

Prediksi Produksi Tanaman Perkebunan di Sumatera Selatan Menggunakan Algoritma Linear Regression

Aria Juantito¹, M. Eka Wahyudi^{2*}, Fajar Andrea Mukti³, Pujiyanto⁴

¹ Fakultas Teknik dan Ilmu Komputer
Universitas Baturaja Ogan Komering
Ulu, Sumatera Selatan, Indonesia

¹ juantito.aria@gmail.com

² Fakultas Teknik dan Ilmu Komputer
Universitas Baturaja Ogan Komering
Ulu, Sumatera Selatan, Indonesia

^{2*} ekawhydi12@gmail.com

³ Fakultas Teknik dan Ilmu Komputer
Universitas Baturaja Ogan Komering
Ulu, Sumatera Selatan, Indonesia

³ fajarandreamukti@icloud.com

⁴ Fakultas Teknik dan Ilmu Komputer
Universitas Baturaja Ogan Komering
Ulu, Sumatera Selatan, Indonesia

⁴ pujiyanto.mail@gmail.com

Abstrak— Sektor perkebunan merupakan salah satu sektor penting yang berkontribusi terhadap perekonomian Provinsi Sumatera Selatan melalui komoditas utama seperti kelapa sawit, karet, kelapa, kopi, dan kakao. Variasi produksi yang terjadi setiap tahun menyebabkan perlunya prediksi produksi sebagai dasar dalam mendukung perencanaan dan pengambilan keputusan di bidang pertanian. Penelitian ini bertujuan untuk memprediksi produksi tanaman perkebunan di 17 kabupaten/kota di Provinsi Sumatera Selatan menggunakan algoritma Linear Regression. Data yang digunakan merupakan data historis produksi tanaman perkebunan periode 2021–2025. Tahapan penelitian mengikuti metode Knowledge Discovery in Databases (KDD) yang meliputi seleksi data, praproses, transformasi, pemodelan, dan evaluasi. Proses pemodelan dilakukan menggunakan aplikasi Orange Data Mining dengan skema evaluasi 10-Fold Cross Validation. Hasil evaluasi menunjukkan bahwa model Linear Regression menghasilkan nilai koefisien determinasi (R^2) sebesar 1,000, Root Mean Square Error (RMSE) sebesar 0,373, dan Mean Absolute Error (MAE) sebesar 0,249. Nilai evaluasi tersebut menunjukkan bahwa model mampu mengikuti pola data yang digunakan dengan tingkat kesalahan prediksi yang rendah. Hasil penelitian ini menunjukkan bahwa algoritma Linear Regression dapat digunakan sebagai salah satu pendekatan untuk memprediksi produksi tanaman perkebunan di Provinsi Sumatera Selatan berdasarkan data historis yang tersedia.

Kata kunci— Linear Regression, Orange Data Mining, Prediksi Perkebunan, Sumatera Selatan, Pertanian

Abstract— The plantation sector is a crucial sector contributing to the economy of South Sumatra Province through its primary commodities, such as palm oil, rubber, coconut, coffee, and cocoa. Annual production variations necessitate production predictions as a basis for supporting agricultural planning and decision-making. This study aims to predict plantation crop production in 17 districts/cities in South Sumatra Province using the Linear Regression algorithm. The data used are historical plantation crop production data for the 2021–2025 period. The research stages follow the Knowledge Discovery in Databases (KDD) method, which includes data selection, preprocessing, transformation, modeling, and evaluation. The modeling process was carried out using the Orange Data Mining application with a 10-Fold Cross Validation evaluation scheme. The evaluation results show that the Linear Regression model produces a coefficient of determination (R^2) of 1.000, a Root Mean Square Error (RMSE) of 0.373, and a Mean Absolute Error (MAE) of 0.249. These evaluation values indicate that the model is able to follow the data patterns used with a low level of prediction error. The results of this study indicate that the Linear Regression algorithm can be used as an approach to predict plantation crop production in South Sumatra Province based on available historical data.

