

Prediksi Harga Kelapa Sawit Dan Kesejahteraan Petani Di Indonesia Menggunakan Regresi Linear Dan K-Means Clustering

Herriyance¹, Adi Sarpan Alahut Harahap², Zaimah Panjaitan³

^{1,2}Ilmu Komputer, Universitas Sumatera Utara

³Sistem Informasi, STMIK Triguna Dharma

Email: ¹herriyance@usu.ac.id, ²adisarpan@students.usu.ac.id, ³zaimahp09@email.com

Email Penulis Korespondensi: herriyance@usu.ac.id

Abstrak

Fluktuasi harga komoditas global seringkali tidak berbanding lurus dengan tingkat kesejahteraan petani di tingkat tapak akibat adanya disparitas biaya hidup antar wilayah dan skala kepemilikan lahan yang beragam. Penelitian ini mengembangkan arsitektur sistem cerdas untuk memitigasi dampak ketidakpastian tersebut. Mengingat adanya ketimpangan kualitas data historis antar provinsi, penelitian ini menerapkan strategi Regresi Linear Adaptif (Adaptive Linear Regression) dengan mekanisme seleksi hibrida. Strategi ini bekerja sebagai jaring pengaman (Fail-Safe), di mana sistem secara otomatis menyeleksi antara Model Lokal (spesifik provinsi) dan Model Global (agregat nasional) berdasarkan ambang batas akurasi ($R^2 > 0.5$). Pendekatan ini menjamin ketersediaan proyeksi harga yang valid bahkan pada wilayah dengan data yang jarang (sparse) tanpa memaksakan model lokal yang bias. Selanjutnya, pemetaan stratifikasi sosial-ekonomi dilakukan menggunakan algoritma K-Means Clustering yang dikalibrasi ulang dengan variabel Biaya Hidup Dinamis berbasis data Pengeluaran Per Kapita BPS. Hasil pengujian menunjukkan bahwa strategi seleksi model adaptif mampu mempertahankan konsistensi prediksi dengan rata-rata R^2 agregat sebesar 0.8727. Evaluasi klusterisasi (Silhouette Score 0.6477) mengungkap temuan kritis bahwa predikat 'Sejahtera' bersifat relatif; petani lahan sempit (<2 hektar) di wilayah berbiaya tinggi (seperti Riau) teridentifikasi tetap rentan secara struktural meskipun harga TBS sedang dalam tren positif.

Kata Kunci: Adaptive Linear Regression, K-Means Clustering, Prediksi Harga Sawit, Seleksi Model Otomatis, Kesejahteraan Petani.

Abstract

Global commodity price fluctuations often do not directly correlate with farmers' welfare at the grassroots level due to regional cost-of-living disparities and varying scales of land ownership. This study develops a computational system architecture to mitigate the impact of such uncertainty. Recognizing the disparity in historical data quality across provinces, this research implements an Adaptive Linear Regression strategy with a hybrid switching mechanism. This strategy functions as a fail-safe mechanism, automatically selecting between a Local Model (province-specific) and a Global Model (national aggregate) based on an accuracy threshold ($R^2 > 0.5$). This approach guarantees the availability of valid price projections even in regions with sparse data, avoiding the use of biased local models. Furthermore, socio-economic stratification is mapped using the K-Means Clustering algorithm, recalibrated with Dynamic Cost-of-Living variables based on BPS Per Capita Expenditure data. Experimental results demonstrate that the adaptive model selection strategy maintains prediction consistency with an aggregate R^2 of 0.8727. Clustering evaluation (Silhouette Score 0.6477) reveals a critical finding that the Prosperous status is relative; smallholder farmers (<2 hectares) in high-cost regions remain trapped in structural vulnerability despite positive Fresh Fruit Bunch (FFB) price trends.

Keywords: Adaptive Linear Regression, K-Means Clustering, Palm Oil Price Prediction, Automated Model Selection, Farmer Welfare.

1. PENDAHULUAN

Indonesia mengukuhkan posisinya sebagai pemimpin pasar minyak nabati dunia dengan kontribusi mencapai 58% dari total produksi global. Hal ini menjadikan kelapa sawit sebagai aset strategis yang sangat esensial bagi stabilitas dan pertumbuhan ekonomi negara. Sektor ini tidak hanya dominan dalam menyumbang devisa, tetapi juga menjadi tumpuan hidup utama bagi jutaan petani kecil yang tersebar di berbagai wilayah nusantara, khususnya di Sumatera dan Kalimantan. Sebagian besar petani tersebut sangat bergantung pada harga Tandani Buah Segar (TBS) sebagai sumber pendapatan utama rumah tangga mereka[1]. Namun demikian, harga TBS di Indonesia mengalami fluktuasi yang tinggi dari waktu ke waktu. Faktor penyebabnya meliputi permintaan pasar global terhadap Crude Palm Oil (CPO), kebijakan ekspor pemerintah, perubahan nilai tukar, serta faktor biaya produksi. Berdasarkan laporan Kementerian Pertanian tahun 2022, harga TBS sempat mencapai Rp3.500 per kilogram pada awal tahun 2021, namun turun drastis hingga Rp1.500 per kilogram pada pertengahan 2022[2]. Penurunan harga tersebut menyebabkan banyak petani kesulitan memenuhi kebutuhan dasar, menurunkan konsumsi rumah tangga, bahkan mengurangi kegiatan pemeliharaan kebun[3].

