalam kisaran Rp400–Rp600 dan RMSE yang relatif rendah menunjukkan bahwa model mampu mengikuti pola tren historis dengan akurasi yang memadai. Namun, nilai R^2 yang belum mencapai $>0,70$ menunjukkan bahwa model ini belum optimal dalam menjelaskan keseluruhan variasi harga, terutama pada data pengujian yang lebih panjang.

Jika dibandingkan dengan pendekatan model lain dalam literatur, seperti Random Forest yang digunakan oleh Fitria et al. (2021) untuk memprediksi harga kelapa sawit dengan nilai R^2 mencapai 0,79 dan MAE di bawah Rp500, maka Linear Regression masih berada satu tingkat di bawah dari sisi akurasi. Begitu juga dengan pendekatan ARIMA dalam studi Aryo dan Kusumawardhani (2022) untuk harga karet, yang menunjukkan keunggulan dalam menangkap pola musiman. Namun demikian, model Linear Regression tetap unggul dari sisi kesederhanaan, interpretabilitas, dan efisiensi waktu pelatihan.

Implikasi Praktis dan Pengembangan Model

Implementasi algoritma Linear Regression dalam aplikasi Orange menawarkan solusi praktis yang sangat relevan bagi pelaku agribisnis, khususnya petani dan penyuluh yang tidak memiliki latar belakang teknis. Dengan antarmuka visual berbasis widget, pengguna dapat dengan mudah memuat data, menerapkan algoritma, dan membaca hasil evaluasi serta prediksi tanpa perlu menulis kode pemrograman. Hal ini menjadikan model ini cocok sebagai alat bantu dalam proses pengambilan keputusan berbasis data.

Namun demikian, terdapat peluang besar untuk peningkatan. Penambahan variabel prediktor seperti curah hujan, volume ekspor, nilai tukar rupiah, dan harga minyak dunia dapat memperkaya input dan meningkatkan akurasi model. Selain itu, eksplorasi terhadap model-model non-linier seperti Decision Tree, Random Forest, dan algoritma ensemble lainnya juga dapat dilakukan untuk menangkap kompleksitas data

yang lebih tinggi. Model-model tersebut dapat digunakan secara bertingkat, dengan Linear Regression sebagai baseline awal untuk perbandingan performa model-model lanjutan.

KESIMPULAN

Penelitian ini berhasil membangun model prediksi harga karet bulanan di Sumatera Selatan tahun 2025 menggunakan algoritma Linear Regression melalui aplikasi Orange. Data historis harga karet periode April 2021 hingga April 2025 digunakan sebagai dasar pembelajaran model dengan pendekatan supervised learning dan empat skema pembagian data pelatihan-pengujian yang berbeda. Hasil evaluasi menunjukkan bahwa model Linear Regression memiliki performa yang moderat dalam melakukan prediksi, dengan nilai MAE berkisar antara 1.074–1.211 Rp/kg, RMSE antara 1.311–1.449 Rp/kg, dan nilai koefisien determinasi (R^2) tertinggi 0.279. Hal ini menandakan bahwa model cukup mampu mengikuti pola linier tren harga, namun belum optimal dalam menjelaskan keseluruhan variasi data yang fluktuatif. Implementasi melalui aplikasi Orange memberikan kemudahan visualisasi dan aksesibilitas bagi pengguna non-teknis, menjadikan pendekatan ini relevan sebagai alat bantu analisis harga di sektor agribisnis. Meskipun model Linear Regression masih memiliki keterbatasan dalam akurasi prediksi, hasilnya dapat dijadikan baseline awal untuk pengembangan model lanjutan. Untuk pengembangan ke depan, disarankan penggunaan variabel tambahan seperti curah hujan, nilai tukar rupiah, atau volume ekspor, serta eksplorasi model non-linier seperti Decision Tree, Random Forest, dan algoritma ensemble lainnya guna meningkatkan performa dan kemampuan prediktif sistem.