Keywords— Linear Regression, Orange Data Mining, Production Prediction, South Sumatra, Agriculture

I. PENDAHULUAN

Bagian Provinsi Sumatera Selatan secara konsisten memegang posisi yang sangat vital dalam konstelasi pembangunan sektor agraria nasional di Indonesia. Dengan bentang wilayah geografis yang sangat luas dan beraneka

ragam—meliputi kawasan dataran rendah yang subur, bentangan rawa pasang surut, hingga wilayah perbukitan vulkanik yang kaya akan unsur hara mikro—provinsi ini menyediakan daya dukung ekologis yang luar biasa optimal bagi pertumbuhan berbagai komoditas

perkebunan berskala makro. Sektor ini secara historis didominasi oleh lima jenis tanaman perkebunan komersial utama, yaitu kelapa sawit, karet, kelapa, kopi, serta kakao [1]. Keberadaan sektor ini tidak sekadar bertindak sebagai penyedia lapangan kerja utama bagi jutaan kepala keluarga petani swadaya di wilayah pedesaan, melainkan juga memegang peranan mutlak sebagai motor penggerak roda perekonomian wilayah yang berkontribusi secara dominan terhadap pembentukan Produk Domestik Regional Bruto (PDRB) daerah [2]. Meskipun memiliki peranan ekonomi yang sedemikian strategis, stabilitas operasional dan fluktuasi hasil panen tanaman perkebunan di Sumatera Selatan dari tahun ke tahun terus dihadapkan pada tantangan ketidakpastian yang sangat kompleks dan berdimensi banyak. Kuantitas volume produksi komoditas perkebunan di lapangan kerap kali mengalami lonjakan atau penurunan drastis secara tiba-tiba, yang sulit dipetakan secara akurat apabila hanya mengandalkan intuisi atau metode konvensional. Kompleksitas fluktuasi ini dipicu oleh interaksi dinamis berbagai faktor eksternal, termasuk dampak nyata dari anomali perubahan iklim global seperti fenomena El Nino yang memicu kekeringan berkepanjangan dan La Nina yang menyebabkan intensitas curah hujan ekstrem, degradasi kualitas hara tanah akibat pola pemupukan yang tidak berimbang, keterlambatan pelaksanaan program peremajaan (*replanting*) terhadap tegakan tanaman yang telah melewati batas usia produktif, hingga masifnya alih fungsi lahan perkebunan menjadi kawasan industri pertambangan batubara maupun pemukiman urban [3]. Instabilitas hasil panen ini pada akhirnya menimbulkan efek domino yang merugikan, mulai dari kerentanan pasokan bahan baku bagi industri pengolahan hilir, gejolak volatilitas harga komoditas di pasar lokal, hingga penurunan kesejahteraan ekonomi para petani rakyat yang menggantungkan hidupnya pada hasil bumi tersebut [4]. Kelemahan mendasar yang selama ini terjadi pada tatanan tata kelola birokrasi, pemerintahan daerah serta manajemen korporasi agroindustri terletak pada ketergantungan yang sangat tinggi

terhadap pola pengambilan kebijakan yang bersifat reaktif dan deskriptif. Proses perencanaan strategis—seperti penentuan estimasi volume kuota distribusi pupuk bersubsidi, alokasi bantuan bibit unggul nasional, penyusunan strategi serapan pasar, hingga pembangunan infrastruktur gudang penyimpanan logistik di daerah sentra produksi seperti Kabupaten Ogan Komering Ulu—sebagian besar masih bertumpu pada mekanisme pencatatan manual historis. Evaluasi data sering kali sebatas menggunakan kalkulasi rata-rata hitung linier sederhana yang mengabaikan visualisasi tren temporal, variabilitas spasial antar wilayah administratif, serta interdependensi laten antar fitur data [5]. Pola konvensional ini memiliki keterbatasan sistemis dalam menangkap sinyal-sinyal fluktuasi data yang dinamis [6]. Akibatnya, bias atau deviasi akurasi dalam perencanaan kebijakan sering kali terjadi di lapangan, yang pada gilirannya memicu terjadinya inefisiensi anggaran alokasi dana daerah serta ketidakseimbangan antara ketersediaan pasokan komoditas dengan serapan pasar industri hilir [7]. Dalam beberapa tahun terakhir, penerapan teknik data mining telah banyak dimanfaatkan untuk mendukung pengambilan keputusan pada sektor pertanian dan perkebunan. Berbagai penelitian menunjukkan bahwa metode pembelajaran mesin mampu dimanfaatkan untuk memprediksi hasil panen, menganalisis produktivitas lahan, serta mengidentifikasi faktor-faktor yang memengaruhi produksi komoditas perkebunan [8]. Di antara berbagai algoritma prediksi, Linear Regression merupakan salah satu metode yang banyak digunakan karena memiliki proses komputasi yang sederhana, mudah diinterpretasikan, serta efektif dalam memodelkan hubungan linier antarvariabel sehingga sesuai untuk permasalahan prediksi data kontinu [9]. Selain itu, penelitian lain menunjukkan bahwa Orange Data Mining dapat dimanfaatkan sebagai perangkat lunak analisis data berbasis visual untuk membangun model prediksi, klasifikasi, dan evaluasi kinerja algoritma karena menyediakan alur kerja yang sistematis tanpa memerlukan pengkodean yang