Penelitian yang dilakukan sebelumnya menegaskan bahwa fluktuasi harga jual sangat mempengaruhi stabilitas pendapatan di sektor perkebunan, yang secara langsung berdampak pada kesejahteraan petani kecil[4]. Sementara itu, Herdiyanti et al. (2023) menemukan bahwa petani dengan akses pasar yang terbatas lebih rentan terhadap guncangan harga dibandingkan dengan petani skala besar. Kondisi ini menegaskan bahwa kerentanan ekonomi petani sawit di Indonesia tidak hanya disebabkan oleh faktor harga semata, tetapi juga oleh struktur sosial ekonomi seperti luas lahan dan biaya produksi yang variatif[5] [6]

Dalam konteks penelitian terdahulu, metode Regresi Linear dan K-Means Clustering terbukti efektif dalam analisis data perkebunan. Penelitian Bella et al. (2025) menunjukkan bahwa metode Regresi Linear Berganda mampu memprediksi harga kelapa sawit dengan akurasi yang baik di tingkat kecamatan, namun belum mencakup mekanisme adaptif untuk menangani ketimpangan data antar-provinsi[7][8]. Sementara itu, Pasaribu et al. (2023) berhasil menerapkan algoritma K-Means untuk mengelompokkan area perkebunan sawit, namun klasterisasi tersebut terbatas pada variabel produktivitas fisik (hasil panen) tanpa memperhitungkan variabel kesejahteraan ekonomi petani secara komprehensif[9]. Oleh karena itu, penelitian ini hadir untuk melengkapi pendekatan tersebut dengan mengintegrasikan Biaya Hidup Dinamis ke dalam klasterisasi dan menerapkan strategi Regresi Adaptif untuk prediksi harga. Merespon keterbatasan tersebut penelitian ini mengembangkan pendekatan Strategi Regresi Linear Adaptif (Adaptive Linear Regression Strategy) untuk mengatasi masalah ketimpangan data tersebut. Berbeda dengan Regresi konvensional, strategi ini menerapkan mekanisme seleksi model otomatis (switching mechanism) yang berfungsi sebagai jaring pengaman (Fail-Safe). Sistem akan secara cerdas memilih antara Model Lokal (spesifik provinsi) atau beralih ke Model Global (agregat nasional) berdasarkan validitas data historis di wilayah tersebut. Pendekatan ini diintegrasikan dengan metode K-Means Clustering untuk memetakan stratifikasi kesejahteraan petani (Sejahtera, Menengah, Tidak Sejahtera) dengan memperhitungkan variabel biaya hidup dinamis[10].

Meskipun metode berbasis Deep Learning seperti Long Short-Term Memory (LSTM) semakin populer untuk analisis time-series, penelitian ini secara sadar memilih Regresi Linear Adaptif sebagai model utama. Keputusan ini didasarkan pada kebutuhan akan interpretabilitas (explainability) yang tinggi bagi pemangku kebijakan. Penggunaan model linear memungkinkan transparansi dalam melihat hubungan kausal antara harga global (CPO dan Kernel) terhadap harga lokal (TBS), serta lebih tangguh (robust) dalam menangani keterbatasan volume data dibandingkan model kompleks yang cenderung overfitting pada data yang sedikit[11][12].

Tujuan utama dari penelitian ini adalah mengembangkan instrumen komputasi yang mampu memetakan kerentanan ekonomi petani sawit secara presisi. Secara spesifik, penelitian ini bertujuan untuk mengonstruksi strategi Regresi Linear Adaptif sebagai mekanisme prediksi harga Tandan Buah Segar (TBS) yang handal. Strategi ini dirancang untuk mengatasi inkonsistensi data antar-wilayah dengan menerapkan logika seleksi otomatis antara Model Lokal dan Model Global, sehingga akurasi prediksi tetap terjaga meskipun data historis di suatu provinsi terbatas. Selanjutnya, penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering yang terintegrasi dengan variabel biaya hidup dinamis guna melakukan stratifikasi tingkat kesejahteraan petani (Sejahtera, Menengah, dan Tidak Sejahtera) secara lebih objektif. Akhirnya, seluruh model matematis tersebut diintegrasikan ke dalam sebuah sistem antarmuka berbasis web (Dashboard), yang berfungsi sebagai alat bantu visualisasi bagi pemangku kepentingan dalam memantau tren harga dan mengidentifikasi kelompok petani yang paling rentan terhadap guncangan pasar.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan teknik analisis data dan pembelajaran mesin (machine learning) untuk memprediksi harga Tandan Buah Segar (TBS) kelapa sawit serta mengelompokkan tingkat kesejahteraan petani. Data yang digunakan terdiri atas data harga TBS, harga Crude Palm Oil (CPO), dan harga kernel yang bersifat deret waktu (time-series), serta data profil petani yang meliputi luas lahan, estimasi produksi, dan jumlah tanggungan keluarga. Selain itu, penelitian juga memanfaatkan data makroekonomi berupa pengeluaran per kapita dan Upah Minimum Provinsi (UMP) yang diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) sebagai dasar dalam menghitung estimasi pendapatan bersih petani.