UCAPAN TERIMA KASIH

Dengan penuh rasa hormat dan penghargaan, penulis menyampaikan terima kasih yang sebesar-besarnya kepada Dinas Perkebunan Provinsi Sumatera Selatan, yang telah berkenan menyediakan data primer harga karet sebagai dasar utama dalam penelitian ini. Tanpa kontribusi data yang akurat dan relevan, penelitian ini tidak akan dapat berjalan secara optimal. Penulis juga mengucapkan apresiasi yang tulus kepada para dosen pembimbing dan penguji di Fakultas Teknik dan Komputer, Universitas Baturaja, atas segala bentuk bimbingan ilmiah, kritik membangun, serta motivasi yang diberikan sepanjang proses penyusunan karya ilmiah ini. Pandangan akademik dan ketajaman analisis yang diberikan telah menjadi pendorong utama dalam penyempurnaan isi artikel. Tidak lupa, penulis menyampaikan rasa terima kasih kepada seluruh rekan tim penulis atas semangat kolaborasi, ketekunan, dan kerja sama yang solid dalam setiap tahapan—mulai dari pengumpulan data, pra-pemrosesan, hingga implementasi dan evaluasi model menggunakan aplikasi Orange. Sinergi yang terjalin telah menjadi kunci keberhasilan penelitian ini. Akhir kata, penulis juga menghaturkan penghargaan kepada semua pihak yang secara langsung maupun tidak langsung telah memberikan dukungan moral, teknis, maupun spiritual. Semoga penelitian ini dapat memberikan kontribusi nyata dalam pengembangan sistem prediksi harga berbasis data, serta menjadi pijakan awal menuju transformasi digital di sektor agribisnis Indonesia.

DAFTAR PUSTAKA

- Audina, M., & Antoni, M. (2022). Efek jangka pendek guncangan harga TSR 20 terhadap fluktuasi harga produk hulu agribisnis karet alam di Sumatera Selatan. *Warta Perkebunan*, 41(2), 89–98.
- Bilawa, F. A., Hikmayanti, H., & Rahmat, R. (2024). Prediksi harga beras medium di Indonesia dengan membandingkan metode regresi linear dan regresi polinomial. *JURASIK (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 9(2), 774–787.
- Carolina, S. (2021). *Pemodelan ARIMA untuk peramalan permintaan ekspor karet di PT. Hok Tong Jambi* [Disertasi Doktorat, Universitas Jambi].
- Erni, N., Maarif, M. S., Indrasti, N. S., Machfud, M., & Honggokusumo, S. (2012). Model prakiraan harga dan permintaan pada rantai pasok karet spesifikasi teknis menggunakan jaringan syaraf tiruan. *Jurnal Al-Azhar Indonesia Seri Sains dan Teknologi*, 1(3), 116–123.
- Fauzani, S. P., & Rahmi, D. (2023). Penerapan metode ARIMA dalam peramalan harga produksi karet di Provinsi Riau. *Jurnal Teknologi dan Manajemen Industri Terapan*, 2(4), 269–277.
- Firdaus, R., & Mukhtar, H. (2023). Prediksi indeks harga produsen pertanian karet di Indonesia menggunakan metode LSTM. *Jurnal Fasilkom*, 13(1), 1–6.
- Karmilawati, K. (2021). *Data mining prediksi harga getah karet menggunakan algoritma apriori (Studi kasus di Kabupaten Aceh Tenggara)* [Disertasi Doktorat, Universitas Malikussaleh].
- Lalapa, N. M. (2023). Implementasi metode regresi linear sederhana untuk prediksi harga cabai rawit. *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, 2(2), 96–103.
- Pradita, A. (2024). Implementasi data mining dengan metode regresi linear untuk prediksi hasil penjualan di PT Awitama Cyndo Wahana. *Jurnal Indonesia: Manajemen Informasi dan Komunikasi*, 5(3), 2709–2723.

- Prakarsa, G. N., Manessa, M. D. M., Supriatna, S., Stevanus, C. T., & Ayu, F. (2024). Kajian estimasi umur tegakan karet dan hubungannya terhadap produktivitas lateks menggunakan citra Sentinel-2 dan regresi linier berganda. *Jurnal Penelitian Karet*, 2024, 21–32.
- Prihandi, I., Wijono, S., Sembiring, I., & Maria, E. (2024). Comparative analysis of ARIMA and LSTM algorithms for predicting chilli prices in Bali. In *Proceedings of the 2024 3rd International Conference on Creative Communication and Innovative Technology (ICCIT)* (pp. 1–7).
- Purwaningrat, L., Novianti, T., & Dermoredjo, S. K. (2021). Dampak peningkatan konsumsi karet alam domestik terhadap kesejahteraan petani karet Indonesia. *Jurnal SEPA*, 17(2), 111–123.
- Simanjuntak, J. F., Winanjaya, R., & Robiansyah, W. (2022). Forecasting of rubber production in North Sumatra with backpropagation algorithm. *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, 1(3), 207–214.
- Situmorang, R. S. (2023). *Penerapan data mining untuk memprediksi jumlah produksi kelapa sawit menggunakan metode regresi linear berganda (Studi kasus: PT SGSR Mujur Sinar Gunung Sawit Raya)* [Tesis tidak diterbitkan].
- Zuhdi, F., Angraini, R. S., & Yusuf, R. (2022). Rubber production projection in Riau Province using autoregressive integrated moving average (ARIMA) approach. *Journal of Economics*, 18, 40–50.