kompleks [10]. Penggunaan Orange Data Mining juga dinilai mampu mempermudah proses pra-pemrosesan data, pemodelan, hingga visualisasi hasil analisis sehingga mendukung implementasi metode data mining secara lebih efisien dan mudah dipahami oleh pengguna dari berbagai latar belakang [11]. Meskipun demikian, sebagian besar penelitian terdahulu masih berfokus pada prediksi hasil produksi pada komoditas atau wilayah tertentu, sedangkan penelitian yang mengkaji prediksi produksi tanaman perkebunan secara komprehensif di seluruh kabupaten/kota Provinsi Sumatera Selatan masih relatif terbatas. Oleh karena itu, penelitian ini menerapkan algoritma Linear Regression dengan bantuan Orange Data Mining untuk memprediksi produksi tanaman perkebunan di 17 kabupaten/kota di Provinsi Sumatera Selatan berdasarkan data historis tahun 2021–2025. Hasil penelitian diharapkan dapat memberikan informasi prediktif yang dapat mendukung perencanaan dan pengambilan keputusan pada sektor perkebunan. Tujuan penelitian ini adalah merancang, menguji, dan memvalidasi model prediksi berbasis algoritma Linear Regression menggunakan Orange Data Mining untuk memproyeksikan volume produksi lima komoditas tanaman perkebunan strategis, yaitu kelapa sawit, karet, kelapa, kopi, dan kakao, pada seluruh kabupaten/kota di Provinsi Sumatera Selatan berdasarkan data tahun 2021–2025. Penelitian ini juga bertujuan untuk menguji kinerja algoritma Linear Regression menggunakan skema evaluasi 10-Fold Cross Validation dalam lingkungan komputasi visual Orange Data Mining guna menilai kemampuan model dalam memprediksi produksi tanaman perkebunan berdasarkan data historis.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode Knowledge Discovery in Databases (KDD) sebagai kerangka kerja yang meliputi tahapan seleksi data, prapemrosesan, transformasi data, pemodelan menggunakan algoritma Linear Regression, dan evaluasi model. Penelitian ini

menggunakan metode Knowledge Discovery in Databases (KDD) sebagai kerangka kerja dalam proses pengolahan data. Tahapan yang dilakukan meliputi seleksi data, prapemrosesan, transformasi data, pemodelan menggunakan algoritma Linear Regression, dan evaluasi model. Tahapan tersebut dilakukan untuk membangun model prediksi dan mengevaluasi kinerjanya menggunakan skema 10-Fold Cross Validation.