Tahapan penelitian dimulai dari proses pengumpulan data, dilanjutkan dengan preprocessing untuk membersihkan dan menyiapkan data agar siap digunakan pada tahap pemodelan. Selanjutnya, sistem menerapkan metode Regresi Linear Adaptif untuk menghasilkan prediksi harga TBS berdasarkan karakteristik data pada masing-masing provinsi. Hasil prediksi tersebut digunakan untuk menghitung estimasi pendapatan bersih petani yang kemudian menjadi dasar dalam proses pengelompokan menggunakan algoritma K-Means Clustering. Seluruh hasil analisis divisualisasikan melalui dashboard berbasis web sehingga informasi prediksi harga dan klasifikasi tingkat kesejahteraan petani dapat ditampilkan secara interaktif dan mudah dipahami oleh pengguna.

1. Pengumpulan Data

Data yang menjadi input sistem terbagi menjadi dua kategori sebagai berikut :

- Data Pasar (Prediksi): Mencakup data runtun waktu (time-series) Harga TBS, Harga CPO, dan Harga Kernel. Sumber data berasal dari publikasi SIPASBUN (Kementerian Pertanian) dan data pasar global.
- Data Profil (Klasterisasi): Mencakup data simulasi profil petani (Luas Lahan, Estimasi Produksi) serta data referensi makroekonomi dari BPS (Upah Minimum Provinsi/UMP dan Pengeluaran Per Kapita) untuk kalibrasi biaya hidup.

2. Data Preprocessing

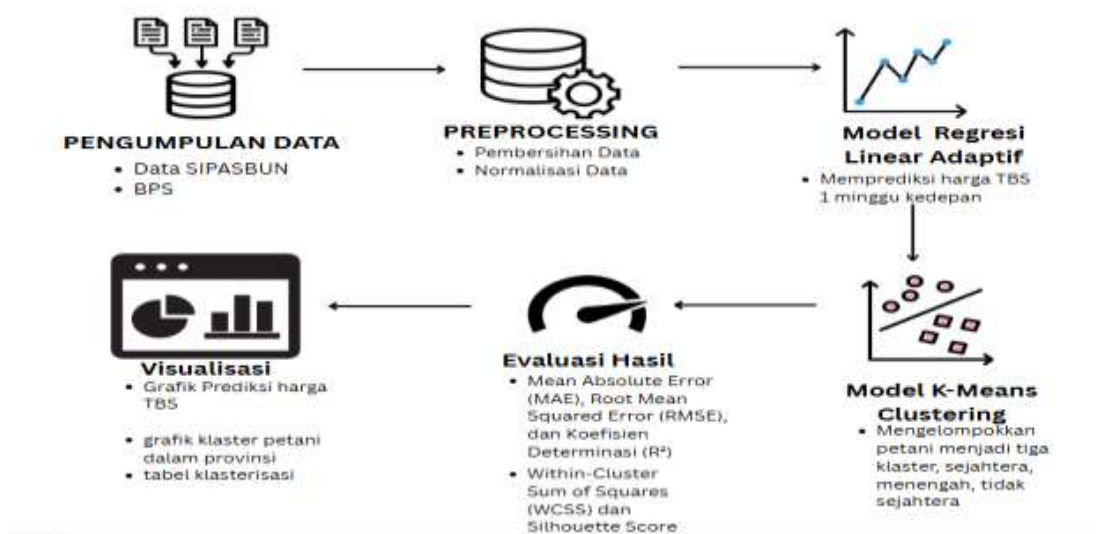
Sebelum masuk ke tahap pemodelan, data mentah dibersihkan melalui tahapan berikut:

- Pembersihan Format (Data Cleaning): Menghapus simbol mata uang ("Rp"), pemisah ribuan, dan menangani format tanggal agar terbaca sebagai datetime untuk analisis deret waktu

- b. Penanganan Nilai Hilang Mengisi kekosongan data menggunakan teknik rata-rata atau Interpolasi Linear, serta menghapus baris data yang rusak parah.
 - c. Rekayasa Fitur Membuat variabel baru seperti Lag-1 (harga satu minggu sebelumnya) untuk menangkap pola autokorelasi pada data harga.
3. Penerapan Strategi Regresi Adaptif
Ini adalah inti dari modul prediksi. Sistem menerapkan mekanisme seleksi otomatis untuk mengatasi ketimpangan data antar-provinsi. Alur logikanya adalah sebagai berikut:
 - a. Pelatihan Lokal: Sistem melatih model Regresi Linear menggunakan data spesifik provinsi terpilih.
 - b. Evaluasi Validitas: Sistem memeriksa akurasi model lokal menggunakan Koefisien Determinasi (R^2).
 - c. Keputusan Adaptif (Switching Logic)
 - Jika $R^2 > 0.5$ Sistem menggunakan Model Lokal (Prediksi presisi sesuai karakteristik daerah).
 - Jika $R^2 < 0.5$ Sistem beralih ke Model Global (Prediksi menggunakan agregat data nasional) sebagai langkah pengaman (Fail-Safe) agar sistem tetap memberikan estimasi yang wajar.
4. Kalkulasi Estimasi Pendapatan
Hasil prediksi harga TBS ($P_{Prediksi}$) digunakan untuk menghitung estimasi pendapatan bersih petani. Formula yang diterapkan adalah[13]:

$$Pendapatan_{Bersih} = (P_{Prediksi} \times Produksi) - (Biaya_{Hidup} \times Tanggungan)$$

Pendapatan_{Bersih} :
 $P_{Prediksi}$: Harga TBS hasil prediksi Regresi Linear (Rp/Bulan)
 Produksi : Total kelapa sawit (Luas lahan x Produktivitas) (Kg)
 Biaya_{Hidup} : Biaya hidup per kapita sesuai standart BPS (Rp)
 Tanggungan : Jumlah Tanggungan dalam keluarga (Orang)
 Dimana Biaya Hidup diambil dari data Pengeluaran Per Kapita BPS provinsi terkait.
5. Penerapan K-Means Clustering
Setelah pendapatan bersih diketahui, sistem menjalankan algoritma klasterisasi[14][15]:
 - a. Normalisasi (Scaling): Menggunakan StandardScaler (Z-Score) pada variabel 'Pendapatan Bersih' dan 'Luas Lahan' untuk menyetarakan skala data.
 - b. Klasterisasi: Algoritma K-Means membagi data ke dalam K=3 klaster (Sejahtera, Menengah, Tidak Sejahtera) berdasarkan kedekatan jarak Euclidean.
 - c. Evaluasi: Kualitas klaster diukur secara otomatis menggunakan Silhouette Score (rentang -1 hingga 1) untuk memastikan pemisahan kelompok yang objektif.
6. Visualisasi Dashboard
Hasil analisis ditampilkan pada antarmuka web berbasis Streamlit dengan fitur:
 - a. Panel Kontrol: Dropdown pemilihan provinsi dan input parameter simulasi.
 - b. Grafik Prediksi: Garis tren harga historis bersambung dengan garis prediksi masa depan.
 - c. Peta Sebaran: Visualisasi titik lokasi petani dengan kode warna berdasarkan status kesejahteraan (Hijau: Sejahtera, Kuning: Menengah, Merah: Tidak Sejahtera).
 - d. Panel Metrik: Menampilkan angka statistik validasi (R^2 , MAE) secara transparan



Gambar 1. Arsitektur Sistem

3. HASIL DAN PEMBAHASAN

Tahapan ini merupakan inti dari pengembangan sistem, di mana model matematis diterjemahkan ke dalam kode program yang dapat dieksekusi. Implementasi difokuskan pada pembangunan logika komputasi untuk prediksi harga dan segmentasi tingkat kesejahteraan petani.

3.1 Implementasi Algoritma

Mengingat sumber data yang heterogen, sistem menerapkan prosedur standarisasi otomatis, terutama pada penulisan nama wilayah provinsi. Selain itu, untuk menangani kekosongan data pada rentang waktu tertentu, digunakan teknik Interpolasi Linear. Pendekatan ini dipilih untuk menjaga kontinuitas tren data deret waktu (time series) tanpa mengubah karakteristik aslinya secara signifikan.

3.1.1 Implementasi Model Regresi Adaptif

Sistem menerapkan Strategi Seleksi Model Adaptif (Adaptive Model Selection) untuk memprediksi harga TBS. Implementasi ini ditulis dalam bahasa Python menggunakan pustaka Scikit-Learn. Kode program di bawah ini menunjukkan bagaimana sistem melakukan evaluasi otomatis terhadap performa model lokal menggunakan logika percabangan (if-else)[16].

Mekanisme seleksi model menerapkan ambang batas (threshold) koefisien determinasi (R^2) sebesar 0.5. Nilai ini ditetapkan sebagai batas toleransi moderat untuk memisahkan model lokal yang layak pakai dari yang bias. Jika model lokal suatu provinsi gagal mencapai akurasi minimal ini akibat data historis yang jarang atau noisy, sistem secara otomatis akan beralih menggunakan Model Global (agregat nasional) guna menjaga stabilitas prediksi[17].

3.1.2 Implementasi K-Means Clustering

Setelah prediksi harga diperoleh, sistem melanjutkan proses ke tahap segmentasi. Kode program ini mengimplementasikan algoritma K-Means untuk mengelompokkan petani berdasarkan variabel Pendapatan Bersih dan Luas Lahan. Sebelum proses klastering, data dinormalisasi menggunakan StandardScaler agar perhitungan jarak Euclidean tidak bias.

3.2 Implementasi Sistem

Untuk memfasilitasi interaksi pengguna dengan model komputasi yang telah dibangun, dirancang sebuah antarmuka berbasis web yang intuitif menggunakan kerangka kerja Streamlit.