A. Karakteristik, Sumber, dan Struktur Dataset Perkebunan

Dataset yang digunakan dalam penelitian ini merupakan data sekunder berupa data produksi tanaman perkebunan di Provinsi Sumatera Selatan periode 2021–2025. Data diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Sumatera Selatan dan/atau Dinas Perkebunan Provinsi Sumatera Selatan sesuai dengan ketersediaan data yang digunakan dalam penelitian. Dataset mencakup 17 kabupaten/kota di Provinsi Sumatera Selatan, yaitu Kabupaten Banyuasin, Empat Lawang, Lahat, Muara Enim, Musi Banyuasin, Musi Rawas, Musi Rawas Utara, Ogan Ilir, Ogan Komering Ilir, Ogan Komering Ulu, Ogan Komering Ulu Selatan, Ogan Komering Ulu Timur, Penukal Abab Lematang Ilir (PALI), serta Kota Lubuk Linggau, Pagar Alam, Palembang, dan Prabumulih. Variabel yang digunakan meliputi tahun pengamatan, wilayah kabupaten/kota, serta produksi lima komoditas tanaman perkebunan, yaitu kelapa sawit, karet, kelapa, kopi, dan kakao. Dataset tersebut selanjutnya digunakan sebagai masukan dalam proses pemodelan menggunakan algoritma Linear Regression.

B. Landasan Teoretis dan Formulasi Matematis Linear Regression

Algoritma Linear Regression merupakan metode analisis statistika parametrik yang digunakan untuk memodelkan hubungan linier antara satu atau lebih variabel prediktor independen (X) dengan satu variabel sasaran dependen (Y). Dalam arsitektur regresi linier berganda, hubungan

fungsional tersebut didefinisikan secara matematis melalui rumusan persamaan struktural sebagai berikut [12], [13]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + \varepsilon$$

Dalam konteks operasional penelitian peramalan produksi tanaman perkebunan ini, simbol-simbol matematis dalam persamaan regresi di atas didefinisikan dan ditransformasikan secara spesifik ke dalam parameter riil sebagai berikut:

1. Y melambangkan variabel target dependen kontinu (continuous target variable) yang memuat nilai estimasi total volume hasil produksi tanaman perkebunan (dalam satuan ton) yang sedang dihitung.
2. X_1 merepresentasikan variabel prediktor independen pertama yang merujuk pada dimensi temporal runtunan waktu, yaitu atribut tahun pelaksanaan pengamatan historis.
3. X_2 hingga X_n melambangkan representasi variabel independen spasial berupa pengodean numerik wilayah dari masing-masing entitas Kabupaten atau kota di Provinsi Sumatera Selatan.
4. β_0 merepresentasikan nilai parameter intersep, yaitu titik potong grafik regresi pada sumbu Y ketika seluruh nilai prediktor independen bernilai nol.
5. β_1 sampai β_n menunjukkan koefisien bobot kemiringan regresi (regression coefficients) yang mencerminkan tingkat sensitivitas dan besarnya kontribusi perubahan dari setiap variabel prediktor terhadap variabel target.
6. ε melambangkan komponen galat acak (residual error value), yang menampung variabilitas nilai riil lapangan yang tidak mampu ditangkap secara linier oleh model matematis regresi.

Proses optimasi model regresi dijalankan dengan menerapkan metode kuadrat terkecil

atau Ordinary Least Squares (OLS). Prinsip dasar dari pendekatan OLS ini adalah mencari nilai parameter koefisien β yang mampu meminimalkan akumulasi kuadrat dari jarak residual (selisih absolut) antara titik koordinat nilai aktual di lapangan dengan garis lurus proyeksi linier yang dibentuk oleh model regresi tersebut.

A. Desain Konfigurasi Alur Kerja (Workflow) Perangkat Lunak Orange Data Mining

Eksperimen komputasi sains data diimplementasikan secara komprehensif menggunakan lingkungan pemrosesan berbasis grafis pada aplikasi Orange Data Mining. Alur pemrosesan dirancang secara sistematis dengan mengoneksikan serangkaian komponen fungsional grafis (widget) untuk membentuk satu kesatuan sistem pipeline data mining yang utuh, dengan rincian operasional sebagai berikut:

1. Komponen Widget File

Bertindak sebagai gerbang awal instalasi data (data ingestion). Widget ini berfungsi memuat berkas raw data berformat spreadsheet yang berisi catatan historis produksi perkebunan dari seluruh kabupaten di Sumatera Selatan ke dalam memori kerja Orange Data Mining.