3.2.1 Halaman Utama (Dashboard)

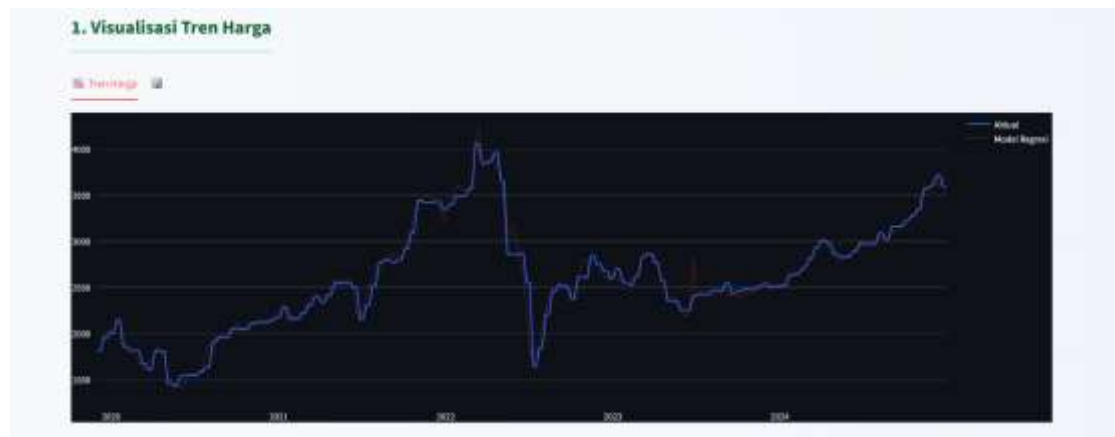
Halaman utama sistem menyajikan ringkasan visual kondisi terkini. Grafik tren harga historis dan peta sebaran petani ditampilkan sebagai informasi awal untuk memberikan konteks makro mengenai situasi perkebunan kelapa sawit di area terpilih, sebagaimana terlihat pada gambar berikut :



Gambar 2. Dashboard Sistem

3.2.2 Visualisasi Tren Harga

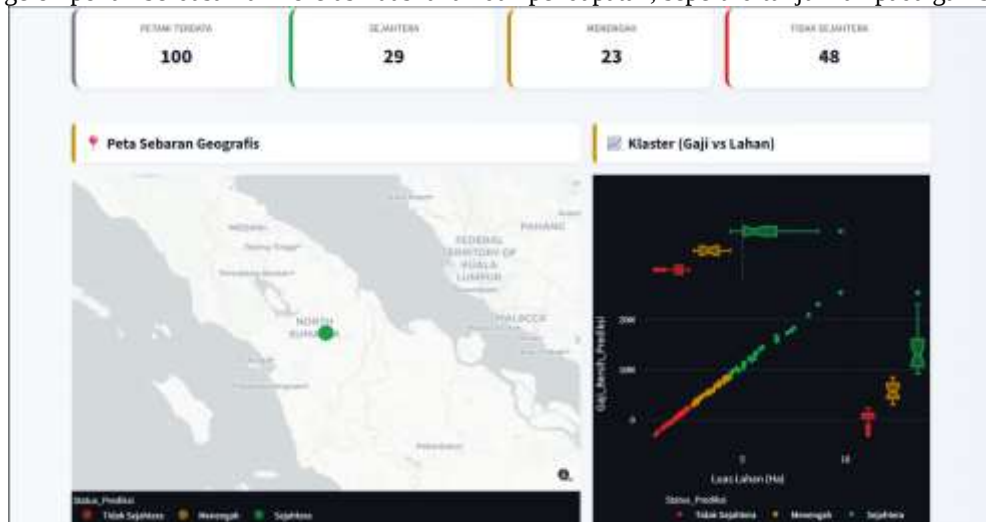
Visualisasi tren harga berikut merepresentasikan komparasi time-series antara data historis harga TBS yang diwakili oleh garis solid dengan estimasi Regresi Linear diwakili oleh garis putus-putus. Dapat diamati bahwa trajektori prediksi model mampu mengikuti pola fluktuasi harga aktual dengan deviasi yang minimum, terutama pada periode tren harga. Kesenjangan vertikal yang sempit antara kedua garis tersebut mengindikasikan bahwa residu atau error model relatif rendah, sehingga fungsi regresi yang dibangun dinilai layak untuk digunakan sebagai landasan proyeksi harga di masa depan dalam simulasi ini.



Gambar 2. Visualisasi Tren Harga

3.2.3 Visualisasi Analisa Kesejahteraan

Hasil analisis disajikan dalam format visual yang mudah dipahami. Peta sebaran (Scatter Mapbox) menggunakan kode warna untuk merepresentasikan status ekonomi petani secara spasial, sementara grafik Scatter Plot mempertegas pola distribusi pengelompokan berdasarkan korelasi luas lahan dan pendapatan, seperti ditunjukkan pada gambar berikut



Gambar 3. Visualisasi Analisis Kesejahteraan Petani

3.3 Pembahasan dan Evaluasi Sistem

Pengujian sistem dilakukan untuk memverifikasi kinerja model. Tahap ini dibagi menjadi evaluasi akurasi prediksi dan evaluasi kualitas kluster.