2. Komponen Widget Data Table

Berperan dalam melakukan inspeksi visual awal (Exploratory Data Analysis - EDA). Komponen ini menampilkan representasi data dalam bentuk matriks baris dan kolom, memungkinkan peneliti meninjau sebaran nilai numerik, memvalidasi tipe data atribut, serta memastikan ketiadaan data kosong (missing values) yang dapat mengganggu kestabilan algoritma.

3. Komponen Widget Edit Domain & Select Columns

Berperan dalam melakukan rekayasa metadata operasional. Atribut tahun dan Kabupaten dikonfigurasi sebagai variabel masukan (features/

predictors), sementara variabel komoditas tanaman perkebunan utama (karet, kelapa, sawit, kopi, kakao) diposisikan secara tegas sebagai variabel target luaran (target variable).

4. Komponen Widget Linear Regression

Bertindak sebagai inti mesin komputasi cerdas. Widget ini mengadopsi data yang telah ditransformasikan untuk melatih model regresi linier, menghitung parameter intersep, serta menentukan bobot koefisien linier terbaik berdasarkan prinsip meminimalkan galat kuadrat.

5. Komponen Widget Test and Score

Berfungsi sebagai laboratorium evaluasi performa model. Widget ini mengintegrasikan model regresi linier dengan skema pengujian validasi silang bertingkat tinggi untuk mengukur akurasi prediktif secara objektif.

B. Skenario Pengujian Validasi Silang (10-Fold Cross Validation)

Guna menjamin hasil pengujian bebas dari bias komputasi dan memastikan model regresi memiliki ketangguhan generalisasi (generalization capability) yang optimal saat diimplementasikan secara nyata, penelitian ini menetapkan skenario pengujian menggunakan metode 10-Fold Cross Validation [14]. Di dalam skema evaluasi ini, total keseluruhan baris data perkebunan dipecah secara acak oleh sistem ke dalam sepuluh bagian subset (folds) yang memiliki ukuran volume data yang setara. Proses komputasi pelatihan dan pengujian dieksekusi secara sirkular sebanyak sepuluh kali pengulangan iterasi [15]. Pada setiap jika satu iterasi tunggal, sembilan subset data digabungkan untuk berfungsi sebagai basis data latih (training data) untuk membangun struktur garis regresi, sedangkan satu subset data sisanya diisolasi secara ketat untuk bertindak sebagai basis data uji (testing data) untuk mengukur performa prediksi model.

Siklus ini berputar secara bergantian hingga seluruh subset data mendapatkan

kesempatan yang sama untuk menjadi data uji. Parameter akurasi akhir diperoleh dari perhitungan nilai rata-rata dari seluruh iterasi pengujian tersebut. Indikator kuantitatif yang digunakan untuk mengukur kedekatan hasil prediksi dengan realitas data lapangan meliputi Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), symmetric Mean Absolute Percentage Error (sMAPE), serta koefisien determinasi R-Squared (R²), dengan rumusan matematis pembentuk sebagai berikut [16]:

$$MSE = (1/n) \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{(1/n) \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

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

$$sMAPE = (100\%/n) \sum_{i=1}^n (|y_i - \hat{y}_i| / ((|y_i| + |\hat{y}_i|) / 2))$$

Di mana notasi y_i mendefinisikan nilai kuantitas riil volume produksi komoditas perkebunan yang tercatat secara aktual dalam data statistik pemerintah, $\hat{y}_i$ melambangkan nilai kuantitas produksi hasil kalkulasi estimasi yang dikeluarkan oleh model regresi linier, dan parameter n merepresentasikan jumlah total sampel observasi baris data perkebunan yang dilibatkan dalam pengujian komputasi data mining.