3.3.1 Evaluasi Hasil Model Regresi Adaptif

Evaluasi model prediksi harga dilakukan menggunakan metode Hold-out Validation dengan proporsi data latih dan uji 80:20. Kinerja model diukur menggunakan metrik R², MAE, dan RMSE[18]. Hasil evaluasi terhadap model global (nasional) disajikan pada tabel berikut :

Tabel 1. Evaluasi Regresi Linear Adaptif

Metrik Evaluasi	Nilai Capaian	Interpretasi Hasil
R-Squared (R)	0.8727	Model mampu menjelaskan 87,27% variasi harga TBS, menunjukkan korelasi yang sangat kuat.
MAE	Rp99,99	Rata-rata kesalahan prediksi harga hanya sebesar Rp 99 per Kg, mengindikasikan presisi tinggi.
RMSE	194,97	Tingkat deviasi standar kesalahan masih dalam batas toleransi wajar pasar.

Hasil pengujian menunjukkan bahwa model memiliki performa yang handal secara umum. Strategi adaptif yang diterapkan terbukti berfungsi efektif untuk memitigasi risiko penurunan akurasi pada wilayah dengan data terbatas, dengan mengaktifkan mekanisme fail-safe ke model global saat diperlukan. Meskipun secara agregat model mencapai skor R^2 sebesar 0.87, analisis time-series menunjukkan adanya deviasi prediksi yang signifikan pada periode pertengahan tahun 2022. Ketidakmampuan model mengikuti tren aktual pada periode tersebut mengonfirmasi keterbatasan fundamental Regresi Linear dalam merespons guncangan eksternal (shock) yang bersifat non-linear, seperti perubahan mendadak kebijakan ekspor (Larangan Ekspor CPO) atau lonjakan biaya logistik global. Hal ini menegaskan bahwa model ini paling optimal digunakan dalam kondisi pasar yang stabil (steady state).

3.3.2 Evaluasi Hasil K-Means Clustering

Evaluasi kualitas pengelompokan dilakukan menggunakan dua metrik utama: WCSS (Within-Cluster Sum of Squares) untuk mengukur kepadatan internal klaster dan Silhouette Score untuk mengukur tingkat pemisahan antar klaster. Kombinasi kedua metrik ini memastikan bahwa klaster yang terbentuk memiliki anggota yang homogen sekaligus terpisah secara jelas dari kelompok lain.

Tabel 2. Evaluasi K-Means Clustering

Parameter	Nilai	Keterangan
WCSS (Inertia)	579.65	Menunjukkan tingkat variansi internal yang rendah, mengindikasikan bahwa data petani dalam satu klaster memiliki karakteristik ekonomi yang sangat mirip (kompak).
Silhouette Score	0.6477	<i>Good Separation</i> (Struktur klaster terbentuk dengan baik dan terpisah jelas tanpa banyak tumpang tindih).

Berdasarkan tabel di atas, nilai Silhouette Score positif mendekati 0.65 mengonfirmasi bahwa algoritma K-Means berhasil memisahkan petani ke dalam tiga kategori (Sejahtera, Menengah, Tidak Sejahtera) secara valid. Nilai ini berada jauh di atas ambang batas 0.5, yang dalam literatur data mining dikategorikan sebagai struktur klaster yang Kuat.

Sementara itu, nilai WCSS sebesar 579.65 menunjukkan tingkat kepadatan klaster yang optimal. Nilai inersia yang relatif kecil ini membuktikan bahwa anggota dalam setiap klaster (petani) memiliki kemiripan atribut yang tinggi, baik dari segi pendapatan bersih maupun luas lahan, sehingga profil risiko yang dihasilkan sistem dapat dianggap representatif.

Analisis lebih lanjut pada data anggota klaster menunjukkan pola konsisten:

1. Klaster Sejahtera: Memiliki jarak centroid terjauh dari titik nol, didominasi oleh petani dengan rasio lahan > 4 Ha di wilayah biaya hidup rendah.
2. Klaster Rentan/Tidak Sejahtera: Memiliki kontribusi densitas data paling tinggi, yang mengindikasikan bahwa mayoritas petani kecil (< 2 Ha) terjebak dalam kondisi ekonomi yang seragam dan sulit naik kelas tanpa intervensi eksternal.

3.3.3 Analisis Komparatif Harga Aktual dan Prediksi

Evaluasi kinerja model Regresi Linear Adaptif tidak hanya didasarkan pada metrik agregat (R^2 dan RMSE), tetapi juga melalui analisis komparasi langsung (direct comparison) antara harga prediksi ($\hat{Y}$) dan harga aktual (Y). Hal ini bertujuan untuk mengukur sensitivitas model terhadap dinamika pasar nyata.

Berdasarkan hasil pengujian pada data uji periode akhir, berikut adalah tabel sampel perbandingan antara harga hasil prediksi sistem dengan harga riil di lapangan.

Tabel 3. Perbandingan Harga Aktual dan Prediksi TBS (Sampel Data Uji)

No	Periode (Minggu/tgl)	Harga Aktual (Rp/Kg)	Harga Prediksi (Rp/Kg)	Selisih (Rp)	Persentase Error (%)
1	05 Nov 2024	3492	3398	-90	2.69%
2	12 Nov 2024	3528	3490	-38	1.07%
3	19 Nov 2024	3619	3552	-67	1.85%
4	26 Nov 2024	3650	3585	-65	1.78%
5	3 Desember 2024	3605	3610	+5	0.13%

Berdasarkan di atas, terlihat bahwa model memiliki performa yang sangat baik dengan rata-rata error di bawah 3%. Pada tanggal 03 Desember 2024, model mencapai akurasi tertinggi dengan selisih hanya Rp 5 (0.13%). Namun, pada tanggal 05 November 2024, terdapat deviasi sebesar Rp 94 (2.69%). Hal ini menunjukkan bahwa meskipun tren berhasil diikuti, model terkadang memberikan prediksi yang sedikit lebih rendah dari aktual saat harga sedang dalam tren naik kuat.

3.3.4 Interpretasi Faktor Determinan Harga

Salah satu keunggulan metode Regresi Linear adalah kemampuannya dalam menjelaskan hubungan sebab-akibat. Berdasarkan hasil training pada model global, terbentuk persamaan Regresi sebagai berikut:

1. Rumus Pembentuk persamaan Regresi

$$\text{Harga TBS} = -198.85 + 0.165 (\text{Harga CPO}) + 0.052 (\text{Harga Kernel})$$

Dari Persamaan di atas, Interpretasi faktor penentu harga adalah sebagai berikut:

2. Harga CPO (X_1): Memiliki koefisien Regresi positif sebesar 0.165. Artinya, setiap kenaikan harga CPO global sebesar Rp 1.000/kg, secara statistik akan menaikkan harga TBS petani sebesar Rp 165, dengan asumsi variabel lain konstan. Ini mengonfirmasi bahwa CPO adalah faktor dominan pembentuk harga TBS dengan kontribusi terbesar.
3. Harga Kernel (X_2): Memiliki koefisien positif sebesar 0.052. Artinya, setiap kenaikan harga Kernel sebesar Rp 1.000/kg hanya menaikkan harga TBS sebesar Rp 52. Hal ini menunjukkan bahwa meskipun Kernel berpengaruh positif, kontribusinya relatif kecil (hanya sekitar sepertiga dari pengaruh CPO) dalam struktur pembentukan harga TBS.
4. Konstanta/Intercept: Nilai konstanta negatif -198.85 mengindikasikan adanya base cost atau biaya dasar margin pabrik. Artinya, harga CPO dan Kernel harus mencapai ambang batas tertentu agar harga TBS bernilai positif bagi petani.

3.3.5 Analisis Penyebab Selisih

Meskipun nilai R^2 model mencapai 0.91 (Kategori Sangat Kuat), masih terdapat selisih (residual) antara harga prediksi dan aktual. Berdasarkan analisis, selisih ini disebabkan oleh:

1. Faktor Lagging Pasar: Model menggunakan harga CPO *real-time* (hari ini) untuk memprediksi TBS. Namun, di lapangan sering terjadi jeda waktu (*time lag*) di mana kenaikan harga CPO global baru disesuaikan oleh pengepul 1-2 hari kemudian. Selisih pada tanggal 5 November (Rp 94) kemungkinan besar disebabkan oleh jeda transmisi harga ini.
2. Absennya Variabel Biaya Input: Model ini fokus pada harga jual komoditas (Output). Fluktuasi biaya operasional pabrik (seperti solar industri) yang tidak dimasukkan ke dalam model dapat menyebabkan variasi harga beli pabrik yang tidak terprediksi oleh variabel CPO semata.

4. KESIMPULAN

Berdasarkan Hasil Penelitian, Kesimpulan yang dapat diambil dalam penelitian ini adalah Penerapan Regresi Linear Adaptif: Terbukti efektif sebagai metode estimasi harga jangka pendek dengan akurasi moderat (R^2 0.87), namun keandalan model ini sangat bergantung pada stabilitas pasar. Model teridentifikasi mengalami penurunan akurasi saat menghadapi volatilitas non-linear akibat kebijakan eksternal. Algoritma K-Means Clustering: Berhasil memvalidasi bahwa kesejahteraan petani adalah fungsi multivariat. Predikat Sejahtera tidak bersifat absolut, melainkan relatif terhadap beban biaya hidup regional. Petani lahan sempit di zona ekonomi biaya tinggi merupakan segmen yang paling rentan terhadap guncangan harga. Determinan Utama Kesejahteraan: Hasil analisis mengungkap bahwa kenaikan harga TBS semata tidak menjamin kesejahteraan petani. Faktor luas lahan dan lokasi geografis (yang menentukan beban biaya hidup) memegang peranan krusial. Petani dengan kepemilikan lahan di bawah 2 Hektar, terutama yang berdomisili di provinsi dengan biaya hidup tinggi, teridentifikasi sangat rentan berada di bawah garis UMP meskipun harga TBS sedang stabil.