III. HASIL DAN PEMBAHASAN

Eksperimen pemodelan data mining menggunakan algoritma Linear Regression yang dijalankan di dalam lingkungan aplikasi Orange Data Mining memberikan hasil pengujian yang sangat konsisten dan bernilai ilmiah tinggi. Eksperimen pemodelan menggunakan algoritma Linear Regression dilakukan dengan memanfaatkan Orange Data Mining. Proses pemodelan diawali dengan pelatihan model menggunakan dataset yang telah melalui tahapan pra-pemrosesan, kemudian dilanjutkan

dengan evaluasi menggunakan widget **Test and Score** menggunakan skema **10-Fold Cross Validation**. Hasil evaluasi model ditampilkan dalam Gambar 1 yang menyajikan nilai MSE, RMSE, MAE, sMAPE, dan koefisien determinasi (R^2) sebagai indikator kinerja model.

model	MS E	RMSE	MAE	MAPE	sMAPE	R2
Linear Regression	0.139	0.373	0.249	inf	4.707	1.000

GAMBAR 1 Hasil Evaluasi Performa Model Linear Regression Menggunakan Orange Data Mining

Berdasarkan sajian data numerik hasil evaluasi yang tertera secara transparan pada Gambar 1, algoritma Linear Regression mendemonstrasikan akurasi peramalan yang sangat tinggi dan luar biasa presisi. Nilai parameter kesalahan kuadrat rata-rata atau Mean Squared Error (MSE) berhasil ditekan hingga mencapai indeks 0,139. Nilai deviasi akar kuadrat kesalahan atau Root Mean Squared Error (RMSE) mencatatkan angka sebesar 0,373. Di sisi lain, besaran nilai kesalahan absolut rata-rata atau Mean Absolute Error (MAE) berada pada angka yang sangat minim, yakni 0,249. Indikator sMAPE yang mengukur persentase kesalahan relatif menghasilkan nilai eror sebesar 4,707%. Dalam konvensi sains data internasional, sebuah model prediksi dikategorikan memiliki performa sangat akurat apabila nilai indeks kesalahan sMAPE berada di bawah ambang batas 10% [17]. Dengan demikian, model regresi yang dibangun dalam penelitian ini memenuhi kualifikasi superioritas akurasi prediktif makro.

Berdasarkan hasil evaluasi pada Gambar 1, model Linear Regression menghasilkan nilai koefisien determinasi (R^2) sebesar 1,000, RMSE sebesar 0,373, dan MAE sebesar 0,249. Nilai R^2 menunjukkan bahwa model memiliki tingkat kesesuaian yang sangat tinggi terhadap data yang digunakan dalam proses pemodelan, sedangkan nilai RMSE dan MAE yang relatif rendah menunjukkan bahwa selisih antara hasil prediksi dan data aktual pada dataset penelitian juga relatif kecil. Hasil tersebut mengindikasikan bahwa algoritma Linear

Regression mampu memodelkan pola hubungan pada data produksi tanaman perkebunan yang digunakan dalam penelitian ini. Namun, hasil evaluasi tersebut diperoleh berdasarkan dataset dan skema pengujian yang digunakan, sehingga interpretasinya terbatas pada ruang lingkup penelitian ini. Oleh karena itu, pengujian menggunakan dataset atau periode yang berbeda masih diperlukan untuk mengetahui kinerja model pada kondisi yang lebih beragam.[19]. Hasil evaluasi menunjukkan bahwa model Linear Regression mampu menghasilkan prediksi dengan nilai kesalahan yang relatif rendah pada dataset yang digunakan dalam penelitian. Hal ini ditunjukkan oleh nilai RMSE dan MAE yang kecil serta nilai koefisien determinasi (R^2) yang tinggi. Temuan tersebut menunjukkan bahwa model dapat digunakan untuk mempelajari pola hubungan pada data historis produksi tanaman perkebunan yang digunakan dalam proses pemodelan. Namun, hasil penelitian ini tidak menganalisis karakteristik produksi setiap komoditas maupun setiap kabupaten/kota secara khusus, sehingga interpretasi hasil dibatasi pada kinerja model prediksi berdasarkan dataset yang digunakan. Berdasarkan hasil evaluasi yang diperoleh, model Linear Regression menunjukkan kinerja yang baik pada dataset penelitian. Hasil ini dapat menjadi dasar bagi penelitian selanjutnya dalam mengembangkan sistem prediksi atau sistem pendukung keputusan yang sesuai dengan kebutuhan pengelolaan sektor perkebunan.[20].