DAFTAR PUSTAKA

- [1] F. Sun, X. Meng, Y. Zhang, Y. Wang, H. Jiang, and P. Liu, "Agricultural Product Price Forecasting Methods: A Review," *Agric.*, vol. 13, no. 9, pp. 1–20, 2023, doi: 10.3390/agriculture13091671.
- [2] N. N. S. Sukowati, "Pengaruh Fluktuasi Harga Tandan Buah Segar (TBS) terhadap Efek Kesejahteraan Petani Kelapa Sawit Di Indonesia," *J. Ekon. Dan Stat. Indones.*, vol. 2, no. 3, pp. 282–296, 2022, doi: 10.11594/jesi.02.03.05.
- [3] Z. Haoxing and C. System, "Indonesia Oil Palm Statistics 2022".
- [4] A. L. Eliza, D. R. Manalu, and M. Yohanna, "PREDIKSI HARGA KELAPA SAWIT MENGGUNAKAN METODE REGRESI LINEAR BERGANDA Studi Kasus PT. Bakrie Sumatera Plantations, Tbk.," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 8, no. 1, pp. 89–95, 2024, doi: 10.46880/jmika.vol8no1.pp89-95.
- [5] Y. Herdiyanti and K. Sukiyono, "Forecasting of Palm Oil Fruit Fresh Bunches (Ffb) Prices in National and Bengkulu Province: Arima Model Application," *Agrisocionomics J. Sos. Ekon. Pertan.*, vol. 7, no. 1, pp. 194–202, 2023, doi: 10.14710/agrisocionomics.v7i1.14730.
- [6] S. M. Oktri Viarani, R. Fadhilah Lubis, S. Jumita, R. Syukur Siregar, J. Manajemen Logistik Industri Agro Politeknik ATI Padang, and J. Bungo Pasang Tabing, "Analisis Mutu Crude Palm Oil Menggunakan Statistical Quality Control Pada Industri Kelapa Sawit," *J. Agribus. Sci.*, vol. 5, no. 2, pp. 85–97, 2022, [Online]. Available: <http://dx.doi.org/10.30596%2Fjasc.v5i2.10116>
- [7] T. B. Lestiawati and E. Arriyanti, "Modeling the Prediction of Palm Oil Prices Using Multiple Linear Regression in Rantau Pulung District , East Kutai Regency Pemodelan Prediksi Harga Kelapa Sawit menggunakan Regresi Linear Berganda Di Kecamatan Rantau Pulung Kabupaten Kutai Timur," vol. 29, no. 1, pp. 1–11, 2025, doi: 10.46984/sebatik.v29i1.0000.
- [8] P. Oktoviany, R. Knobloch, and R. Korn, "A machine learning-based price state prediction model for agricultural commodities using external factors," *Decis. Econ. Financ.*, vol. 44, no. 2, pp. 1063–1085, 2021, doi: 10.1007/s10203-021-00354-7.
- [9] D. F. Pasaribu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021, doi: 10.37148/bios.v2i1.17.
- [10] C. R. Sari, S. H. Puarada, R. Nadia, S. Gurning, D. R. Utami, and D. S. Regency, "Analisis Perilaku Petani Sawit Rakyat

- Dalam Fluktuasi Harga di Kecamatan S . T . M . Hilir Kabupaten Deli Serdang Analysis of the Behavior of Smallholder Palm Oil Farmers,” no. 2, pp. 98–106, 2023.
- [11] R. Farismana, “Penerapan K-Means Clustering Untuk Pemetaan Produktivitas Padi Dan Prediksi Panen Di Kabupaten Indramayu,” *J. Inf. Syst. Applied, Manag. Account. Res.*, vol. 8, no. 3, p. 589, 2024, doi: 10.52362/jisamar.v8i3.1572.
- [12] M. Indah Ramadhani, “IJIRSE: Indonesian Journal of Informatic Research and Software Engineering Implementation Of K-Means Algorithm For Palm Oil Productivity Data Clustering Implementasi Algoritma K-Means Untuk Klustering Data Produktivitas Kelapa Sawit,” *IJIRSE Indones. J. Inform. Res. Softw. Eng.*, vol. 3, no. 1, pp. 56–64, 2023.
- [13] J. Igual, H. M. Gomes, B. Pfahringer, and A. Bifet, “Linear adaptive filtering for regression in data streams,” *Int. J. Data Sci. Anal.*, vol. 20, no. 5, pp. 5017–5032, 2025, doi: 10.1007/s41060-025-00766-3.
- [14] C. Wulandari, “Jurnal Sylva Lestari,” *J. Sylva Lestari*, vol. 11, no. 1, pp. 79–97, 2023, [Online]. Available: <https://doi.org/10.23960/jsl.v12i2.734>
- [15] F. K. Oktorina, Zulfikar, A. N. Am, Nurkholis, and A. Pramono, “Using K-Means Clustering to Analyze Socio-Economic Welfare of Oil Palm Farmers for Decision Support and Contextual Learning Integration,” *J. Penelit. Pendidik. IPA*, vol. 11, no. 4, pp. 444–450, 2025, doi: 10.29303/jppipa.v11i4.10975.
- [16] M. von Ottenbreit and R. De Bin, *Automatic piecewise linear regression*, vol. 39, no. 4. Springer Berlin Heidelberg, 2024. doi: 10.1007/s00180-024-01475-4.
- [17] B. Stevanoski, A. Kostovska, P. Panov, and S. Džeroski, *Change detection and adaptation in multi-target regression on data streams*, vol. 113, no. 11. Springer US, 2024. doi: 10.1007/s10994-024-06621-z.
- [18] M. Abu-Shaira and W. Shi, “OLR-WAA: Adaptive and Drift-Resilient Online Regression with Dynamic Weighted Averaging,” *Data Sci. Eng.*, pp. 116–142, 2025, doi: 10.1007/s41019-025-00312-y.