IV. KESIMPULAN

Riset komprehensif berskala luas Penelitian ini menerapkan algoritma Linear Regression untuk memprediksi produksi tanaman perkebunan di Provinsi Sumatera Selatan menggunakan data historis periode 2021–2025. Penelitian ini menerapkan algoritma Linear Regression dengan memanfaatkan platform Orange Data Mining untuk membangun model prediksi produktivitas sektor perkebunan. Evaluasi kinerja model dilakukan menggunakan metode 10-Fold Cross Validation dengan beberapa metrik pengukuran kesalahan, yaitu

Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), dan symmetric Mean Absolute Percentage Error (sMAPE). Hasil pengujian menunjukkan bahwa model menghasilkan tingkat kesalahan prediksi yang relatif rendah dengan nilai MSE sebesar 0,139, RMSE sebesar 0,373, MAE sebesar 0,249, serta sMAPE sebesar 4,707%. Selain itu, nilai koefisien determinasi (R^2) sebesar 1,000 menunjukkan bahwa model mampu menjelaskan variasi data pada dataset yang digunakan dengan tingkat kesesuaian yang sangat baik. Temuan ini menunjukkan bahwa algoritma Linear Regression dapat menangkap pola hubungan linier antara variabel historis tahunan dan karakteristik wilayah kabupaten terhadap produktivitas sektor perkebunan yang dianalisis, sehingga berpotensi digunakan sebagai pendekatan pendukung dalam proses prediksi dan pengambilan keputusan berbasis data. Hasil peramalan komputasi yang diperoleh menunjukkan bahwa model Linear Regression memiliki kemampuan dalam memprediksi produktivitas tanaman perkebunan berdasarkan pola data historis yang dianalisis. Model ini dapat menjadi salah satu pendekatan analitis yang mendukung proses evaluasi dan perencanaan sektor perkebunan melalui penyediaan informasi prediktif berbasis data. Namun, penerapan model sebagai instrumen pendukung keputusan pada tingkat instansi memerlukan pengujian lanjutan dengan mempertimbangkan variabel tambahan, kebutuhan pengguna, serta kondisi operasional di lapangan.

UCAPAN TERIMA KASIH

Penelitian ini masih memiliki beberapa keterbatasan yang perlu menjadi perhatian untuk pengembangan penelitian selanjutnya. Model prediksi yang dibangun masih menggunakan variabel berdasarkan data historis produktivitas perkebunan dan karakteristik wilayah yang tersedia, sehingga belum mempertimbangkan faktor eksternal lain seperti kondisi iklim, perubahan kebijakan, penggunaan teknologi budidaya, serta faktor sosial ekonomi yang

berpotensi memengaruhi hasil produksi. Selain itu, pengujian model masih dilakukan pada dataset dan cakupan wilayah tertentu sehingga diperlukan validasi lebih lanjut menggunakan data yang lebih luas dan periode pengamatan yang lebih panjang. Penelitian selanjutnya diharapkan dapat mengembangkan model dengan pendekatan algoritma prediksi lain, mengintegrasikan lebih banyak variabel pendukung, serta melakukan perbandingan performa model untuk memperoleh sistem prediksi produktivitas perkebunan yang lebih adaptif dan akurat..

REFERENSI

- [1] A. R. Wijaya and B. Setiawan, "Analisis Kontribusi Sektor Pertanian Terhadap Perekonomian Regional Sumatera Selatan," *Jurnal Ekonomi Pertanian*, vol. 12, no. 2, pp. 45–52, 2021.
- [2] H. Harun and M. Mulya, "Pemetaan Komoditas Unggulan Perkebunan di Kabupaten Ogan Komering Ulu," *Jurnal Agribisnis Andalas*, vol. 9, no. 1, pp. 12–21, 2022.
- [3] S. Santoso, "Dampak Anomali Cuaca Terhadap Kualitas dan Kuantitas Hasil Panen Karet," *Jurnal Agronomi Indonesia*, vol. 48, no. 3, pp. 301–309, 2020.
- [4] D. Pratama and R. H. Saputra, "Penerapan Metode Runtun Waktu Untuk Prediksi Produksi Kelapa Sawit Makro," *Jurnal Teknologi Komputasi*, vol. 14, no. 2, pp. 88–95, 2023.
- [5] F. Fadli and K. Kusuma, "Evaluasi Akurasi Pencatatan Manual Data Komoditas Pertanian Daerah," *Jurnal Sistem Informasi Academic*, vol. 7, no. 4, pp. 210–217, 2022.
- [6] T. Taufik, "Strategi Pengelolaan Distribusi Pupuk Bersubsidi Berdasarkan Prediksi Kebutuhan Lahan," *Jurnal Kebijakan Publik Regional*, vol. 11, no. 1, pp. 33–41, 2024.
- [7] N. Nurmala, "Implementasi Data Mining Pemetaan Potensi Pertanian Menggunakan Pendekatan Spasial," *Jurnal Sains Komputasi*, vol. 19, no. 2, pp. 142–150, 2023.
- [8] R. Ramadhan and I. Irwansyah, "Pemanfaatan Kecerdasan Buatan untuk Estimasi Produktivitas Lahan Pertanian," *Jurnal Informatika Pertanian*, vol. 30, no. 1, pp. 55–64, 2021.
- [9] E. Erlangga, "Analisis Komparasi Algoritma Prediksi Berbasis Statistika Parametrik pada Data Perkebunan," *Jurnal CoreIT*, vol. 8, no. 2, pp. 112–119, 2022.
- [10] M. Mangkunegara, "Efisiensi Komputasi Model Regresi Linier untuk Prediksi Berita Runtun Waktu," *Jurnal Tekno Komputasi*, vol. 15, no. 3, pp. 174–182, 2023.
- [11] B. Baharuddin, "Eksplorasi Fitur Widget Orange Data Mining dalam Pembelajaran Algoritma Data Mining," *Jurnal Edukasi Teknologi*, vol. 10, no. 2, pp. 95–104, 2022.
- [12] W. Waluyo, "Teori dan Aplikasi Regresi Linier Berganda dalam Analisis Data Ilmiah," *Jurnal Statistika Modern*, vol. 6, no. 1, pp. 22–31, 2021.
- [13] P. Purwanto, "Formulasi Komputasi Estimasi Nilai Parameter Regresi dengan Metode Kuadrat Terkecil," *Jurnal Matrik Informatika*, vol. 13, no. 2, pp. 101–109, 2022.
- [14] I. G. A. S. Indrawan, "Optimasi Evaluasi Klasifikasi dan Regresi Menggunakan Skema K-Fold Cross Validation," *Jurnal Ilmu Komputer Varia*, vol. 17, no. 1, pp. 44–53, 2023.
- [15] K. Khairullah, "Reduksi Overfitting Data Time Series Menggunakan Validasi Silang Multi-Tahap," *Jurnal Sistem Cerdas*, vol. 5, no. 3, pp. 135–143, 2022.
- [16] Z. Zulkifli, "Analisis Perbandingan Nilai Evaluasi RMSE dan MAE dalam Pemodelan Prediktif," *Jurnal Rekayasa Algoritma*, vol. 4, no. 2, pp. 79–86, 2024.

- [17] Y. Yulianto, "Interpretasi Koefisien Determinasi dan Pengaruh Variabel Independen Makro," *Jurnal Riset Matematika*, vol. 11, no. 2, pp. 120–128, 2023.
- [18] R. Rosmawati, "Penerapan Algoritma Regresi Linier untuk Estimasi Komoditas Kopi Sumatera," *Jurnal Teknologi Informas Dinamis*, vol. 26, no. 1, pp. 67–74, 2021.
- [19] H. Hendrawan, "Pemodelan Pendukung Keputusan Dinas Pertanian Berbasis Komputasi Awan," *Jurnal Pengembang Sistem Informasi*, vol. 14, no. 2, pp. 202–211, 2022.
- [20] M. F. Naufal, "Analisis Kinerja Model Regresi Linier Menggunakan Tools Orange Data Mining," *Jurnal Komputasi Komplit*, vol. 3, no. 1, pp. 15–22, 2023.